Plant proteolytic breakthroughs: determining metacaspases functions from substrate identification to activity visualization.

Álvaro Daniel Fernández Fernández

Academic Year: **2020-2021**

Ghent University

Faculty of Sciences

Department of Plant Biotechnology and Bioinformatics

Plant proteolytic breakthroughs: determining metacaspases functions from substrate identification to activity visualization.

Álvaro Daniel Fernández Fernández

Thesis submitted as partial fulfilment for the doctor degree (PhD) in Sciences, Biochemistry and Biotechnology.

Academic Year: **2020-2021**

Promoter

Prof. Dr Frank Van Breusegem

Faculty of Sciences, Department of Plant Biotechnology and Bioinformatics, Ghent University, Belgium

Co-Promoter

Dr. Simon Stael

Faculty of Sciences, Department of Plant Biotechnology and Bioinformatics, Ghent University, Belgium

On the cover: Capture of a confocal image of *Nicotiana benthamiana* leaf transiently co-expressing pUBI10::ZipGFP$^{1-10}$ and pUBI10::ZipGFP^{11} and p35S::NES-TEVp-tagBFP.

On the back-cover: Capture of a Z-Stack maximal projection of a CPR line expressed in *Arabidopsis thaliana*.

Examination committee

Prof. Dr. Sofie Goormachtig (Chair)

Faculty of Sciences, Department of Plant Biotechnology and Bioinformatics, Ghent University, Belgium

Prof. Dr. Frank Van Breusegem (Promoter)

Faculty of Sciences, Department of Plant Biotechnology and Bioinformatics, Ghent University, Belgium

Dr. Simon Stael (Co-Promoter)

Faculty of Sciences, Department of Plant Biotechnology and Bioinformatics, Ghent University, Belgium

Dr. Jefri Heyman (Secretary)

Faculty of Sciences, Department of Plant Biotechnology and Bioinformatics, Ghent University, Belgium

Prof. Dr. Danny Geelen

Faculty of Bioscience Engineering, Department of Plants and Crops, Ghent University, Belgium

Prof. Dr. Jens Staal

Faculty of Sciences, Department of Biochemistry and microbiology and Department of Biomedical molecular biology, Ghent University, Belgium

Prof. Dr. Kris Gevaert

Faculty of Medicine and Health Sciences, Department of Biomolecular Medicine, Ghent University, Belgium

Dr. Núria Sánchez Coll

CSIC Scientist, Centre for Research in Agricultural Genomics, Barcelona, Spain

Table of Contents

Summary 6

Overzicht 8

List of Abbreviations 10

Introduction 13

Chapter 1: “Introduction to plant metacaspases” 15

Chapter 2: “Proteolytic activation mechanisms in plants” 33

Chapter 3: “Caught green-handed: methods for in vivo detection and visualization of protease activity” 61

Research results 101

Chapter 4: “Damage on plants activates Ca^{2+}-dependent metacaspases for release of immuno-modulatory peptides” 103

Chapter 5: “Identification of molecular partners and substrates of AtMC4 during wounding stress” 149

Chapter 6: “Genetically encoded fluorescent reporters for the detection of proteolysis in plants” 209

Discussion and perspectives 247

Acknowledgements 255

Curriculum vitae 257

Publications 259

Summary

Proteases, proteinases and peptidases are all synonyms used to refer to enzymes performing hydrolysis of the amino acid chains of their substrates in different fashions. While proteases are commonly used in a research environments to digest proteins to peptides or to stop the function of other enzymes, the contribution of most proteases for the functioning of the whole organisms in their *in vivo* environment remain undiscovered. Surprisingly, proteases happen to be a good fraction of the total genome encoding for proteins. Proteases are found in every living organism and they are even present in most types of viruses. Proteolytic enzymes are characterized by their capacity to perform a nucleophilic attack over the carbon of the scissile bond, leading to its hydrolysis and subsequent disruption of the amino acid chain. Thereby, proteolysis as a post-transcriptional modification is capable to transform the protein composition, and consequently this may lead to variation of its subcellular localization, adjustments in the protein lifetime and modulation of the substrate functions. Traditionally, proteases have been perceived as main components of protein degradation. It is at least paradoxical that most studies at proteome level focus in other PTMs, but that approaches to study proteases that can modulate part of the proteome are not exactly abundant.

They are highly controlled by the cellular environment in their surrounds from their enzymatic induction to control, modulation and depletion of their functioning. In plants, proteases are important for many processes from development to responses to stresses and pathogens. Metacaspases are probably one of the most studied proteases in plants.

In this PhD thesis I tackle the enzymatic activity of the plant proteases focusing in the enzymatic functions of AtMC4 in biological relevant conditions. The PhD thesis is organized in 3 introductory chapters and 3 research chapters (Fig. 1). In the first chapter I introduced metacaspases and centred in their functions in the plant field. In the second chapter, I have updated the latest discoveries in the plant field in relation to plant protease mode of action. I have focused on reviewing the activation mechanisms and substrate specificity of well-studied proteases, and I have also commented on bacterial proteases during infection working as effectors. In the third chapter I have reviewed the current mechanisms used to visualize proteolytic activity. While most of these sensors have been developed, I aimed to approach these genetic tools to the plant field. The fourth chapter constitute a study on the calcium activation of *Arabidopsis thaliana* METACASPASE 4 (AtMC4) in wounding and its

subsequent cleavage of PROPEP1. This study reports the first bona fide substrate for AtMC4 and opens new directions for metacaspases as possible components in peptide maturation of the PROPEP family, of high importance in plant immunity. The fifth chapter is a proteomic approach aiming to detect additional AtMC4 substrates in wounding events. This chapter combine the use of N-terminomics and shotgun proteomics in order to compare candidate substrates and total protein abundance in damage Arabidopsis plants. I successfully detected and validated *in vitro* and *in vivo* cleavage of different substrates and further efforts will elucidate the substantial contribution of the substrates in wounding stresses. In the sixth and last research chapter, proof-of-concept of the use of proteolytic genetically encoded sensors in plants is shown. The experiments include implementation in plants of ZipGFP, a reporter conceived in the animal field, but also two original sensors that report the activity by changes in localization or by FRET variations using a fluorescent protein pair not yet used in the detection of proteolytic activity. I have shown positive results for known proteases and report activity that correspond with the localization of metacaspases in Arabidopsis root and wounding events.

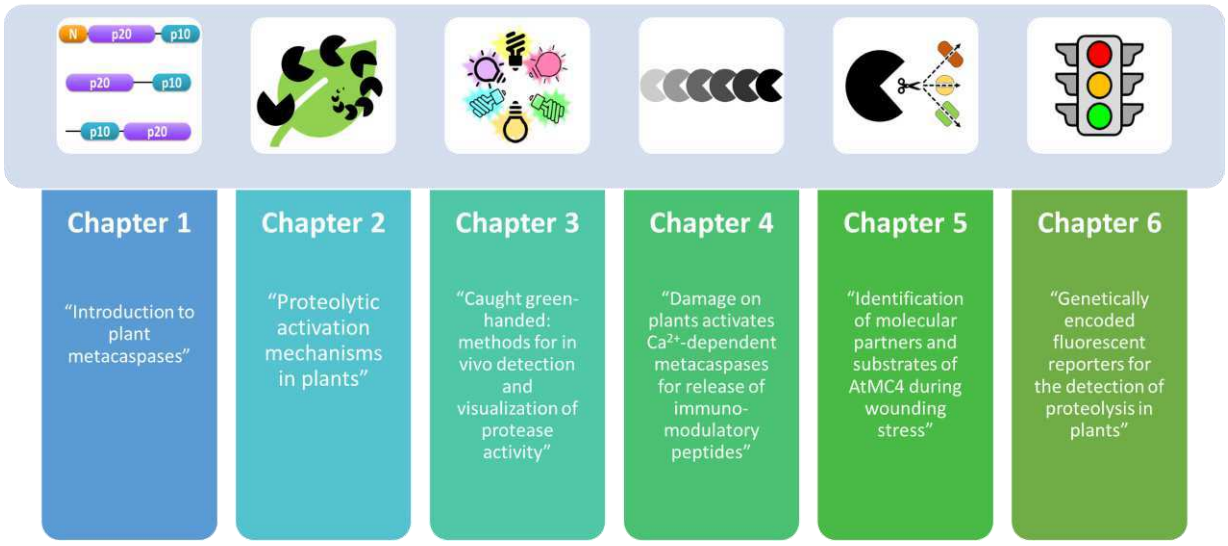


Figure1. Content overview of the chapters in this PhD thesis.

Figuur 1. Inhoudsoverzicht van de hoofdstukken in deze doctoraalscriptie

Overzicht

Proteasen, proteïnasen en peptidasen zijn synoniemen die worden gebruikt om enzymen aan te duiden die op verschillende manieren hydrolyse van de aminozuurketens van hun substraten uitvoeren. Hoewel proteasen algemeen worden gebruikt in een onderzoeksomgeving om eiwitten te verteren tot peptiden of om de functie van andere enzymen te stoppen, blijven de functies van de meeste proteasen in hun *in vivo* omgeving onontdekt. Verrassend genoeg maken proteasen een groot deel uit van het totale genoom van alle levende organismen en zijn zij ook aanwezig in de meeste soorten virussen. Hun efficiënte werkingsmechanisme berust op een nucleofiele aanval op de koolstof van de Peptide verbinding, die leidt tot de hydrolyse ervan en vervolgens tot het verbreken van de aminozuurketen. Daarbij is proteolyse als een post-transcriptionele modificatie in staat de eiwitsamenstelling te veranderen, en als gevolg daarvan kunnen deze veranderingen leiden tot re-lokalisatie, verschillen in de levensduur van de eiwitten en modulatie van de substraatfuncties. Het is op zijn minst paradoxaal dat er veel studies bestaan op het niveau van het proteoom om de functie van een gen te begrijpen, maar dat algemene studies over enzymen die een deel van het proteoom onafhankelijk van de translatie kunnen moduleren, niet echt overvloedig zijn.

Proteasen worden traditioneel gezien als de belangrijkste componenten van eiwitafbraak, maar zij worden in hoge mate gecontroleerd door de cellulaire omgeving, van hun activiteitsinductie tot latere controle van hun activiteit en depletie van de protease functie. In planten zijn proteasen belangrijk voor vele processen, van ontwikkeling tot reacties op stress en pathogenen. Metacaspases zijn waarschijnlijk één van de meest bestudeerde proteases in planten. In dit proefschrift hebben we ons gericht op de enzymatische activiteit van de plantproteasen die de enzymatische functies van AtMC4 in de hand werken. We hebben het proefschrift georganiseerd in 3 inleidende hoofdstukken en 3 onderzoekshoofdstukken (Fig. 1). In het eerste hoofdstuk heb ik metacaspasen geïntroduceerd en het centrum in hun functies in het plantenveld. In het tweede hoofdstuk heb ik de laatste ontdekkingen op plantengebied met betrekking tot het werkingsmechanisme van plantenproteasen bijgewerkt. We hebben ons gericht op het herzien van de activeringsmechanismen en substraatspecificiteit van goed bestudeerde proteasen, en we hebben ook een blik geworpen op bacteriële proteasen die van invloed zijn

tijdens infectie. In het deerde hoofdstuk hebben we de huidige mechanismen herzien die gebruikt worden om proteolytische activiteit te visualiseren. Terwijl de meeste van deze sensoren zijn ontwikkeld in dierlijke of humane studies, willen we de genetische hulpmiddelen naar het plantenveld vertalen. Het vierde hoofdstuk vormt een grondige studie van de calciumactivering van *Arabidopsis thaliana* METACASPASE 4 (AtMC4) bij verwonding en de daaropvolgende splitsing van PROPEP1. Deze studie rapporteert het eerste *bona fide* substraat voor AtMC4 en opent nieuwe richtingen voor metacaspases als mogelijke componenten in peptide maturatie van de PROPEP familie, van groot belang in plantimmuniteit. Het vijfde hoofdstuk is een proteomische benadering gericht op het detecteren van additionele AtMC4 substraten bij verwondingen. Dit hoofdstuk combineert het gebruik van N-terminomics en shotgun proteomics om kandidaat substraten en de totale eiwit abundantie in beschadigde Arabidopsis planten te vergelijken. We hebben met succes *in vitro* en *in vivo* splitsing van verschillende substraten gedetecteerd en gevalideerd en verdere inspanningen zullen de substantiële bijdrage van de substraten in schade veroorzakende stress ophelderen. In het zesde en laatste onderzoekshoofdstuk tonen we proof-of-concept voor het gebruik van proteolytische genetisch gecodeerde sensoren in planten. Onze experimenten omvatten de implementatie in planten van ZipGFP, een reporter die in dierlijke studies is bedacht, maar we kwamen ook met twee originele sensoren die de activiteit rapporteren door veranderingen in lokalisatie of door FRET variaties met behulp van een fluorescerend eiwitpaar dat nog niet gebruikt werd in de detectie van proteolytische activiteit. Wij hebben positieve resultaten laten zien voor bekende proteasen en rapporteren activiteit die overeenkomt met de lokalisatie van metacaspasen in Arabidopsis wortel en verwonding.

List of abbreviations

ABP: Activity-based Probes.

ATG4: Autophagy-related protein 4 (ATG4) cysteine proteases

AtLEG: Arabidopsis Legumain

AtPep1: Arabidopsis Elicitor Peptide 1

BAK1: BRASSINOSTEROID INSENSITIVE 1-associated receptor kinase 1

BRET: Bioluminescent Resonance Energy Transfer.

CA-GFP: Caspase Activatable GFP.

CFP: Cyan Fluorescent Protein.

CLEL: CLAVATA3 EMBRYO SURROUNDING REGION (CLE)-like family peptides (

COFRADIC: Combined Fractional Diagonal Chromatography

DEVD: minimal consensus cleavage site for caspase-3 activity composed by aspartate-glutamate-valine-aspartate, processed after the C-terminal aspartate of the motif.

EFR: Arabidopsis EF-Tu receptor

eGFP: Enhanced variant of GFP

EPI: Extracellular Protease Inhibitors

ER: Endoplasmic reticulum

FLS2: Flagellin Sensing 2

FRET: Förster Resonance Energy Transfer.

GAL4: Galactose 4 transcription factor.

GFP: Green fluorescent protein

GFP: Green Fluorescent Protein derived from jellyfish *Aequorea victoria*.

GLV1: GOLVEN1

GUS: β-glucuronidase

H_2O_2: Hydrogen peroxide

HCV: Hepatitis C Virus.

HIV: Human Immunodeficiency Virus.

IDA: Inflorescence Deficient in Abscission

JA: Jasmonic acid

kDa: kilodalton

MAPK: Mitogen Activated Protein Kinase

mRFP / RFP: monomeric Red Fluorescent Protein; Red Fluorescent Protein.

NLS: Nuclear Localization Signal.

PCD: Programmed Cell Death.

PEPR1 : Pep1 receptor 1

PEPR2: Pep1 receptor 2

PROPEP1: Precursor of Pep1

PSK: Phytosulfokine

PTM: post-translational modification

RALF: Rapid Alkalinization Factor

RLK: Receptor Like Kinase

SA: Salicylic acid

SCX: Strong Cation Exchange

SDS: Sodium Dodecyl Sulphate

SERPIN1: Serine Protease inhibitor

TEVp: Tobacco Etch Virus protease.

UAS: Upstream Activating Sequence.

UV-C: Ultraviolet C radiation which can act as germicidal.

VC3AI: Venus Caspase-3 Activation Indicator.

YFP: Yellow Fluorescent Protein with a T203Y mutation related to GFP. Improved versions of YFP include Citrine and Venus.

Introduction

Chapter 1

"Introduction to plant metacaspases".

Álvaro Daniel Fernández-Fernández[1,2], Simon Stael[1,2,] and Frank Van Breusegem[1,2]

[1] Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium

[2] Center for Plant Systems Biology, VIB, 9052 Ghent, Belgium

Author contributions:

ADFF wrote the manuscript with the help of SS and FVB. SS and FVB revised and corrected the manuscript.

Aim and context

This chapter serves as a guide to the less acquainted reader to metacaspases. The features associated to the different types of metacaspases are described as well as a comparison to other cysteine proteases. Also metacaspases are compared to their distant relative caspases in terms of activation and substrate affinity. In the last part, the current knowledge and the latest discoveries on plant metacaspases are treated in detail.

Abstract:

Metacaspases are a family of cysteine-dependent proteases present in fungi, plants and protists. The function of metacaspases across species are divergent and even a two metacaspase family members in an organism can conduct contrasting activities to each other. Here, we discuss the main differences of metacaspases to its orthologues caspases, and define the types of metacaspases according to their domain arrangement. Moreover, we focus on the function of metacasapases in plants and comment on their contribution for plant fitness.

Keywords: metacaspases, type-I, type-II, type-III, cysteine proteases.

Introduction

Proteases are enzymes with the capacity to hydrolyze peptidic bonds. Proteases are classified according to the amino acids performing their nucleophilic attack on the carbonyl group (Erez *et al.*, 2009). In plants, we can find aspartic, cysteine, serine, threonine or metalloproteases in case activation is performed by a metal ion (van der Hoorn, 2008; Lallemand *et al.*, 2015). Proteases are classified based on their structure resemblance in families and those, by homology to other families, into clans according to the MEROPS database http://www.ebi.ac.uk/merops (Rawlings *et al.*, 2018).

Metacaspases are a family of cysteine proteases originally identified by homology search to animal caspases (Uren *et al.*, 2000). Metacaspases and seven other families of cysteine endopeptidases are classified in the CD clan (Chen *et al.*, 1998; Cambra *et al.*, 2010). Generally, all the families in this clan contain a His-Cys catalytic dyad in this sequential order and they present hydrophobic amino acids preceding both active sites. Within the CD clan, metacaspases, caspases and paracapases are grouped in the C14 family with a further subdivision of caspases in the subfamily C14A and metacaspases and paracapases in the subfamily C14B (Rawlings *et al.*, 2010).

Metacaspases versus Caspases

The presence of metacaspases in eukaryotes is believed to have occurred by endosymbiosis of mitochondrial and plastids or by horizontal gene transfer from α-proteobacteria (Koonin and Aravind, 2002; Choi and Berges, 2013). Despite sharing the catalytic dyad and the structural similarity of the hemoglobinase fold, metacaspase's functions, substrate specificity and mode of action clearly differs from those of caspases (Salvesen *et al.*, 2016; Minina *et al.*, 2017). Both caspases and metacaspases show a strict requirement at the amino acid preceding the processed peptidyl bond. Caspases, as reflected in their name (*C*ysteine-dependent *ASP*artate-specific prote*ASE*), cleave their substrates after aspartic acid (Earnshaw *et al.*, 1999). Contrarily, metacaspases process their substrate following an arginine or a lysine residue and they cannot process capase-3 synthetic substrates (Vercammen *et al.*, 2004; Bozhkov *et al.*, 2005; Watanabe and Lam, 2011b). In fact screening of the substrates for metacaspases resulted in an optimal substrate VRPR↓X and DEVD↓X for caspases (Talanian *et al.*, 1997). It is noticeable the preference of metacaspases

for a double basic motif (XRXR) at the positions P1 and P3 before cleavage according to the Schechter and Berger nomenclature for amino acids in the cleavage zone (P4P3P2P1↓P1'P2'P3'P4') (Schechter and Berger, 1967). In addition, although both subfamilies display a hemoglobinase folding loop, small differences are found in their structure. For example, in metacaspases the structure is shaped by eight β-strands, while in caspases only six β-strands form the tri-dimensional arrangement (McLuskey and Mottram, 2015). Additionally, the activation of most metacaspases requires high levels of calcium (Gilio *et al.*, 2017; Hander *et al.*, 2019; Zhu *et al.*, 2020) with the exception of the pH-dependent type-II metacaspases, like Arabidopsis AtMC9 (Vercammen *et al.*, 2004) and AtMC9-like metacaspases in other plant species. Whereas most metacaspases have an optimal pH at neutral values, AtMC9 activity is most favourable around pH 5,5 and does not require calcium. Caspase-3, -6, -7 and -8 are optimally active at a pH value of 7-8 (Stennicke and Salvesen, 1997) but pH acidification is one of the footprints of the early events during apoptosis (Sergeeva *et al.*, 2017), meaning that caspases retain part of their activity in such environment. In fact, caspase activation is known to encompass a cascade where initiators like caspase-2, 8 ,9 and 10 process and activate the effector caspases (caspase-3, 6 and 7) that are directly responsible for the execution of cell death (Parrish *et al.*, 2013). Another difference with caspases is that, metacaspases function as monomers (McLuskey and Mottram, 2015; Zhu *et al.*, 2020), while caspases can shape dimers and in some cases can assemble into a superstructure designated apoptosome (Yuan *et al.*, 2010).

Along evolution, metacaspase genes have undergone several events of duplication, originating variation and diversification, which has resulted in a certain type of metacaspase specialization across species (Tsiatsiani *et al.*, 2011; McLuskey and Mottram, 2015). When comparing the distribution of the C14 proteases, metacaspases are present in plants, fungi, protists, archaea and bacteria; paracaspases can be found in animals, protists and fungi; and caspases are exclusively represented in metazoa (Fig. 1A).

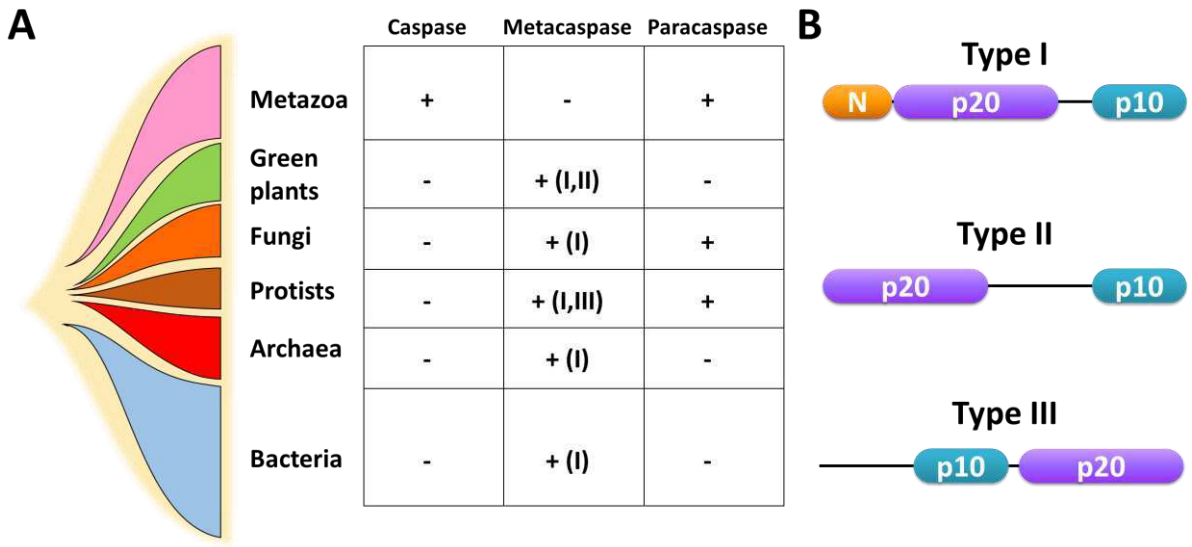


Figure 1. A) Presence of caspases, metacaspases, and paracaspases along different organisms. +/-: denotes presence and absence respectively. I, II and III denotes presence of type I, type II or type III metacaspases. B) Metacaspase classification according to their domain arrangement. (N: indicates N-terminal prodomain, p20: 20kDa subunit and p10: 10kDa subunit).

Caspases are well-known regulators of programmed cell death events in animals (Thornberry *et al.*, 1997). Although apoptosis does not exist in plants, metacaspases have been found to be involved in multiple processes involving cell death (Bozhkov *et al.*, 2005; He *et al.*, 2008; Coll *et al.*, 2010) and other stresses like pathogen response, or wounding (Coll *et al.*, 2014; Lema Asqui *et al.*, 2018; Hander *et al.*, 2019; Salguero-Linares and Coll, 2019).

Metacaspases are divided in type-I, type-II and type-III subclasses, depending on the structural arrangement of their large (p20) and small subunits (p10) (Fig. 1B). In type-I and type-II metacaspases the large subunit precedes the small subunit, while the order is inverted in type-III metacaspases (Choi and Berges, 2013; Klemencic and Funk, 2018). Type-I metacaspases bear a distinctive N-terminal prodomain with a proline-rich repeat region, a Zinc finger like motif and a relatively shorter linker between their domains compared to type-II metacaspases. The N-terminal domain of type-I metacaspases is important for its function. For example, the N-terminal parts are required for shaping insoluble aggregates in yeast (Lee *et al.*, 2010).

The distribution of the metacaspase types among different phylogenetic lineages is highly variable. For example, protozoa and fungi exclusively contain type-I metacaspases as seen in the green algae *Ostreococcus spp.* Other green algae

species like *Volvox carteri* and *Chlamydomonas reinhardtii* present one type-I and one type-II metacaspase. Land plants have both type-I and type-II metacaspases in different numbers (Tsiatsiani *et al.*, 2011; Minina *et al.*, 2017). For example, in the moss *Physcomitrella patens* there are 2 type-I and 5 type-II metacaspases and in the vascular plant *Selaginella moellendorffii* 3 type-I and 3 type-II metacaspases are present. In species like *Populus trichocarpa* the number of metacaspases can rise to 14 with 10 type-I and 5 type-II (Bollhoner *et al.*, 2018). While the *Arabidopsis thaliana* genome contains 3 type-I and 6 type-II metacaspases, its close relative *Arabidopsis lyrata* encodes for 5 type-I and 7 type-II. The genome duplication events occurred during speciation and how close species present different number of metacapase's types has not been studied but would render information in the specialization of the different subgroups.

Type-III metacaspases were discovered in algae and bacteria (Choi and Berges, 2013). Typically, they show an inversion in the order of the prodomains, p10 at the N-terminal side and p20 at the C-terminal part of the protein. They are by far the least studied metacaspases, with novel functions for the p10 domain. The type-III metacaspases in the flagellate cryptomonas algae *Guillardia theta* GtMC2, not only has the typical arginine and lysine processing specificity but also shows two calcium binding sites with different affinities that are lost in the metacaspases activated at low pH (Klemencic and Funk, 2018)

Metacaspases in plants

In Arabidopsis, the type-I metacaspases AtMC1 and AtMC2 function in an antagonistic manner during the hypersensitive response to bacterial pathogens (Coll *et al.*, 2010). AtMC1 is a positive regulator of a hypersensitive response, which is a form of programmed cell death. On the contrary, AtMC2 show opposite results to AtMC1 as shown by conductivity and number of cell death analysis using a salicylic acid analog or *Pseudomonas syringae* pt. tomato (Pst) DC3000 carrying avrRpm1. The regulator of cell death LESION STIMULATING DISEASE 1 (LSD1) (Dietrich *et al.*, 1994), strongly interacts with AtMC1 at a shared Zinc finger N-terminal motif. AtMC2 showed weak interaction to LSD1 and AtMC1.The antagonistic effects of AtMC1 and AtMC2 on cell death are enhanced when the prodomain sequence is absent; indicating that accessibility of their catalytic sites or proteolysis of their substrates affects the

hypersensitive response. Additionally, AtMC1 participates in clearance of aggregates during aging. AtMC1 accumulates in aggregates based on its N-termini in a parallel way to autophagy processes (Coll *et al.*, 2014). The other type-I metacaspase in Arabidopsis, AtMC3, remains unreported so far.

Type-II metacaspases lack a prodomain but instead they entail longer linker sequences between the p20 and the p10 subunits (Tsiatsiani *et al.*, 2011). Type-II metacaspases have an additional cysteine in the beginning of their p20 that can partially take over the activity of the typical catalytic cysteine (Belenghi *et al.*, 2007). An additional sub-classification among type-II metacaspases could be done by the mechanism contributing to their function whether it can be calcium rise (AtMC4-like) or pH drop (AtMC9-like) (Vercammen *et al.*, 2004). Modulation of the calcium or pH-dependent activation is determined by the linker sequence and their proteolytic activation can be tuned by domain swap between both type-II subclasses (Fortin and Lam, 2018; Zhu *et al.*, 2020). AtMC4 is a strongly expressed calcium-dependent type-II metacaspase, reaching high levels in Arabidopsis roots (Watanabe and Lam, 2011a; Hander *et al.*, 2019). AtMC4 is reported to enter the nuclei in genotoxic treatments like zeocin, allowing release to the cytosol of Arabidopsis Lamin 1 (AtLa1) (Huang *et al.*, 2018), a RNA binding protein that initiates the expression of WUSCHEL (WUS) in the shoot apical meristem (Cui *et al.*, 2015). AtLa1 interacts with AtMC4, and it is proposed as AtMC4 substrate. However, *in vitro* AtMC4 protease assays showed faint processing of AtLa1, and neither confirmation of AtMC4 footprints were indicated like calcium dependency, nor cleavage abolition by mutation on the substrate or inhibitor application. In chapter 3, we report on PROPEP1 as a *bona fide* substrate of AtMC4 both *in vitro* and *in vivo* conditions, and by determining the necessity of elevated calcium levels after wounding (Hander *et al.*, 2019). PROPEP processing did not occur when metacaspase activity was depleted by treatment with specific inhibitors or calcium chelating agents. Moreover, PROPEP cleavage was shown to be inhibited when the catalytic cysteine performing the nucleophilic attack was mutated as well as when the arginine at position P1 preceding the processing in PROPEP1 was exchanged for an amino acid of different nature. As a consequence of AtMC4 hydrolysis, PROPEP1 releases a 23 amino acids length peptide denominated Pep1, a Damage Associated Molecular Pattern (Huffaker *et al.*, 2006). Cleavage of PROPEP1 requires elevated levels of calcium, and while AtMC4 is capable of cleaving other

PROPEP members *in vitro* or in wounded and protoplast cells (Hander *et al.*, 2019; Shen *et al.*, 2019). The AtMC4 structure was recently solved by X-ray crystallography, confirming the critical importance of calcium binding to AtMC4 (Zhu *et al.*, 2020). Calcium ions bind to aspartic rich motifs in the linker region, displacing the proximity of the catalytic cysteine and subsequent cleavage and disengagement of the p20 domain from the p10 part. Upon high calcium availability, successive cleavage rapidly occur obtaining an active AtMC4 that can exert its activity on its substrates. Depletion of available calcium with the use of chelators or by mutation of the calcium binding amino acids abolishes AtMC4 processing. In Norway spruce (*Picea abies*), the metacaspase McII-Pa, has been shown necessary for the correct function of the suspensor during germination stages (Bozhkov *et al.*, 2005). McII-Pa is a type-II metacaspase that requires calcium for its activity and, similarly to AtMC4 as reported by Huang *et al.*, (2018), it re-localizes to the nuclei. In addition, McII-Pa can process Tudor Staphylococcal Nuclease (TSN), a shared substrate with mammalian caspases, but in this case McII-Pa -II hydrolyses after arginine and lysine instead of hydrolysis after aspartic acid (Sundstrom *et al.*, 2009).

The AtMC8 protein sequence is similar to the one of AtMC4, but in contrast, AtMC8 transcript levels under non-stressed conditions are very low. Its expression is strongly induced by oxidative conditions like methyl viologen or H_2O_2 along with exposure to other stress inductors like copper oxide nanoparticles (He *et al.*, 2008; Tang *et al.*, 2016). The ROS dependent expression for AtMC8 seems to be controlled by RCD1, a master regulator of the transcriptional responses in oxidative stresses (Ahlfors *et al.*, 2004). Using FR↓ or GRR↓ synthetic substrates AtMC8 was characterized showing an optimal pH of 7 to 8,5 on the recombinant protein and an activity induction when plants were exposed to H_2O_2 or UV (He *et al.*, 2008). An independent study on cell death induced by UV exposure reported DEVDase activity by the use of a sensor (Zhang *et al.*, 2009). However, AtMC8 did not show caspase-like activity and rather a metacaspase-like activity on arginine synthetic substrates at P1 (FR-AMC and GRR-AMC, AMC: aminomethylcoumarin). Therefore, it is unlikely that AtMC8 is the cause of DEVDase increase activity in plants during UV exposure, and the responsible protease remain unknown.

AtMC9 is the only type-II Arabidopsis metacaspase known to be active at low pH values. AtMC9 is expressed in cotyledon cells, protoxylem, root cap, tracheary

elements and petals prior to abscission and it is involved in the autolysis of cells during xylem cell death (Bollhoner *et al.*, 2013; Tsiatsiani *et al.*, 2013). While AtMC9 activity is coordinated with other cysteine proteases, such as Xylem Cysteine Protease 2 (XCP2) in protoxylem vessels clearance, the molecular basis triggering this cell-type specific cell death is not totally clear (Escamez and Tuominen, 2014). AtMC9 can process XCP2 *in vitro*, but *in vivo* validation remains unconfirmed. In a first plant N-terminome study, several potential substrates of AtMC9 were identified (Tsiatsiani *et al.*, 2013). From the validated substrates *in vitro*, PEPCK1 proteolysis resulted in the increase of its overall activity. Expression of PEPCK1 occurs at early stages of development, matching the patterns for AtMC9 (Rojas *et al.*, 2020). PEPCK proteolysis evidenced the contribution of AtMC9 during the remobilization of nutrients at sink-to-source transition stage of Arabidopsis seedlings. AtMC9 is also involved in the production of peptides with functions in cell survival, *i.e.* KRATOS, and cell death, *i.e.* BIA (Escamez *et al.*, 2019). These peptides were differentially abundant in cell cultures undergoing xylem differentiation when AtMC9 is normally expressed compared to when AtMC9 levels were reduced using an RNAi line. Exogenous application of KRATOS to leaf disks reduced ion leakage and BIA application enhanced the effect of stress induced ion leakage in wound-induced cell death or superoxide-inducing reaction. Maturation of both peptides, if they are direct targets of AtMC9 and how they contribute to the formation of the xylem needs more research for its understanding. Another metacaspase-derived peptide is GRIM REAPER, which binds to the receptor kinase PRK5 promoting cell death through a ROS dependent system (Wrzaczek *et al.*, 2015). AtMC9 is responsible for various processing sites in the protein sequence generating small size peptides influencing cell death. GRI is perceived by the Leucine Rich Repeat Receptor Like Kinase POLLEN RECEPTOR KINASE 5 (PRK5). Neither *prk5* nor *atmc9* mutants display an altered phenotype in relation to pathogen infection, while *gri* mutant is less susceptible to colonization by the virulent pathogen *Pseudomonas syringae* pv. tomato DC3000 (Wrzaczek *et al.*, 2009). The expression of the *Capsicum annuum CaMC9*, homolog to AtMc9, increases in presence of bacterial pathogens (Kim *et al.*, 2013). Silencing of *CaMC9* reduced plant cell death and its overexpression in *N. benthamiana* enhances ROS accumulation leading to a higher disease susceptibility and increased cell death. Two AtMC9-like proteases, *PttMC13* and *PttMC14*, are expressed in xylem tissues in poplar (Bollhoner *et al.*, 2018). Silencing of these genes modulate protein levels of proteases such as the

papain like cysteine protease RESPONSIVE TO DESICCATION 21 (RD21), a PUTATIVE ASPARTIC PROTEASE 3 (PASPA3) and a homologue of the known metacaspase target TSN in poplar. A metacaspase from pear tree is involved in lignification processes by expression induction of lignin biosynthetic genes affecting lignin accumulation in fruit organs (Gong *et al.*, 2020). In maize, metacaspase family members are expressed differentially in response to salicylic acid and jasmonate, and additionally they bind to Rp1-D21, a nucleotide-binding, leucine-rich repeat (NLR) capable to induce a necrotic response (Ma *et al.*, 2021).

Some protease inhibitors are able to block the proteolytic activity of both metacaspases type-II and I. The promiscuous *SERINE PROTEASE INHIBITOR 1* (SERPIN1) targets both cysteine and serine proteases (Lampl *et al.*, 2013; Ghorbani *et al.*, 2016; Rustgi *et al.*, 2017) and is capable to diminish AtMC1 and AtMC9 activity (Vercammen *et al.*, 2006; Lema Asqui *et al.*, 2018). SERPINs capture proteases by a one-pass suicide-like mechanism that involves interaction of the inhibitor reactive center loop with the protease. During the nucleophilic attack, the catalytic pocket of the protease is in proximity to the C-terminal part of the inhibitor. During the interaction, structural changes occur, and the C-terminal part relocates towards the rest of the inhibitory structure, dragging the protease with it. The resulting inhibitor-protease complex remains inactive while they bind covalently, given that the catalytic pocket is shed, and it is prone to degradation by other proteases or the proteasomal system (Huntington *et al.*, 2000; Cohen *et al.*, 2019). Therefore, SERPIN1 and homologues could control the activity of type-I metacaspases limiting the hypersensitive response to restricted areas where the proteases have been activated. Further, the activity of type-II metacaspases can be controlled by post-translational modifications. An example is the S-nitrosylation of the catalytic cysteine of AtMC9, which interfere with its enzymatic activity (Belenghi *et al.*, 2007). Similarly, the catalytic cysteine of AtMC4 has been identified in proteomic screen as being sulfenylated in H_2O_2 stressed cell cultures (Huang *et al.*, 2019). The significance of this modification for the AtMC4 performance *in vivo* activity remains to be studied. It seems that there are multiple layers controlling the enzymatic activity of type-I and type-II metacaspases, including proteinase inhibitors and PTMs on the catalytic pocket derived from the cellular redox balance.

In conclusion, metacaspases are proteases with evident contribution to plant performance in different species, but the molecular characterization is still partially

missing. Type I proteases are genetic regulators of hypersensitive response in plants, but their targets and interactors remain unidentified. On the other hand, there are more studies on the function of type-II proteases. For some type-II metacaspases, N-terminomics and peptidomics studies have allowed identification of substrates and differential regulated proteins (Tsiatsiani *et al.*, 2013; Bollhoner *et al.*, 2018; Escamez *et al.*, 2019), but their downstream characterization is still missing. Regulation of the type-II metacaspase activity can be controlled by a red-ox balance (Belenghi *et al.*, 2007; Huang *et al.*, 2019) and some of these stimuli induce the expression of other metacaspases (He *et al.*, 2008). At the moment, there is not a clear link between the functions of type-I and type-II metacaspases other than they share some inhibitors like SERPIN1 (Vercammen *et al.*, 2006; Lema Asqui *et al.*, 2018). The ability to recombinantly purify and study AtMC4 *in vitro* resulted in the elucidation of its structure and into a better understanding of its activation mechanism of AtMC4 (Zhu *et al.*, 2020). The AtMC4 structure elegantly explained its calcium dependency for self-processing and subsequent cleavage of its substrates such as PROPEP1. On the other hand, it has not been tested how calcium might influence the activity of type-I metacaspases and its relation in pathogen infection. Lastly, processes that are metacaspase dependent like PROPEP maturation and processes of other substrates in stress signalling seems like an open field where to search for metacaspase contribution to plant performance. We predict that in the coming years, additional studies combining of traditional -omics like transcriptomic response of metacaspases in stress in combination with positional proteomics and interactomics will place metacaspases as an alluring topic to investigate.

REFERENCES

Ahlfors R, Lang S, Overmyer K*, et al.* 2004. Arabidopsis RADICAL-INDUCED CELL DEATH1 belongs to the WWE protein-protein interaction domain protein family and modulates abscisic acid, ethylene, and methyl jasmonate responses. *Plant Cell* **16,** 1925-1937.

Belenghi B, Romero-Puertas MC, Vercammen D, Brackenier A, Inze D, Delledonne M, Van Breusegem F. 2007. Metacaspase activity of Arabidopsis thaliana is regulated by S-nitrosylation of a critical cysteine residue. *J Biol Chem* **282,** 1352-1358.

Bollhoner B, Jokipii-Lukkari S, Bygdell J*, et al.* 2018. The function of two type II metacaspases in woody tissues of Populus trees. *New Phytol* **217,** 1551-1565.

Bollhoner B, Zhang B, Stael S, Denance N, Overmyer K, Goffner D, Van Breusegem F, Tuominen H. 2013. Post mortem function of AtMC9 in xylem vessel elements. *New Phytol* **200,** 498-510.

Bozhkov PV, Suarez MF, Filonova LH, Daniel G, Zamyatnin AA, Jr., Rodriguez-Nieto S, Zhivotovsky B, Smertenko A. 2005. Cysteine protease mcII-Pa executes programmed cell death during plant embryogenesis. *Proc Natl Acad Sci U S A* **102,** 14463-14468.

Cambra I, Garcia FJ, Martinez M. 2010. Clan CD of cysteine peptidases as an example of evolutionary divergences in related protein families across plant clades. *Gene* **449,** 59-69.

Chen JM, Rawlings ND, Stevens RA, Barrett AJ. 1998. Identification of the active site of legumain links it to caspases, clostripain and gingipains in a new clan of cysteine endopeptidases. *FEBS Lett* **441,** 361-365.

Choi CJ, Berges JA. 2013. New types of metacaspases in phytoplankton reveal diverse origins of cell death proteases. *Cell Death Dis* **4,** e490.

Cohen M, Davydov O, Fluhr R. 2019. Plant serpin protease inhibitors: specificity and duality of function. *J Exp Bot* **70,** 2077-2085.

Coll NS, Smidler A, Puigvert M, Popa C, Valls M, Dangl JL. 2014. The plant metacaspase AtMC1 in pathogen-triggered programmed cell death and aging: functional linkage with autophagy. *Cell Death Differ* **21,** 1399-1408.

Coll NS, Vercammen D, Smidler A, Clover C, Van Breusegem F, Dangl JL, Epple P. 2010. Arabidopsis type I metacaspases control cell death. *Science* **330,** 1393-1397.

Cui Y, Rao S, Chang B*, et al.* 2015. AtLa1 protein initiates IRES-dependent translation of WUSCHEL mRNA and regulates the stem cell homeostasis of Arabidopsis in response to environmental hazards. *Plant Cell Environ* **38,** 2098-2114.

Dietrich RA, Delaney TP, Uknes SJ, Ward ER, Ryals JA, Dangl JL. 1994. Arabidopsis mutants simulating disease resistance response. *Cell* **77,** 565-577.

Earnshaw WC, Martins LM, Kaufmann SH. 1999. Mammalian caspases: structure, activation, substrates, and functions during apoptosis. *Annu Rev Biochem* **68,** 383-424.

Erez E, Fass D, Bibi E. 2009. How intramembrane proteases bury hydrolytic reactions in the membrane. *Nature* **459,** 371-378.

Escamez S, Stael S, Vainonen JP*, et al.* 2019. Extracellular peptide Kratos restricts cell death during vascular development and stress in Arabidopsis. *J Exp Bot* **70,** 2199-2210.

Escamez S, Tuominen H. 2014. Programmes of cell death and autolysis in tracheary elements: when a suicidal cell arranges its own corpse removal. *J Exp Bot* **65,** 1313-1321.

Fortin J, Lam E. 2018. Domain swap between two type-II metacaspases defines key elements for their biochemical properties. *Plant J* **96,** 921-936.

Ghorbani S, Hoogewijs K, Pecenkova T*, et al.* 2016. The SBT6.1 subtilase processes the GOLVEN1 peptide controlling cell elongation. *J Exp Bot* **67,** 4877-4887.

Gilio JM, Marcondes MF, Ferrari D, Juliano MA, Juliano L, Oliveira V, Machado MFM. 2017. Processing of metacaspase 2 from Trypanosoma brucei (TbMCA2) broadens its substrate specificity. *Biochim Biophys Acta Proteins Proteom* **1865,** 388-394.

Gong X, Xie Z, Qi K*, et al.* 2020. PbMC1a/1b regulates lignification during stone cell development in pear (Pyrus bretschneideri) fruit. *Hortic Res* **7,** 59.

Hander T, Fernandez-Fernandez AD, Kumpf RP*, et al.* 2019. Damage on plants activates Ca(2+)-dependent metacaspases for release of immunomodulatory peptides. *Science* **363**.

He R, Drury GE, Rotari VI, Gordon A, Willer M, Farzaneh T, Woltering EJ, Gallois P. 2008. Metacaspase-8 modulates programmed cell death induced by ultraviolet light and H2O2 in Arabidopsis. *J Biol Chem* **283,** 774-783.

Huang J, Willems P, Wei B*, et al.* 2019. Mining for protein S-sulfenylation in Arabidopsis uncovers redox-sensitive sites. *Proc Natl Acad Sci U S A* **116,** 21256-21261.

Huang Y, Cui Y, Hou X, Huang T. 2018. The AtMC4 regulates the stem cell homeostasis in Arabidopsis by catalyzing the cleavage of AtLa1 protein in response to environmental hazards. *Plant Sci* **266,** 64-75.

Huffaker A, Pearce G, Ryan CA. 2006. An endogenous peptide signal in Arabidopsis activates components of the innate immune response. *Proc Natl Acad Sci U S A* **103,** 10098-10103.

Huntington JA, Read RJ, Carrell RW. 2000. Structure of a serpin-protease complex shows inhibition by deformation. *Nature* **407,** 923-926.

Kim SM, Bae C, Oh SK, Choi D. 2013. A pepper (Capsicum annuum L.) metacaspase 9 (Camc9) plays a role in pathogen-induced cell death in plants. *Mol Plant Pathol* **14,** 557-566.

Klemencic M, Funk C. 2018. Type III metacaspases: calcium-dependent activity proposes new function for the p10 domain. *New Phytol* **218,** 1179-1191.

Koonin EV, Aravind L. 2002. Origin and evolution of eukaryotic apoptosis: the bacterial connection. *Cell Death Differ* **9,** 394-404.

Lallemand J, Bouche F, Desiron C, Stautemas J, de Lemos Esteves F, Perilleux C, Tocquin P. 2015. Extracellular peptidase hunting for improvement of protein production in plant cells and roots. *Front Plant Sci* **6,** 37.

Lampl N, Alkan N, Davydov O, Fluhr R. 2013. Set-point control of RD21 protease activity by AtSerpin1 controls cell death in Arabidopsis. *Plant J* **74,** 498-510.

Lee RE, Brunette S, Puente LG, Megeney LA. 2010. Metacaspase Yca1 is required for clearance of insoluble protein aggregates. *Proc Natl Acad Sci U S A* **107,** 13348-13353.

Lema Asqui S, Vercammen D, Serrano I, Valls M, Rivas S, Van Breusegem F, Conlon FL, Dangl JL, Coll NS. 2018. AtSERPIN1 is an inhibitor of the metacaspase AtMC1-mediated cell death and autocatalytic processing in planta. *New Phytol* **218,** 1156-1166.

Ma S, Shi H, Wang GF. 2021. The potential roles of different metacaspases in maize defense response. *Plant Signal Behav***,** 1906574.

McLuskey K, Mottram JC. 2015. Comparative structural analysis of the caspase family with other clan CD cysteine peptidases. *Biochem J* **466,** 219-232.

Minina EA, Coll NS, Tuominen H, Bozhkov PV. 2017. Metacaspases versus caspases in development and cell fate regulation. *Cell Death Differ* **24,** 1314-1325.

Parrish AB, Freel CD, Kornbluth S. 2013. Cellular mechanisms controlling caspase activation and function. *Cold Spring Harb Perspect Biol* **5**.

Rawlings ND, Barrett AJ, Bateman A. 2010. MEROPS: the peptidase database. *Nucleic Acids Res* **38,** D227-233.

Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD. 2018. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res* **46,** D624-D632.

Rojas BE, Hartman MD, Figueroa CM, Iglesias AA. 2020. Proteolytic cleavage of Arabidopsis thaliana phosphoenolpyruvate carboxykinase-1 modifies its allosteric regulation. *J Exp Bot*.

Rustgi S, Boex-Fontvieille E, Reinbothe C, von Wettstein D, Reinbothe S. 2017. Serpin1 and WSCP differentially regulate the activity of the cysteine protease RD21 during plant development in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* **114,** 2212-2217.

Salguero-Linares J, Coll NS. 2019. Plant proteases in the control of the hypersensitive response. *J Exp Bot* **70,** 2087-2095.

Salvesen GS, Hempel A, Coll NS. 2016. Protease signaling in animal and plant-regulated cell death. *FEBS J* **283,** 2577-2598.

Schechter I, Berger A. 1967. On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* **27,** 157-162.

Sergeeva TF, Shirmanova MV, Zlobovskaya OA, Gavrina AI, Dudenkova VV, Lukina MM, Lukyanov KA, Zagaynova EV. 2017. Relationship between intracellular pH, metabolic co-factors and caspase-3 activation in cancer cells during apoptosis. *Biochim Biophys Acta Mol Cell Res* **1864,** 604-611.

Shen W, Liu J, Li JF. 2019. Type-II Metacaspases Mediate the Processing of Plant Elicitor Peptides in Arabidopsis. *Mol Plant* **12,** 1524-1533.

Stennicke HR, Salvesen GS. 1997. Biochemical characteristics of caspases-3, -6, -7, and -8. *J Biol Chem* **272,** 25719-25723.

Sundstrom JF, Vaculova A, Smertenko AP*, et al.* 2009. Tudor staphylococcal nuclease is an evolutionarily conserved component of the programmed cell death degradome. *Nat Cell Biol* **11,** 1347-1354.

Talanian RV, Quinlan C, Trautz S, Hackett MC, Mankovich JA, Banach D, Ghayur T, Brady KD, Wong WW. 1997. Substrate specificities of caspase family proteases. *J Biol Chem* **272,** 9677-9682.

Tang Y, He R, Zhao J, Nie G, Xu L, Xing B. 2016. Oxidative stress-induced toxicity of CuO nanoparticles and related toxicogenomic responses in Arabidopsis thaliana. *Environ Pollut* **212,** 605-614.

Thornberry NA, Rano TA, Peterson EP*, et al.* 1997. A combinatorial approach defines specificities of members of the caspase family and granzyme B. Functional relationships established for key mediators of apoptosis. *J Biol Chem* **272,** 17907-17911.

Tsiatsiani L, Timmerman E, De Bock PJ*, et al.* 2013. The Arabidopsis metacaspase9 degradome. *Plant Cell* **25,** 2831-2847.

Tsiatsiani L, Van Breusegem F, Gallois P, Zavialov A, Lam E, Bozhkov PV. 2011. Metacaspases. *Cell Death Differ* **18,** 1279-1288.

Uren AG, O'Rourke K, Aravind LA, Pisabarro MT, Seshagiri S, Koonin EV, Dixit VM. 2000. Identification of paracaspases and metacaspases: two ancient families of caspase-like proteins, one of which plays a key role in MALT lymphoma. *Mol Cell* **6,** 961-967.

van der Hoorn RA. 2008. Plant proteases: from phenotypes to molecular mechanisms. *Annu Rev Plant Biol* **59,** 191-223.

Vercammen D, Belenghi B, van de Cotte B*, et al.* 2006. Serpin1 of Arabidopsis thaliana is a suicide inhibitor for metacaspase 9. *J Mol Biol* **364,** 625-636.

Vercammen D, van de Cotte B, De Jaeger G*, et al.* 2004. Type II metacaspases Atmc4 and Atmc9 of Arabidopsis thaliana cleave substrates after arginine and lysine. *J Biol Chem* **279,** 45329-45336.

Watanabe N, Lam E. 2011a. Arabidopsis metacaspase 2d is a positive mediator of cell death induced during biotic and abiotic stresses. *Plant J* **66,** 969-982.

Watanabe N, Lam E. 2011b. Calcium-dependent activation and autolysis of Arabidopsis metacaspase 2d. *J Biol Chem* **286,** 10027-10040.

Wrzaczek M, Brosche M, Kollist H, Kangasjarvi J. 2009. Arabidopsis GRI is involved in the regulation of cell death induced by extracellular ROS. *Proc Natl Acad Sci U S A* **106,** 5412-5417.

Wrzaczek M, Vainonen JP, Stael S*, et al.* 2015. GRIM REAPER peptide binds to receptor kinase PRK5 to trigger cell death in Arabidopsis. *EMBO J* **34,** 55-66.

Yuan S, Yu X, Topf M, Ludtke SJ, Wang X, Akey CW. 2010. Structure of an apoptosome-procaspase-9 CARD complex. *Structure* **18,** 571-583.

Zhang L, Xu Q, Xing D, Gao C, Xiong H. 2009. Real-time detection of caspase-3-like protease activation in vivo using fluorescence resonance energy transfer during plant programmed cell death induced by ultraviolet C overexposure. *Plant Physiol* **150,** 1773-1783.

Zhu P, Yu XH, Wang C, Zhang Q, Liu W, McSweeney S, Shanklin J, Lam E, Liu Q. 2020. Structural basis for Ca(2+)-dependent activation of a plant metacaspase. *Nat Commun* **11,** 2249.

Chapter 2

"Proteolytic activation mechanisms in plants"

Álvaro Daniel Fernández-Fernández[1,2] , Simon Stael[1,2] and Frank Van Breusegem[1,2]

[1] Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium

[2] Center for Plant Systems Biology, VIB, 9052 Ghent, Belgium

Author contributions:

ADFF wrote the manuscript with the help of SS and FVB. SS and FVB revised and corrected the manuscript.

Aim and context

This chapter collects the knowledge on the mode of action and pathways involving proteolysis in plant systems. To date, there are not many evidences of the mechanistic models inducing activation of certain type of proteases. Here we review the processes known involving induction of hydrolytic activation such as calcium, pH oscillation or dimerization. The function of proteases by co-expression in certain tissues together with their substrates or sequential processing of substrates during trafficking different cellular compartments is also commented. Moreover, we address plant-pathogen interactions and how the presence of infectious organisms can induce plant proteolytic activity in pathways not totally deciphered. Lastly, we study the action of effector agents that happen to carry proteolytic activity and how it is perceived by the plants and answered in different fashion.

Abstract:

Proteases alter the fate of their substrates by hydrolysis of their peptidyl bonds. Due to the inherent irreversibility of this post-translational modification, proteolysis needs to be strictly regulated in the cellular environment. Several mechanisms ensure that proteolytic activities occur in a highly specific spatiotemporal manner in order to minimize ectopic proteolysis events that might disturb plant growth and fitness. The different control layers include the production of many proteases as inactive versions, dependence on the environmental conditions of a specific cellular location or signals derived from a specific stress and control of the protease stability and turnover. Despite the involvement of proteases in a broad scale of cellular processes regulating plant growth and development, abiotic stress and immune responses, the knowledge of their regulatory mechanisms is scarce. Here, we review the signals and processes leading to activation of plant proteases.

Keywords: Plant Proteases, Substrate cleavage, proteolysis, signalling, metacaspases, plant pathogens.

Introduction

Proteases are widespread and abundant enzymes with the capacity to hydrolyse peptidyl bonds between specific amino acids in their substrates (Rawlings et al., 2012; Rawlings et al., 2018; Jones, 2021). This irreversible post-translational modification (PTM) results in the formation of new carboxyl- and amino-termini of the substrate protein. As such, proteolytically cleaved proteins may not only lose their functional properties, but in some cases also acquire new functionalities. Besides their, best known, role in substrate destruction, proteases or protease networks are often key components within signalling cascades, by producing functional isoforms from inactive substrates, like bio-active peptides (Paulus and Van der Hoorn, 2019). Therefore, proteases can act as molecular “switchers” that modulate diverse cellular processes through the cleavage of their substrate proteins. In plants, the number of proteins with predicted peptidase activity is extensive. For example, in *Arabidopsis* and poplar between 600 and 800 proteases are annotated (Garcia-Lorenzo et al., 2006; van der Hoorn, 2008; Lallemand et al., 2015). Plant proteases are integrated in a broad scale of cellular events: from their relevance in organellar protein import (van Wijk, 2015) growth and development (Schaller, 2004; Liu et al., 2018) plant adaptation and responses to abiotic stresses, immunity and the hypersensitive response (Balakireva and Zamyatnin, 2018; Salguero-Linares and Coll, 2019), to developmental programmed cell death events (Salvesen et al., 2016; Buono et al., 2019). However, in contrast with the size of their gene families and vital importance in many cellular processes, plant protease’s mode of action and activation mechanisms remain largely enigmatic.

Proteases are mostly translated as proenzymes, also called zymogens, requiring appropriate environmental conditions or triggers to function. Proteolytic activity can be induced by changes in the cellular environment as derived from stress induction or by changes in the surrounding due to protease translocation to a new location with dissimilar conditions. In parallel, protease activity can be controlled by protease inhibitors that can trap them and drive them to the protease-inhibitor complex towards their degradation (Vercammen et al., 2006; Lema Asqui et al., 2018). In some cases, this interaction occurs only when the protease is active, showing a certain permissibility of the proteolytic action followed by a controlled restriction in the enzymatic activity. Other proteases are self-processed, which can limit their lifetime

and functions in the cells (Kapust et al., 2001). So, protease activity is controlled at multiple levels; at gene expression level, prior to initiation of their activity, during hydrolysis and after hydrolysis events take place. As proteolysis is irreversible, these events are believed to minimize unspecific cleavage of over-active proteases that otherwise may disturb the balance of the cellular environment.

External cues controlling proteolytic activity: Ca^{2+}, pH and redox balance.

Most proteases are translated as inactive zymogens with a prodomain that blocks its activity and dependence on external factors to be activated. The activation of most plant type-II metacaspases is dependent on increased calcium levels to initiate self-processing and subsequent activation (Vercammen et al., 2004; Watanabe and Lam, 2011; Wen et al., 2013). Metacaspases structurally resemble mammalian caspases and entail two conserved subunits, p20 and p10, named according to their 20 and 10 kDa sizes, that are interconnected by a linker sequence (Uren et al., 2000; Tsiatsiani et al., 2011; Minina et al., 2017; Klemencic and Funk, 2019). In contrast to caspases, type II metacaspases have a predilection to cleave substrates after arginine and lysine amino acids. In *Arabidopsis*, studied type-II metacaspases are dependent on increased calcium levels, in the millimolar range, for activation, with the exception of AtMC9 whose activation depends on acidic pH (Fig. 1A) (Vercammen et al., 2004; Watanabe and Lam, 2011). The high calcium levels required to activate the calcium dependent metacaspases *in vitro* and *in vivo* can be reached in physiological conditions by tissue wounding. The cellular disruption allows calcium to reach the cytoplasmic area where AtMC4 resides, cleaving its substrates PROPEPs (Hander et al., 2019; Shen et al., 2019). Although not experimentally demonstrated, most likely other stimuli can also infuse enough intracellular calcium to activate metacaspases. Proteolytic processing of PROPEP1 leads to the release of Pep1 that binds to the receptors PEPR1 and PEPR2, thereby promoting interaction with their co-receptor kinase BAK1 to initiate a phosphorylation cascade and subsequent transcriptional activation of a defense response (Yamada et al., 2016). Such a chain of events indicates that metacaspases act as Ca^{2+} signature decoders to transduce perceived stress signals into response pathways. Recently, the crystal structure of AtMC4 revealed the details on the Ca^{2+}-dependency of the activation mechanism (Zhu et al., 2020). Calcium cations interact with the negatively charged domain in the linker region and thereby destabilize the basic side chain the lysine residue at position 225, which

otherwise blocks the catalytic pocket. The release of the two catalytic residues enables self-processing and subsequent cleavage at other sites in the C- and N-terminal parts of the linker for a full or partial activation, respectively.

For AtMC9 there is experimental evidence that its activation is regulated by different means. Recombinant AtMC9 is highly active at 5 to 5.5 pH values, while its self-processing activity is impeded by S-nitrosylation of the catalytic cysteine. Interestingly, due to a “back-up” cysteine in the proximity of the catalytic pocket, S-nitrosylation has no effect on AtMC9 activity towards substrates once self-processed (Belenghi et al., 2007). AtMC9 activity is irreversibly blocked by SERPIN1, inhibiting its substrate processing and subsequently leading to its degradation (Vercammen et al., 2006). A plethora of substrates for AtMC9 are identified and validated, providing useful information for AtMC9 target prediction (Tsiatsiani et al., 2013). While there is no crystal structure for AtMC9 yet, this information would be useful to integrate a model with the pH-dependent activation, the AtMC9-SERPIN1 interaction and the redox control on the AtMC9 catalytic site.

The Autophagy-related protein 4 (ATG4) cysteine proteases A and B are required to process the C-terminal end of the ubiquitin-like protein ATG8, which enables its cargo binding in the nascent autophagosome and consecutive transport of proteins to the lytic vacuole (Fig. 1D). In Arabidopsis, the *in vitro* activity of ATG4 is reversibly inhibited by H_2O_2 (Woo et al., 2014). The enzymatic activity of ATG4A is more affected by the presence of H_2O_2 than the activity of its homolog ATG4B. In yeast, a redox-sensitve cysteine pair affects ATG4 proteolytic activity (Perez-Perez et al., 2014). The presence of oxidative environment promotes a single disulphide between two cysteines at position 338 and 394 in yeast ATG4 that inhibits its activity. Mutation of the any of the cysteine for a serine turns ATG4 redox independent and highly active towards its substrate ATG8. Thioredoxin can efficiently reduce the disulphide bond between the two cysteines, inducing ATG4 activity. In addition, ATG4 activity in Arabidopsis is induced during osmotic stress and nitrogen starvation by a hitherto unknown mechanism, while both in yeast and in *Chlamydomonas* a redox-dependent activation function was demonstrated *in vitro* (Perez-Perez et al., 2016). Independent of reactive oxygen species, intracellular hydrogen sulphide negatively regulates autophagy via a reversible persulfidation of the cysteine within the catalytic cysteine-histidine-aspartic triad (Laureano-Marin et al., 2020).

Some serine proteases can also process propeptide hormones with different functions to Pep1. This is the case of a subtilase named phytaspase, that can cleave after an aspartic residue like mammalian capases. Tobacco phytaspase can cleave both in *cis* for its prodomain release, and in *trans* for its substrates like systemin. Systemin was the first identified peptide hormone in plants composed of 18 amino acid length which was found in jasmonate-dependent wounding events (Pearce et al., 1991). Phytaspase is able to process proSystemin, releasing a Leu-Systemin precursor (Beloshistov et al., 2018). For full functionality, the N-terminal leucine requires to be removed by other protease than phytaspase, likely a leucine aminopeptidase, which is also transcriptionally induced by wounding (Chao et al., 1999).Cleavage and release of systemin leads to recognition by its specific receptors, initiating secondary signals like an oxidative burst and ethylene production (Wang et al., 2018). Interestingly, phytaspases can re-localise from the apoplast to the intracellular space to initiate programmed cell death under some stresses (Chichkova et al., 2010). In the absence of obvious wounding stress, methyl viologen treatment and viral infection trigger the engulfement of the apoplastic phytaspase into the cytoplasm by clathrin dependent endocytosis (Trusova et al., 2019). This process shows selectivity for phytaspase remobilization when compared to other proteases like cathepsins (Chichkova et al., 2010). Whether this process also plays during wounding responses has not been reported.

A third of the precursor peptides from the *Arabidopsis* Rapid Alkalinisation-Like Factor (RALF) family bear a canonical subtilase cleavage site cleaved by SBT6.1, also named Site-1 Protease (Srivastava et al., 2009). RALF23 overexpression leads to a bushy and dwarf phenotype while overexpression in the subtilase mutant showed no visible differences to wild type, indicating that high levels of processable RALF delay normal growth possibly by its unrestricted cleavage. RALF peptides acts as ligands binding to FERONIA, a malectin-like RLK that acts as a scaffold for multiple signalling pathways (Zhang et al., 2020). RALF mature peptides binding to FERONIA induce reorganization of other RLK present in plants membranes like the FLAGELIN SENSITIVE2 (FLS2: LRR-RLK that binds to the bacterial PAMP flg22 of 22 amino acid length), EF-TU RECEPTOR (EFR: LRR-RLK receptor of bacterial elf18) and BRI1-ASSOCIATED RECEPTOR KINASE (BAK1: LRR serine/threonine protein kinase that acts as co-receptor for multiple LRR-RLK) FER (Fig. 1B) (Stegmann et al., 2017).

Destabilization of the interactions between the different RLK plant receptors impairs the plant immunity response, and possibly other mechanisms involved in development, being a plausible explanation for the acute phenotype observed in RALF overexpression lines. These and later studies showed that the proteolytic activity of SBT6.1 needs be in endomembranous system (Stuhrwohldt et al., 2020a), hinting to a possible regulation by the pH levels. Perhaps other modifications or post-translational modifications occurring at the different stages of the protease transport through the different membrane fractions can induce its activity.

Protease status change by mono/di-merization

The activity of the vacuolar *Arabidopsis* Legumain-γ (AtLEGγ) is controlled through a pH-dependent dimerization. In a very acidic environment, AtLEGγ monomerizes and operates as a peptidase. In a neutral environment, AtLEGγ dimerizes to act mainly as a ligase. In an intermediate pH level, both activities can co-exist, and are modulated by the monomer-dimer variants and nucleophilicity of the catalytic domain in the peptidase pocket side (Fig. 1C). Hence, a model is proposed in which AtLEGγ can work as a gear of different mechanisms in discrete subcellular environments such as vacuole, apoplast or cytosol, adjusting to varying pH conditions (Zauner et al., 2018).

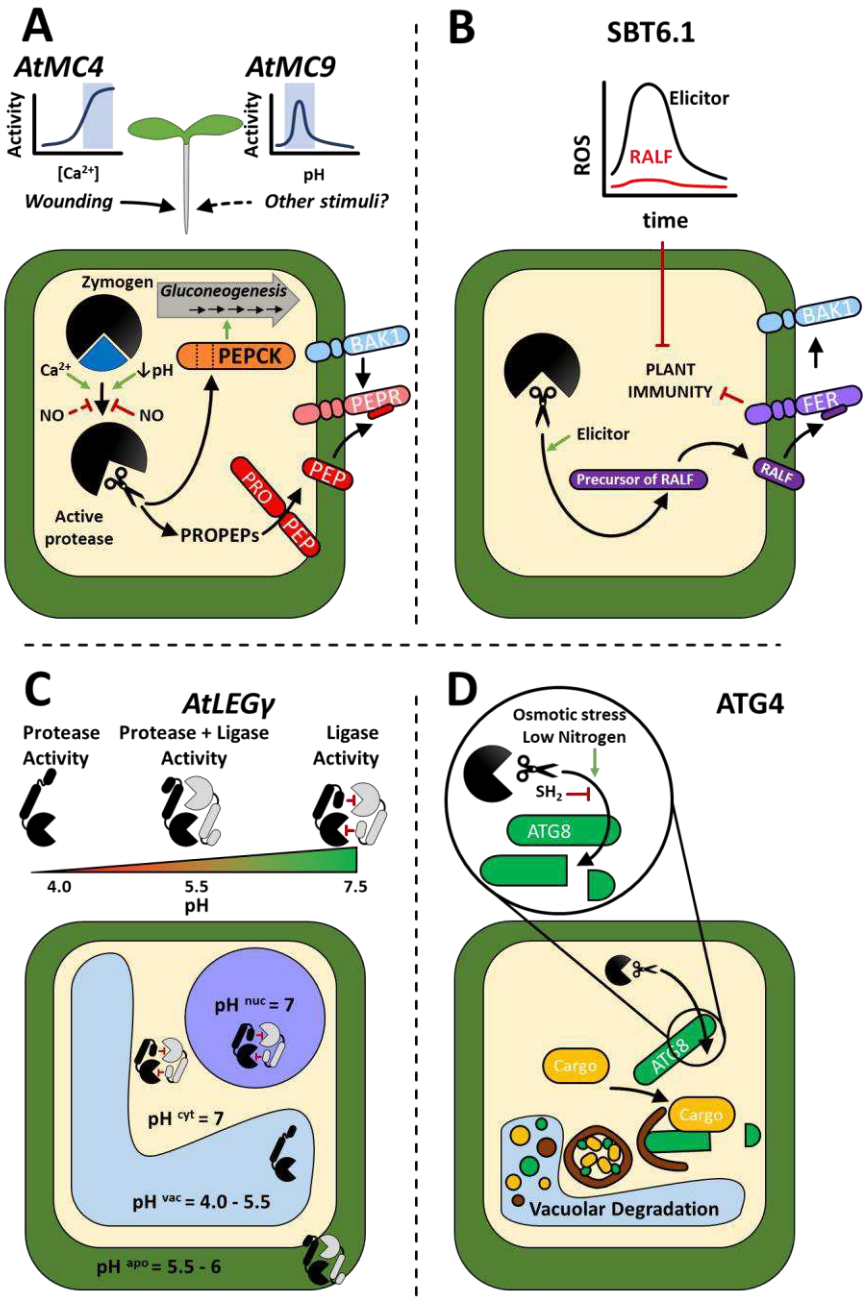


Figure 1: Direct environmental cues affecting proteolytic activity in plants. A) Type-II metacaspase activation can occur via calcium increase of or pH drop. Calcium dependent metacaspases can be activated through the calcium released after wounding, while pH dependent metacaspases mechanism remain to be further explained. Activation from a zymogen to active protease is also regulated by nitrosylation of a conserved catalytic cysteine. After self-processing metacaspases can cleave PROPEPs family members which can then leave the cell and bind its receptor PEPR1 and PEPR2 promoting interaction with BAK1. AtMC9 can cleave PEPCK enhancing its activity and promoting gluconeogenesis during seedling germination. **B)** *Arabidopsis* SBT6.1 can process RALF precursor which is a negative regulator of growth and immune system by rearranging FER interaction, including a decrease in the interaction with BAK1. **C)** *Arabidopsis* legumains are able to switch activities from proteolytic to ligase activity. At lower pH legumains function as peptidases while at neutral pH the main activity is ligase. Still, in an intermediate status, they can combine both activities. This flexibility allows legumains to work as different enzymes depending on the pH suborganelle in which they are located (pH shown in pictures). **D)** The cysteine protease ATG4A and ATG4B control processing of ATG which is induced under environmental stimuli like nitrogen starvation and osmotic stress. Upon cleavage of ATG8, it can recruit adaptor and cargo proteins to the autophagosome (brown) and direct these proteins to a lytic vacuole where they are degraded.

Protease activation by post-translational modifications

DA1 peptidase, named after the Chinese word “big” (大: phonetically pronounced “dà”), contributes to leaf growth and size together with the DA1-related 1 and 2 (Peng et al., 2015). The C-terminus of these proteases encodes a zinc metallopeptidase controlled by a “cysteine switch” that coordinates a cysteine residue with a zinc atom at the active site to block it. DA1 is mono-ubiquitylated by interaction with E2 and E3 ligases. Ubiquitin changes releases the zinc lock and activates the DA1 peptidase activity, subsequently processing its substrates UBP15, TCP14/15 and TCP22 (Dong et al., 2017). DA1 and alike activity is controlled by UBP12 and UBP13, which de-ubiquitinate the protease, counteracting the availability of active DA1 proteases and diminishing its substrates processing, which is needed for normal leaf size development (Vanhaeren et al., 2020).

Substrate availability by post-translational modification

Protease-substrate interactions can also be regulated in a substrate-centric way. The precursor of the *Arabidopsis* CLE40 peptide, which contributes to root stem cell maintenance, can be processed by a triplet of subtilases (SBT1.4, SBT1.7 and SBT4.13) at two different sites. Interestingly, cleavage of the second site is hampered by proline hydroxylation, a common PTM in many secreted peptides, and thereby dampening peptide activity (Stuhrwohldt et al., 2020a). The above findings indicate that peptide processing occurs in a controlled manner at different stages during transport across organelles and that some PTMs can regulate not only the activity of the proteases by affecting its enzymatic activities, but also as additional signals that can shape the cleavability of the substrate repertoire.

Coordinated proteases: Sequential processing of substrates.

Many plant bioactive peptides are first embedded within a precursor protein, needing in some cases several processing at C and N-termini, or sequential cleavage for their maturation (Tavormina et al., 2015). These action can be performed by one protease or by coordinated activities of proteases, where an initial processing allows subsequent access of another protease to the sequence. *E.g.,* the generation of mature systemin requires additional cleavage of the observed by phytaspase as indicated above. Confirmation of the leucine residue removal by a leucine

aminopeptidase, would place it as a good example of coordinated processing with an additional activation of wound-elicited signal. There are many peptides and substrates that are processed by different proteases. For instance, the maturation of the secreted peptides from the CLAVATA3 EMBRYO SURROUNDING REGION (CLE)-like family peptides (CLEL) in *Arabidopsis* is performed by subtilases (Ghorbani et al., 2016). A CLEL gene termed *GOLVEN1* (*GLV1*) plays an important function in development. Upon overexpression, GLV1 causes a wavy gravitropic effect in roots and increased hypocotyl elongation. However, a normal root appearance was reconstituted when *GLV1* was overexpressed in subtilase SBT6.1 and SBT6.2 loss-of-function mutants and co-overexpression of the serine protease inhibitor *SERPIN1* suppressed the hypocotyl phenotype. *SERPIN1* and subtilase expression patterns partially overlapped, and they interact with each other. Later, SBT6.1 dependent processing in a conserved cleavage motif of other secreted peptides, CLEL6 and CLEL 9, was observed at the Golgi (Stuhrwohldt et al., 2020b). After cleavage by SBT6.1, CLEL peptides require additional processing by the apoplastic subtilase SBT3.8, which is controlled by the environmental pH at the apoplast location.

Proteolysis regulated by protease/substrate expression and co-localization

While protease regulation can depend largely on external factors, some other proteases are regulated predominantly by their spatio-temporal expression characteristics. In this situation, processing would rely on the encounter of protease and substrate. In *Arabidopsis*, abscission of the flower organs and other cell separation processes like lateral root emergence depend on the peptide hormone *Inflorescence Deficient in Abscission* (IDA) (Shi et al., 2019). *IDA* loss-of-function plants fail to drop their sepals and petals after fertilization and reveal a role of processed peptide by proteases in cell wall loosening during lateral root emergence (Kumpf et al., 2013). Subtilase inhibition can be achieved by expressing Extracellular Protease Inhibitor (EPI) identified in oomycetes-plant interaction (Tian et al., 2004; Tian et al., 2005). When EPI inhibitor is expressed under control of the *IDA* promoter, both petal and anther detachments were impaired (Schardon et al., 2016) (Fig. 2A), and the phenotype was overruled by local application of a peptide containing the active IDA sequence. The gene expression patterns of a cohort of subtilases (*SBT2.2, SBT2.6, SBT3.1, SBT4.6, SBT4.8, SBT4.10, SBT4.12, SBT4.13* and *SBT5.2*) in the basipetal zone during flower development nicely overlapped with the *IDA* expression

patterns, therefore necessitating the targeted overexpression of EPI inhibitors to overcome genetic redundancy of subtilases. Similarly, premature whole flower abscission and decreased fruit yield in drought stressed tomato plants, could be dampened by reducing the expression of Phytaspase2 (Reichardt et al., 2018). Subtilase expression is stimulated by drought stress in the proximal areas to the flower abscission zone and vasculature. Phytosulfokine (PSK) was identified as the responsible hormone as exogenous application of PSK suffices to control the expression of abscission related genes and ultimately induce the flower drop (Fig. 2B). Despite the similarities, it is important to highlight that subtilase cleavage of IDA in Arabidopsis induces petal and sepal abscission, while SlPHYT2 activity in tomato is responsible for PSK processing resulting in total flower drop at the pedicel. For both cases, cell separation is the final causative process, showing that peptide maturation of different peptide hormones and signalling through independent receptors can converge in parallel abscission mechanisms at various plant organs (Meng et al., 2016; Santiago et al., 2016; Zhang et al., 2018).

Plants are highly dependent for their survival and reproduction on how they interpret the external signals they are facing during their growth and development. Yet, some parasitic plants have developed a dependency to other plants to initiate germination and growth. After colonization, a specialized type of cells in these parasitic plants will promote cell differentiation into a xylem bridge, which performs the initial contact and enables the access to host nutrients (Fig. 2C). Recent studies identified four subtilases expressed specifically in these intrusive cells of *P. japonicus* (SBT1.1.1, SBT1.2.3, SBT1.7.2, SBT1.7.3). Inhibition of the subtilase activity by expressing the EPI inhibitor under the control of a promoter expressed at the host-parasite contact point, led to reduced levels of colonization, lower expression of the intrusive cell marker PjICSL1 and a decrease in total auxin levels, all features necessary for the transition of parasite tracheary elements into xylem bridge cells. During parasite-host colonization, plant parasitic subtilases hence aid the correct differentiation of cells necessary for xylem bridge formation and cell specificity. The peptides involved in this signalling process and the possible triggers of their activity are not identified yet, but the fact that the proteolytic inhibitor reduced the efficiency of host colonization hints that this activity is dependent by the subtilases at this location (Ogawa et al., 2020).

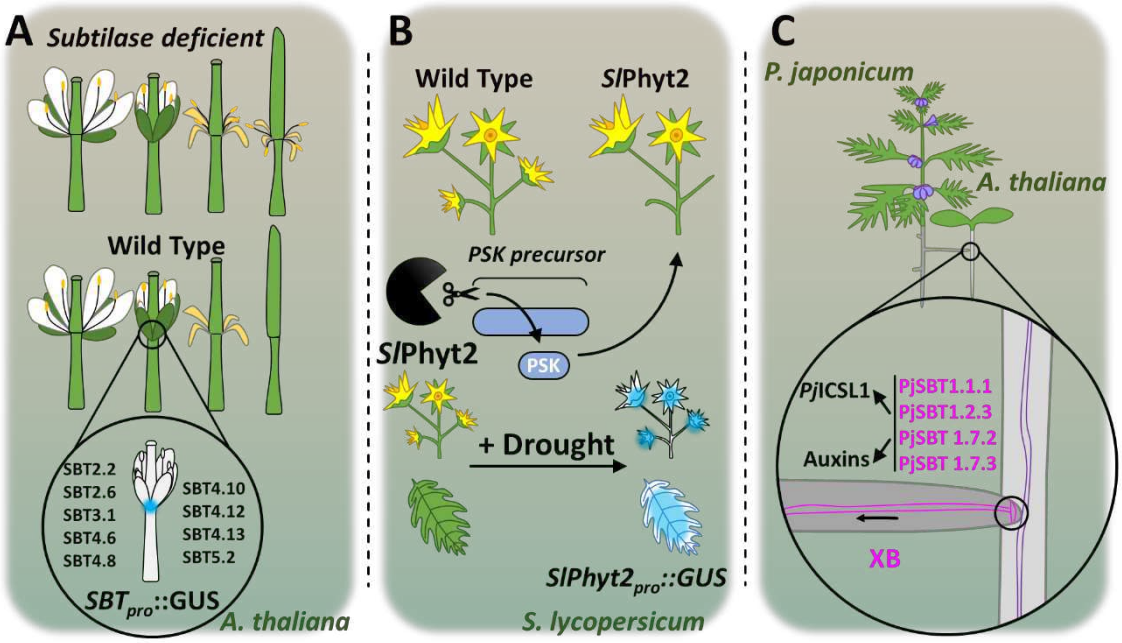


Figure2: Phenotypes representation and localization of responsible proteases for organ abscission and plant-host plant interaction. A) *Arabidopsis* flower phenotypes showing the depletion of the subtilase activity with the use of a broad subtilase inhibitor under the control of IDA promoter abolished the dehiscence of petals and anthers during fertilization stages. Candidate subtilases were identified by expression analysis using specific promoter regions fusions to β-glucuronidase (GUS) reporting gene expression in the basal apex of *Arabidopsis* flowers of possible responsible executors of IDA processing. **B)** Phenotype of tomato plants undergoing drought stress resulting in lower number of total flowers in the inflorescence. Drought stress is responsible for the expression of SlPhyt2 in leaves and regions proximal to flower buds during its development. SlPhyt2 cleaves PSK and initiates a series of processes leading to cell separation and flower loss. **C)** Subtilases expressed at the contact points of the parasitic plant *Phtheirospermum japonicus* assist in the early events of plant interaction with its host and XB formation, through maintenance of necessary genes like *PJ*ICSL1 and XB formation preserving Auxin levels.

Plant-pathogen relationships through proteolysis

Early detection and identification of possible pathogen assaults, can lead to a better preparation of plant responses and enhances the possibility of survival. Peptide generation through proteolysis and its perception is an important instrument used by plants to sense bacteria (Zipfel et al., 2004; Zipfel et al., 2006). Many of the transcriptional responses to bacterial infection have been thoroughly studied depicting an orchestrated expression of defence genes. But, not surprisingly, proteases at both the plant and pathogen side are involved in early recognition events. Recently, Zip1 (*Zea mays* immune peptide 1) was identified upon SA treatment of maize as a secreted

peptide in biotrophic presence (Ziemann et al., 2018). Mutational analysis and specific inhibitor application confirmed that the Papain like Cysteine Proteases (PLCPs) CP1 and CP2 are accountable for Zip1 maturation. More interestingly, external Zip1 application induced endogenous cleavage of its precursor proZip1 but the activation mechanisms of CP1 and CP2 are unknown. It seems unlikely that Zip1 has a traditional receptor. Instead, the authors suggested that Zip1 might work by activating the proteases CP1 and CP2 by binding to an exosite and promoting cleavage of its precursor proZip. Zip1 induction of SA-responsive genes can be explained as a sophisticated mechanism where plants have overcome SA dependence and its suppression by bacterial effectors (Fig. 3A). For example, some biotrophs, like *Ustillago maydis*, impair SA accumulation by secreting a chorismate mutase that lowers SA precursor availability and thereby minimize plant defense response (Djamei et al., 2011). Similarly, some pathogens target PLCPs, such as the *U. maydis* effector Pit2 or the endogenous ZmCC9 (van der Linde et al., 2012; Mueller et al., 2013). The abundance of bacterial effectors targeting specific groups of proteases pinpoints to the importance of their activity for plant defence against pathogens.

In another plant-bacteria interaction study, *Arabidopsis* leaves incubated with *P. aeruginosa* experienced shared pathogenic response to bacterial elicitor treatments. Further investigation identified the *Pseudomonas* serine protease (PrpL) and its closest homolog ArgC from *Xanthomonas* as responsible for the induction of pathogenic genes (Cheng et al., 2015). In fact, only the protease presence sufficed to match pathogenic response as seen by the bacterial derived peptide flg22 in terms of gene expression, ROS burst and MAPK3/6 phosphorylation, but not MAPK4. The authors identified an impaired oxidative burst in G-protein mutant lines and they discovered a cascade reactions involving RACK1 as scaffold where a subset of MAPKK and MAPK, including MAPK3/6 interact with RACK1 and G-proteins (Fig. 3B). The model leaves some open questions such as, what are the substrates that induce a flg22-like response in Arabidopsis leaves? How are the detection mechanisms translating apoplastic signals through MAPK phosphorylation and gene expression? Is there a specific receptor for the derived substrates?

Other plant immunity related proteases can cleave bacterial proteins to control their growth during invasion. This is the case for the Secreted Aspartic Proteases 1 and 2 (SAP1 and SAP2) which process *in planta* MucD, a conserved HtrA-like protease,

necessary for bacterial growth (Wang et al., 2019). Although his processing would affect indifferently to both pathogenic and commensal microorganisms on the surface of the plants, authors suggest that it could work as a surveillance mechanism to keep excessive bacterial growth at bay (Fig. 3C).

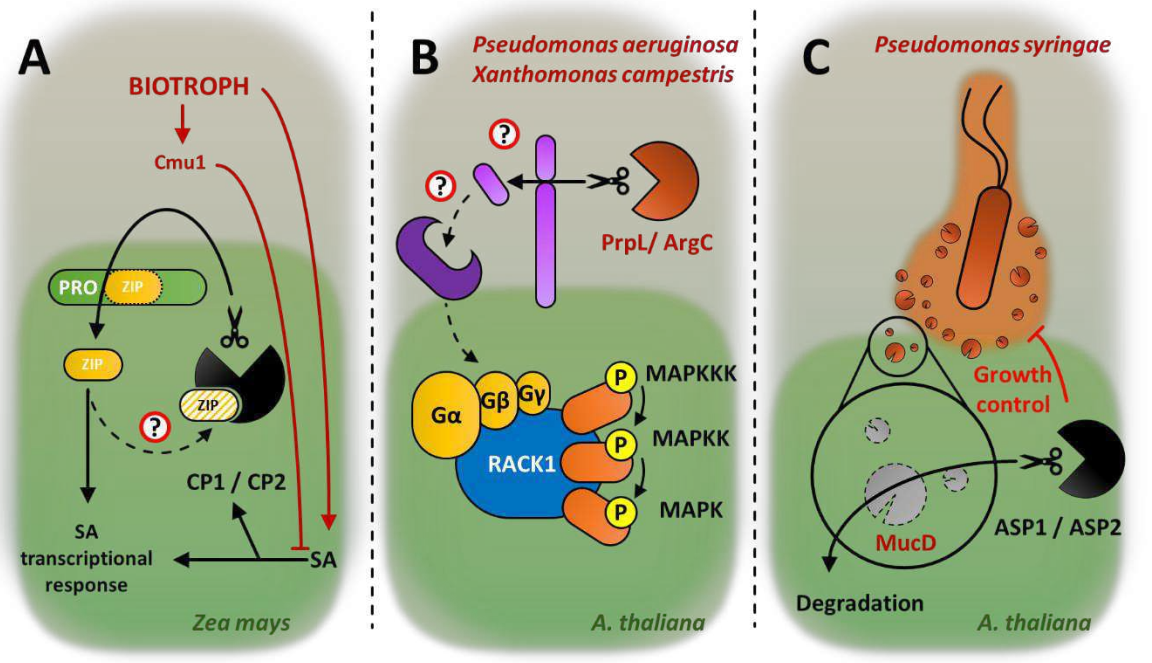


Figure3: Novel identified plant-bacteria interaction mechanism based on proteolysis. A) Plant biotrophs induce SA response in *Z. mays* activating the PLCP CP1 and CP2, which cleaves the precursor of Zip1 (ProZip1). Zip1 can provoke similar responses to SA, including transcriptional induction of SA-dependent genes and activation of their processing proteases CP1 and CP2 in a positive feedback loop. Zip1-proteolytic activation remains to be confirmed by direct interaction with CP1 and CP2. Zip1 SA-independent pathway would disentangle maize plants from the necessity to generate SA which in some situations can be dampened by bacterial effectors such as Cmu1. **B)** PrpL and ArgC bacterial proteases homologs from *P. aeruginosa* and *X. campestris* trigger phosphorylation of MAPK3 and MAPK6 through a previously unknown mechanism where G proteins couple a yet undeciphered signal to MAPKKK, MAPKK and MAPK phosphorylation through RACK1 structural scaffold. **C)** ASP1 and ASP2 process the extracellular bacterial protein MucD which contributes to the bacterial growth. By degrading MucD, plants can keep at bay the growth of their colonizing bacterial in the plant apoplast maintaining a balance with both pathogenic and commensal bacterial.

Detection of pathogens by proteolytic events

As indicated above, proteolysis during plant-pathogen interactions are used on both sides of the plant-pathogen frontline. Proteolytic activity can be perceived indirectly, through decoys that sense proteolysis-derived products. Some bacterial effectors, once injected in the plant host environment, aim to minimize the plant defense response to allow the pathogen to remain undercover as long as possible. Undetected pathogens are more likely to spread through the plant tissues, reproduce and colonize other parts of the plant. On the other hand, plant mechanisms include perception of the signals and when necessary, plant cells will undertake a drastic decision as a hypersensitive response where a group of cells died in a programmed way, isolating a region with an infective host. In this host-pathogen race, a myriad of mechanisms have emerged dampening the colonization, detection and response. Proteolysis and protease-substrate pairs are also involved. One example is the recognition of the effector pathogens via the membrane protein RIN4. Various bacterial effectors target RIN4, for example AvrB and AvrRpm1 disrupt RIN4 interaction with RPM1, which is then activated (Mackey et al., 2002). RIN4 is also a target of AvrRpt2, a bacterial cysteine protease that is injected in the plant cell through a typical Type-3 Secretion System. In the presence of AvrRpt2, RIN4 is degraded resulting in the perception of processed RIN4 by the NB-LRR RPS2, which can activate cell death pathways (Mackey et al., 2003). RIN4 degradation is caused through direct cleavage by AvrRpt2 at two conserved motifs (RCS1 and RCS2) at the N and C-terminal locations of RIN4 (Kim et al., 2005). Cleavage at the most C-terminal site RCS2 turns out to be indispensable for RIN4 de-attachment from the membrane domain and following degradation, which activates RPS2 while cleavage of the RCS1 had no effect in the plant response (Fig. 4B).

Another *P. syringae* cysteine protease (AvrPphB) effector has been shown to function inside of plant cells where it processes multiple pathogen sensing kinases like BIK1, PBL1 and PBL2 (Zhang et al., 2010). A target of AvrPphB is PBS1, another plant kinase that triggers hypersensitive response. Cell death response in plants occurs after AvrPphB cleaves PBS1 and the perception of the hydrolysis by RPS5 (Resistance to *P. syringae* 5) (Shao et al., 2003). Immunoprecipitation results showed that both active and inactive versions of AvrPphB are able bind to PBS1. The cleavage site was determined using Edman sequencing identifying a motif glycine-aspartic-

lysine (↓GDK), showing possible preference for glycine at P1', as AvrPphB self - processing occurs between a lysine and a glycine (K ↓ G). Although the GDK motif was conserved and found in other kinases, its presence was a necessary premise but not sufficient for their cleavage. RPS5 gene is a nucleotide-binding leucine-rich repeat (NLR) controlling infection by AvrPphB (Swiderski and Innes, 2001) with a N-terminal coiled coil and Nucleotide Binding Sites domains indispensable for HR response, and a C-terminal LRR domain that inhibits HR in normal conditions in companion of uncleaved PBS1 (Ade et al., 2007). In absence of the effector, RPS5 and PBS1 remain in a pre-assembled complex. During the presence of the effector, RPS5 detects the complex of the cleaved PBS1 and AvrPphB (Fig. 4C). The LRR domain binds to the C-terminal part of PBS1, possibly retaining AvrPphB, which leads to a change in the structure of RPS5. That physical change switches RPS5A affinity from ADP to ATP, with subsequent changes in direct partner interactions that pilots the plant to enter HR.

In tomato, the fungal effector Avr2 targets the secreted RCR3 Papain Like Cysteine protease. After secretion, RCR3 is activated in the apoplast by cleavage of its prodomain. Initially, it was thought that RCR3 was self-activated, but the presence of prodomain processing in plants expressing a catalytically inactive RCR3 version showed that other proteases were responsible for its maturation in *trans.* In practise, P69B was verified as one of the upstream proteases contributing to RCR3 maturation, but RCR3 activation was shown to be induced by other subtilases (Paulus et al., 2020). This event shows the first validated proteolytic cascade in plants, with multiple proteases that work as initiator proteases, and modulating RCR3 activity which is the final proteolytic effector of the path. Once RCR3 is processed, it can interact with Avr2, shaping a complex that is recognized by the leucine rich repeat receptor Cf-2 initiating a signalling response that finalizes in localized cell death. Interestingly, it is not a requirement that RCR3 is catalytically active to establish an Avr2-RCR3 complex, but its prodomain requires it to be removed enabling Avr2 accessibility. The pathway requires RCR3 to interact both with Avr2 and Cf-2 to signal hypersensitive response and it seemed of great importance for survival of a singular clade among Solanaceae species (Kourelis *et al.*, 2019). This pathway describes the first real proteolytic cascade in plants (Paulus and van der Hoorn, 2019), and serves as an illustration of nature's complexity evolution system as response to a plant pathogen.

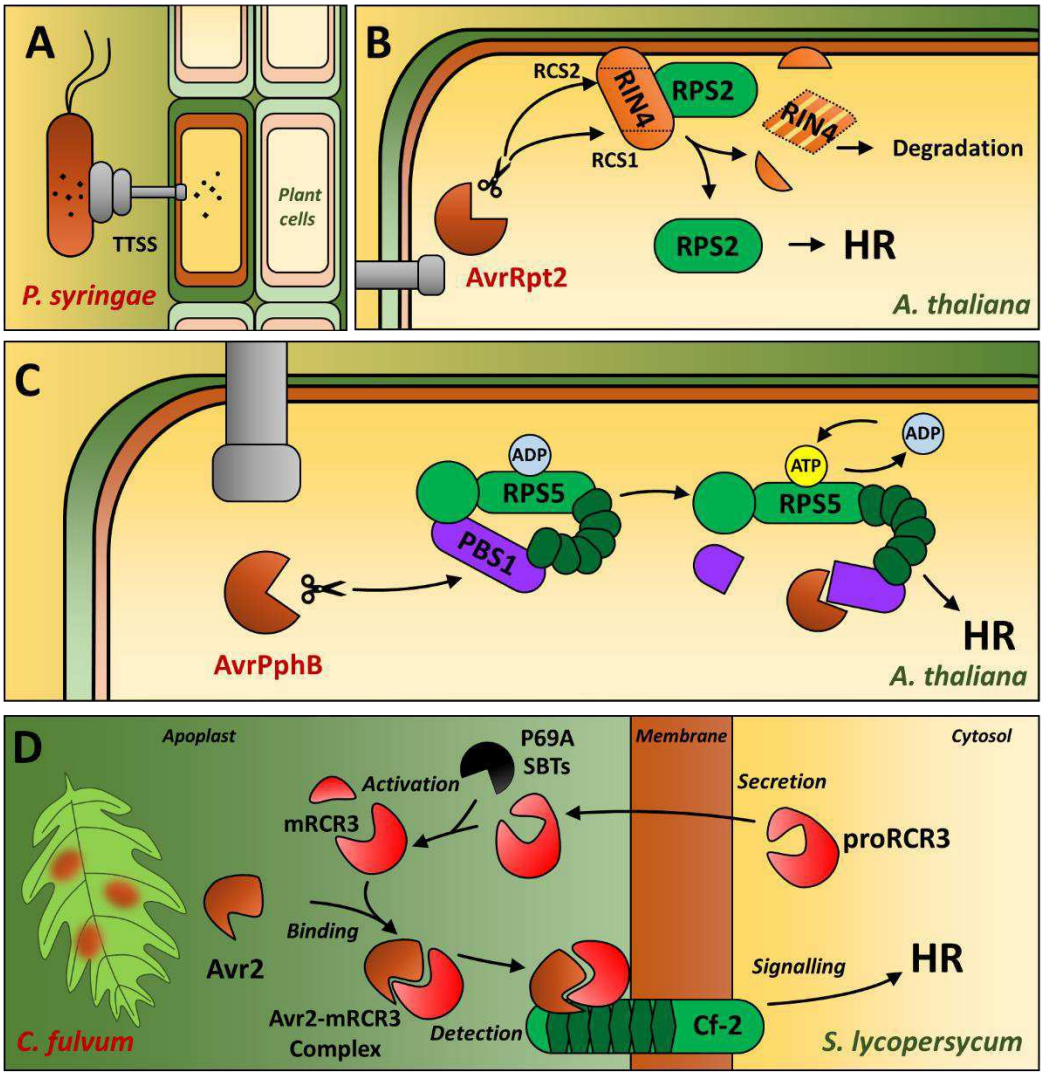


Figure 4. Proteolysis sensing mechanisms in plants based on decoys. A) Representation of bacterial Type Three Secretion System (TTSS) commonly used by Gram negative bacteria to inject effectors from the bacterial cell to the host. **B)** Resistant mode of action for plants recognition of AvrRpt2 cleavage of RIN4. Upon AvrRpt2 presence, RIN4 is processed at two sites called (RCS1 and RCS2). Processing of RCS2 will lead to separation of RIN4 from the membrane and its destabilization through degradation. RPS2 acts a decoy by being release of its initial interaction with RIN4, change localization and is able to signal and initiate an HR response. **C)** Mechanism for the perception of the cysteine protease AvrPphB from *P. syringae* pv. phaseolicola. RPS5 and PBS1 are interacting upon normal conditions in a primed status. During infection the effector AvrPphB cleaves PBS1 which produces a conformational change in the decoy RPS5 and after exchange of ADP for ATP triggers HR. **D)** Perception and signalling pathway for the *C. fulvum* effector Avr2 in tomato plants. RCR3 PLCP is secreted to the apoplast as a zymogen where subtilases like P69B process it by release of RCR3 prodomain. When Avr2 is secreted to the apoplast from the fungi, it targets the active RCR3and shape a complex with it. The complex is recognized by Cf-2 only upon the presence of both RCR3 and Avr2 and only then Cf-2 activates HR.

Conclusions

In this chapter, we reviewed the proteolytic events deciphered in the last years including direct events modulating proteases activity, coordinated work of proteases for multi-step maturation, tissue specific proteolytic activity in biotic and abiotic stresses, and indirect detection of proteolytic events. Proteases are important actors in multiple cellular pathways ranging from recognition of external signals to protein dismantlers in degradation processes. In animal systems, proteases have extensively attracted researchers' attention, mainly due to their functions in cell death and as potential drug targets in biomedical applications. In addition, proteases of diverse origin serve a purpose as research tools and useful enzymes in multiple industrial applications (Folgado and Abranches, 2020; Pottinger and Innes, 2020).

In contrast, our current knowledge built on the mechanistic processes that induce plant proteases remains scarce, as is the functional understanding of the cleavage of the majority of substrates. In comparison to mammalian systems, only a handful of plant proteomic studies towards an unbiased identification of protease targets are reported (Tsiatsiani et al., 2013; Venne et al., 2015; Demir et al., 2018).

Deciphering the proteolytic conditions required for peptide synthesis and their maturation will lead to better understanding of each pathway. Here, we tried to review the current knowledge of plant protease regulatory mechanisms. A better understanding of proteolytic activation mechanisms would require description of the surrounding conditions affecting protease and modulation of the protease activity. At the same time identification of the location sharing active protease and substrate. Overall, the activity shall be determined by processing of a protease, or its family members, to a substrate in an in *situ* context. The enzymatic processing could be validated in an *ex situ* framework by cleavage assays *in vitro* of recombinant proteins or by other means in alternative model organisms, lysates or reticulocyte systems. Determination of the substrate-processing site might not be indispensable, but facilitates the information on the preferential amino acids of a functional protease. Mutation of the protease catalytic site is also a strategy to verify that the protease is enzymatically responsible and that the predicted catalytic site is *de facto* involved in its processing. The surrounding circumstances should be also carefully regarded while looking at enzymatic activity. For example, metacaspases would only work under specific calcium and pH levels, while legumains might change enzymatic functions and

subtilases will only be active in the acidic pH at the apoplast. Some cues directly modify the structure of proteases activating them, other ones will facilitate the substrate-protease encounter, and still without a full understanding of the mechanisms they will be main driving forces affecting protease activity.

A hurdle when working with protease families is the redundant activity of relative proteases towards individual substrates. In the last years, protease-class inhibitors showed to be effective tools surpass redundancy and discover hidden phenotypes that with other means, such as single mutant lines, would remain undetectable (Schardon et al., 2016; Ogawa et al., 2020; Reichardt et al., 2020). Alternatively, the generation of high order mutants is now possible through genome editing tools, like targeting several members of the type-II metacaspase family in *Arabidopsis* using CRISPR (Shen et al., 2019). This application led to discovery of an enhanced phenotype when tetramutants were challenged with pathogens. For some proteases their recombinant production and purification might be challenging. While this might complicate *in vitro* studies, their production in cell-free system individually or together with the substrate to test might come up as a solution for these challenges.

Organism	Protease type	Protease name	Species	Activation/Regulation	Substrates	Reference
Plant proteases	Cysteine protease	AtMC4	Arabidopsis	Calcium	Processing after Arginine/Lysine	Vercammen et al., 2004 Watanabe et al., 2011
					PROPEPs	Hander et al., 2019 Shen et al., 2019 Zhu et al., 2020
				Genotoxic treatment	Lamin1	Huang et al., 2018
		AtMC9	Arabidopsis	Low pH	-	Vercammen et al., 2004
				S-nytrosilation	-	Belenghi et al., 2007
				SERPIN1 inhibition	-	Vercammen et al., 2006
				-	Multiple	Tsiatsiani et al., 2013
		Legumain-γ	Arabidopsis	pH controlled dimerization	-	Zauner et al., 2018
		ATG4	Arabidopsis	Osmotic stress	-	Laureano-Marin et al., 2020
				Starvation	-	Laureano-Marin et al., 2020
				SH_2	-	Laureano-Marin et al., 2020
			Chlamydomonas	Redox inactivation	-	Perez-Perez et al., 2016
				H_2O_2	-	Woo et al., 2014
			Yeast	Thioredoxins	-	Perez-Perez et al., 2014
			(Multiple)	-	ATG8	Reviewed in Seo et al., 2016
		CP1 CP2	*Zea mays*	SA/Biotroph derived signal	Zip1	Ziemann et al., 2018
		RCR3	Tomato	Infection by *C. fulvum* / Avr2	-	Paulus et al., 2020
	Subtilase	P69A	Tomato	-	RCR3	
		Phytaspase	Tomato	-	Systemin	Beloshistov et al., 2018
			Tomato	Wound derived signals	-	Pearce et al., 1991
			N. tabacum	Apoplast to cytosol mobilization(?)	-	Chichkova et al., 2010
			N. tabacum	Selective clathrin endocytosis(?)	-	Trusova et al., 201
		SBT6.1 (S1P)	Arabidopsis	-	RALF	Srivastava et al., 2009
		SBT6.1 (S1P)	Arabidopsis	-	CLEL6 and CLEL 9	Stuhrwohldt et al., 2020a
		SBT6.1 SBT6.2	Arabidopsis	SERPIN1	GLV1	Ghorbani et al., 2016
		SBT3.8	Arabidopsis	pH dependent	CLEL6 and CLEL 9	Stuhrwohldt et al., 2020b

		SBT1.4 SBT1.7 SBT4.13	Arabidopsis	Substrate hydroxylation	CLE40	Stuhrwohldt et al., 2020a
		SBT2.2 SBT2.6 SBT3.1 SBT4.6 SBT4.8 SBT4.10 SBT4.12 SBT4.13 SBT5.2	Arabidopsis	Localized expression (?)	IDA	Schardon et al., 2016
		Phytaspase2	Tomato	Localized expression (?)	PSK	Reichardt et al., 2018
		SBT1.1.1 SBT1.2.3 SBT1.7.2 SBT1.7.3	*P. japonicus*	Plant parasite during colonization	-	Ogawa et al., 2020
	Metalloprotease	DA1	Arabidopsis	Ubiquitination	-	Dong et al., 2017
			Arabidopsis	UBP12/UBP13	-	Vanhaeren et al., 2020
	Aspartic protease	SAP1 SAP2	Arabidopsis	Plant-bacteria colonization	MucD	Wang et al., 2019
Bacterial proteases	Serine protease	PrpL	*Pseudomonas spp.*	Plant-bacteria colonization	-	Cheng et al., 2015
	Cysteine protease	AvrPphB	*P. syringae*	Plant-bacteria colonization	PBS1	Shao et al., 2003
					BIK1, PBL1 and PBL2	Zhang et al., 2010
		AvrRpt2	*P. syringae*	Plant-bacteria colonization	RIN4	Mackey et al., 2002

Table 1. Classification of the main proteases mentioned in this chapter. Proteases were classified in columns as plant or bacteria protease according to the nature of the organism encoding it. Additional classification included Protease type (Cysteine, Serine, subtilase, aspartic or metalloprotease), protease name, the species used for the study, the activation/regulation components and identified substrates followed by the bibliographic reference. When mechanisms where not clear, they were indicated with interrogation mark (?) or with a hyphen (-).

References

Ade, J., DeYoung, B.J., Golstein, C., and Innes, R.W. (2007). Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. Proc Natl Acad Sci U S A **104:** 2531-2536.

Balakireva, A.V., and Zamyatnin, A.A. (2018). Indispensable Role of Proteases in Plant Innate Immunity. Int J Mol Sci **19**.

Belenghi, B., Romero-Puertas, M.C., Vercammen, D., Brackenier, A., Inze, D., Delledonne, M., and Van Breusegem, F. (2007). Metacaspase activity of Arabidopsis thaliana is regulated by S-nitrosylation of a critical cysteine residue. J Biol Chem **282:** 1352-1358.

Beloshistov, R.E., et al. (2018). Phytaspase-mediated precursor processing and maturation of the wound hormone systemin. New Phytol **218:** 1167-1178.

Buono, R.A., Hudecek, R., and Nowack, M.K. (2019). Plant proteases during developmental programmed cell death. J Exp Bot **70:** 2097-2112.

Chao, W.S., Gu, Y.Q., Pautot, V.V., Bray, E.A., and Walling, L.L. (1999). Leucine aminopeptidase RNAs, proteins, and activities increase in response to water deficit, salinity, and the wound signals systemin, methyl jasmonate, and abscisic acid. Plant Physiol **120:** 979-992.

Cheng, Z., Li, J.F., Niu, Y., Zhang, X.C., Woody, O.Z., Xiong, Y., Djonovic, S., Millet, Y., Bush, J., McConkey, B.J., Sheen, J., and Ausubel, F.M. (2015). Pathogen-secreted proteases activate a novel plant immune pathway. Nature **521:** 213-216.

Chichkova, N.V., Shaw, J., Galiullina, R.A., Drury, G.E., Tuzhikov, A.I., Kim, S.H., Kalkum, M., Hong, T.B., Gorshkova, E.N., Torrance, L., Vartapetian, A.B., and Taliansky, M. (2010). Phytaspase, a relocalisable cell death promoting plant protease with caspase specificity. EMBO J **29:** 1149-1161.

Demir, F., Niedermaier, S., Villamor, J.G., and Huesgen, P.F. (2018). Quantitative proteomics in plant protease substrate identification. New Phytol **218:** 936-943.

Djamei, A., et al. (2011). Metabolic priming by a secreted fungal effector. Nature **478:** 395-398.

Dong, H., et al. (2017). Ubiquitylation activates a peptidase that promotes cleavage and destabilization of its activating E3 ligases and diverse growth regulatory proteins to limit cell proliferation in Arabidopsis. Genes Dev **31:** 197-208.

Folgado, A., and Abranches, R. (2020). Plant Aspartic Proteases for Industrial Applications: Thistle Get Better. Plants (Basel) **9**.

Garcia-Lorenzo, M., Sjodin, A., Jansson, S., and Funk, C. (2006). Protease gene families in Populus and Arabidopsis. BMC Plant Biol **6:** 30.

Ghorbani, S., Hoogewijs, K., Pecenkova, T., Fernandez, A., Inze, A., Eeckhout, D., Kawa, D., De Jaeger, G., Beeckman, T., Madder, A., Van Breusegem, F., and Hilson, P. (2016). The SBT6.1 subtilase processes the GOLVEN1 peptide controlling cell elongation. J Exp Bot **67:** 4877-4887.

Hander, T., et al. (2019). Damage on plants activates Ca(2+)-dependent metacaspases for release of immunomodulatory peptides. Science **363**.

Jones, R.A.C. (2021). Global Plant Virus Disease Pandemics and Epidemics. Plants (Basel) **10**.

Kapust, R.B., Tozser, J., Fox, J.D., Anderson, D.E., Cherry, S., Copeland, T.D., and Waugh, D.S. (2001). Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. Protein Eng **14:** 993-1000.

Kim, H.S., Desveaux, D., Singer, A.U., Patel, P., Sondek, J., and Dangl, J.L. (2005). The Pseudomonas syringae effector AvrRpt2 cleaves its C-terminally acylated target, RIN4, from Arabidopsis membranes to block RPM1 activation. Proc Natl Acad Sci U S A **102:** 6496-6501.

Klemencic, M., and Funk, C. (2019). Evolution and structural diversity of metacaspases. J Exp Bot **70:** 2039-2047.

Kumpf, R.P., Shi, C.L., Larrieu, A., Sto, I.M., Butenko, M.A., Peret, B., Riiser, E.S., Bennett, M.J., and Aalen, R.B. (2013). Floral organ abscission peptide IDA and its HAE/HSL2 receptors control cell separation during lateral root emergence. Proc Natl Acad Sci U S A **110:** 5235-5240.

Lallemand, J., Bouche, F., Desiron, C., Stautemas, J., de Lemos Esteves, F., Perilleux, C., and Tocquin, P. (2015). Extracellular peptidase hunting for improvement of protein production in plant cells and roots. Front Plant Sci **6:** 37.

Laureano-Marin, A.M., Aroca, A., Perez-Perez, M.E., Yruela, I., Jurado-Flores, A., Moreno, I., Crespo, J.L., Romero, L.C., and Gotor, C. (2020). Abscisic Acid-Triggered Persulfidation of Cysteine Protease ATG4 Mediates Regulation of Autophagy by Sulfide. Plant Cell.

Lema Asqui, S., Vercammen, D., Serrano, I., Valls, M., Rivas, S., Van Breusegem, F., Conlon, F.L., Dangl, J.L., and Coll, N.S. (2018). AtSERPIN1 is an inhibitor of the metacaspase AtMC1-mediated cell death and autocatalytic processing in planta. New Phytol **218:** 1156-1166.

Liu, H., Hu, M., Wang, Q., Cheng, L., and Zhang, Z. (2018). Role of Papain-Like Cysteine Proteases in Plant Development. Front Plant Sci **9:** 1717.

Mackey, D., Holt, B.F., 3rd, Wiig, A., and Dangl, J.L. (2002). RIN4 interacts with Pseudomonas syringae type III effector molecules and is required for RPM1-mediated resistance in Arabidopsis. Cell **108:** 743-754.

Mackey, D., Belkhadir, Y., Alonso, J.M., Ecker, J.R., and Dangl, J.L. (2003). Arabidopsis RIN4 is a target of the type III virulence effector AvrRpt2 and modulates RPS2-mediated resistance. Cell **112:** 379-389.

Meng, X., Zhou, J., Tang, J., Li, B., de Oliveira, M.V.V., Chai, J., He, P., and Shan, L. (2016). Ligand-Induced Receptor-like Kinase Complex Regulates Floral Organ Abscission in Arabidopsis. Cell Rep **14:** 1330-1338.

Minina, E.A., Coll, N.S., Tuominen, H., and Bozhkov, P.V. (2017). Metacaspases versus caspases in development and cell fate regulation. Cell Death Differ **24:** 1314-1325.

Mueller, A.N., Ziemann, S., Treitschke, S., Assmann, D., and Doehlemann, G. (2013). Compatibility in the Ustilago maydis-maize interaction requires inhibition of host cysteine proteases by the fungal effector Pit2. PLoS Pathog **9:** e1003177.

Ogawa, S., Wakatake, T., Spallek, T., Ishida, J.K., Sano, R., Kurata, T., Demura, T., Yoshida, S., Ichihashi, Y., Schaller, A., and Shirasu, K. (2020). Subtilase activity in intrusive cells mediates haustorium maturation in parasitic plants. Plant Physiology.

Paulus, J.K., and Van der Hoorn, R.A.L. (2019). Do proteolytic cascades exist in plants? J Exp Bot **70:** 1997-2002.

Paulus, J.K., et al. (2020). Extracellular proteolytic cascade in tomato activates immune protease Rcr3. Proc Natl Acad Sci U S A **117:** 17409-17417.

Pearce, G., Strydom, D., Johnson, S., and Ryan, C.A. (1991). A polypeptide from tomato leaves induces wound-inducible proteinase inhibitor proteins. Science **253:** 895-897.

Peng, Y., Chen, L., Lu, Y., Wu, Y., Dumenil, J., Zhu, Z., Bevan, M.W., and Li, Y. (2015). The ubiquitin receptors DA1, DAR1, and DAR2 redundantly regulate endoreduplication by modulating the stability of TCP14/15 in Arabidopsis. Plant Cell **27:** 649-662.

Perez-Perez, M.E., Lemaire, S.D., and Crespo, J.L. (2016). Control of Autophagy in Chlamydomonas Is Mediated through Redox-Dependent Inactivation of the ATG4 Protease. Plant Physiol **172:** 2219-2234.

Perez-Perez, M.E., Zaffagnini, M., Marchand, C.H., Crespo, J.L., and Lemaire, S.D. (2014). The yeast autophagy protease Atg4 is regulated by thioredoxin. Autophagy **10:** 1953-1964.

Pottinger, S.E., and Innes, R.W. (2020). RPS5-Mediated Disease Resistance: Fundamental Insights and Translational Applications. Annu Rev Phytopathol **58:** 139-160.

Rawlings, N.D., Barrett, A.J., and Bateman, A. (2012). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res **40:** D343-350.

Rawlings, N.D., Barrett, A.J., Thomas, P.D., Huang, X., Bateman, A., and Finn, R.D. (2018). The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Res **46:** D624-D632.

Reichardt, S., Piepho, H.P., Stintzi, A., and Schaller, A. (2020). Peptide signaling for drought-induced tomato flower drop. Science **367:** 1482-1485.

Reichardt, S., Repper, D., Tuzhikov, A.I., Galiullina, R.A., Planas-Marques, M., Chichkova, N.V., Vartapetian, A.B., Stintzi, A., and Schaller, A. (2018). The tomato subtilase family includes several cell death-related proteinases with caspase specificity. Sci Rep **8:** 10531.

Salguero-Linares, J., and Coll, N.S. (2019). Plant proteases in the control of the hypersensitive response. J Exp Bot **70:** 2087-2095.

Salvesen, G.S., Hempel, A., and Coll, N.S. (2016). Protease signaling in animal and plant-regulated cell death. FEBS J **283:** 2577-2598.

Santiago, J., Brandt, B., Wildhagen, M., Hohmann, U., Hothorn, L.A., Butenko, M.A., and Hothorn, M. (2016). Mechanistic insight into a peptide hormone signaling complex mediating floral organ abscission. Elife **5**.

Schaller, A. (2004). A cut above the rest: the regulatory function of plant proteases. Planta **220:** 183-197.

Schardon, K., Hohl, M., Graff, L., Pfannstiel, J., Schulze, W., Stintzi, A., and Schaller, A. (2016). Precursor processing for plant peptide hormone maturation by subtilisin-like serine proteinases. Science **354:** 1594-1597.

Shao, F., Golstein, C., Ade, J., Stoutemyer, M., Dixon, J.E., and Innes, R.W. (2003). Cleavage of Arabidopsis PBS1 by a bacterial type III effector. Science **301:** 1230-1233.

Shen, W., Liu, J., and Li, J.F. (2019). Type-II Metacaspases Mediate the Processing of Plant Elicitor Peptides in Arabidopsis. Mol Plant **12:** 1524-1533.

Shi, C.L., Alling, R.M., Hammerstad, M., and Aalen, R.B. (2019). Control of Organ Abscission and Other Cell Separation Processes by Evolutionary Conserved Peptide Signaling. Plants (Basel) **8**.

Srivastava, R., Liu, J.X., Guo, H., Yin, Y., and Howell, S.H. (2009). Regulation and processing of a plant peptide hormone, AtRALF23, in Arabidopsis. Plant J **59:** 930-939.

Stegmann, M., Monaghan, J., Smakowska-Luzan, E., Rovenich, H., Lehner, A., Holton, N., Belkhadir, Y., and Zipfel, C. (2017). The receptor kinase FER is a RALF-regulated scaffold controlling plant immune signaling. Science **355:** 287-289.

Stuhrwohldt, N., Ehinger, A., Thellmann, K., and Schaller, A. (2020a). Processing and formation of bioactive CLE40 peptide are controlled by post-translational proline hydroxylation. Plant Physiol.

Stuhrwohldt, N., Scholl, S., Lang, L., Katzenberger, J., Schumacher, K., and Schaller, A. (2020b). The biogenesis of CLEL peptides involves several processing events in consecutive compartments of the secretory pathway. Elife **9**.

Swiderski, M.R., and Innes, R.W. (2001). The Arabidopsis PBS1 resistance gene encodes a member of a novel protein kinase subfamily. Plant J **26:** 101-112.

Tavormina, P., De Coninck, B., Nikonorova, N., De Smet, I., and Cammue, B.P. (2015). The Plant Peptidome: An Expanding Repertoire of Structural Features and Biological Functions. Plant Cell **27:** 2095-2118.

Tian, M., Benedetti, B., and Kamoun, S. (2005). A Second Kazal-like protease inhibitor from Phytophthora infestans inhibits and interacts with the apoplastic pathogenesis-related protease P69B of tomato. Plant Physiol **138:** 1785-1793.

Tian, M., Huitema, E., Da Cunha, L., Torto-Alalibo, T., and Kamoun, S. (2004). A Kazal-like extracellular serine protease inhibitor from Phytophthora infestans targets the tomato pathogenesis-related protease P69B. J Biol Chem **279:** 26370-26377.

Trusova, S.V., Teplova, A.D., Golyshev, S.A., Galiullina, R.A., Morozova, E.A., Chichkova, N.V., and Vartapetian, A.B. (2019). Clathrin-Mediated Endocytosis Delivers Proteolytically Active Phytaspases Into Plant Cells. Front Plant Sci **10:** 873.

Tsiatsiani, L., Van Breusegem, F., Gallois, P., Zavialov, A., Lam, E., and Bozhkov, P.V. (2011). Metacaspases. Cell Death Differ **18:** 1279-1288.

Tsiatsiani, L., Timmerman, E., De Bock, P.J., Vercammen, D., Stael, S., van de Cotte, B., Staes, A., Goethals, M., Beunens, T., Van Damme, P., Gevaert, K., and Van Breusegem, F. (2013). The Arabidopsis metacaspase9 degradome. Plant Cell **25:** 2831-2847.

Uren, A.G., O'Rourke, K., Aravind, L.A., Pisabarro, M.T., Seshagiri, S., Koonin, E.V., and Dixit, V.M. (2000). Identification of paracaspases and metacaspases: two ancient families of caspase-like proteins, one of which plays a key role in MALT lymphoma. Mol Cell **6:** 961-967.

van der Hoorn, R.A. (2008). Plant proteases: from phenotypes to molecular mechanisms. Annu Rev Plant Biol **59:** 191-223.

van der Linde, K., Hemetsberger, C., Kastner, C., Kaschani, F., van der Hoorn, R.A., Kumlehn, J., and Doehlemann, G. (2012). A maize cystatin suppresses host immunity by inhibiting apoplastic cysteine proteases. Plant Cell **24:** 1285-1300.

van Wijk, K.J. (2015). Protein maturation and proteolysis in plant plastids, mitochondria, and peroxisomes. Annu Rev Plant Biol **66:** 75-111.

Vanhaeren, H., Chen, Y., Vermeersch, M., De Milde, L., De Vleeschhauwer, V., Natran, A., Persiau, G., Eeckhout, D., De Jaeger, G., Gevaert, K., and Inze, D. (2020). UBP12 and UBP13 negatively regulate the activity of the ubiquitin-dependent peptidases DA1, DAR1 and DAR2. Elife **9**.

Venne, A.S., Solari, F.A., Faden, F., Paretti, T., Dissmeyer, N., and Zahedi, R.P. (2015). An improved workflow for quantitative N-terminal charge-based fractional diagonal chromatography (ChaFRADIC) to study proteolytic events in Arabidopsis thaliana. Proteomics **15:** 2458-2469.

Vercammen, D., van de Cotte, B., De Jaeger, G., Eeckhout, D., Casteels, P., Vandepoele, K., Vandenberghe, I., Van Beeumen, J., Inze, D., and Van Breusegem, F. (2004). Type II metacaspases Atmc4 and Atmc9 of Arabidopsis thaliana cleave substrates after arginine and lysine. J Biol Chem **279:** 45329-45336.

Vercammen, D., Belenghi, B., van de Cotte, B., Beunens, T., Gavigan, J.A., De Rycke, R., Brackenier, A., Inze, D., Harris, J.L., and Van Breusegem, F. (2006). Serpin1 of Arabidopsis thaliana is a suicide inhibitor for metacaspase 9. J Mol Biol **364:** 625-636.

Wang, L., Einig, E., Almeida-Trapp, M., Albert, M., Fliegmann, J., Mithofer, A., Kalbacher, H., and Felix, G. (2018). The systemin receptor SYR1 enhances resistance of tomato against herbivorous insects. Nat Plants **4:** 152-156.

Wang, Y., Garrido-Oter, R., Wu, J., Winkelmuller, T.M., Agler, M., Colby, T., Nobori, T., Kemen, E., and Tsuda, K. (2019). Site-specific cleavage of bacterial MucD by secreted proteases mediates antibacterial resistance in Arabidopsis. Nat Commun **10:** 2853.

Watanabe, N., and Lam, E. (2011). Calcium-dependent activation and autolysis of Arabidopsis metacaspase 2d. J Biol Chem **286:** 10027-10040.

Wen, S., Ma, Q.M., Zhang, Y.L., Yang, J.P., Zhao, G.H., Fu, D.Q., Luo, Y.B., and Qu, G.Q. (2013). Biochemical evidence of key residues for the activation and autoprocessing of tomato type II metacaspase. FEBS Lett **587:** 2517-2522.

Woo, J., Park, E., and Dinesh-Kumar, S.P. (2014). Differential processing of Arabidopsis ubiquitin-like Atg8 autophagy proteins by Atg4 cysteine proteases. Proc Natl Acad Sci U S A **111:** 863-868.

Yamada, K., Yamashita-Yamada, M., Hirase, T., Fujiwara, T., Tsuda, K., Hiruma, K., and Saijo, Y. (2016). Danger peptide receptor signaling in plants ensures basal immunity upon pathogen-induced depletion of BAK1. EMBO J **35:** 46-61.

Zauner, F.B., Dall, E., Regl, C., Grassi, L., Huber, C.G., Cabrele, C., and Brandstetter, H. (2018). Crystal Structure of Plant Legumain Reveals a Unique Two-Chain State with pH-Dependent Activity Regulation. Plant Cell **30:** 686-699.

Zhang, H., et al. (2018). A Plant Phytosulfokine Peptide Initiates Auxin-Dependent Immunity through Cytosolic Ca(2+) Signaling in Tomato. Plant Cell **30:** 652-667.

Zhang, J., et al. (2010). Receptor-like cytoplasmic kinases integrate signaling from multiple plant immune receptors and are targeted by a Pseudomonas syringae effector. Cell Host Microbe **7:** 290-301.

Zhang, X., Yang, Z., Wu, D., and Yu, F. (2020). RALF-FERONIA Signaling: Linking Plant Immune Response with Cell Growth. Plant Commun **1:** 100084.

Zhu, P., Yu, X.H., Wang, C., Zhang, Q., Liu, W., McSweeney, S., Shanklin, J., Lam, E., and Liu, Q. (2020). Structural basis for Ca(2+)-dependent activation of a plant metacaspase. Nat Commun **11:** 2249.

Ziemann, S., et al. (2018). An apoplastic peptide activates salicylic acid signalling in maize. Nat Plants **4:** 172-180.

Zipfel, C., Robatzek, S., Navarro, L., Oakeley, E.J., Jones, J.D., Felix, G., and Boller, T. (2004). Bacterial disease resistance in Arabidopsis through flagellin perception. Nature **428:** 764-767.

Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D., Boller, T., and Felix, G. (2006). Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts Agrobacterium-mediated transformation. Cell **125:** 749-760.

Chapter 3

"Caught green-handed: methods for *in vivo* detection and visualization of protease activity."

Álvaro Daniel Fernández-Fernández [1, 2], Renier A. L. van der Hoorn [3], Kris Gevaert [4, 5], Frank Van Breusegem [1, 2 *] and Simon Stael [1, 2, 4, 5 *]

[1] Department of Plant Biotechnology and Bioinformatics, Ghent University, Technologiepark 927, 9052 Ghent, Belgium

[2] VIB Center for Plant Systems Biology, Technologiepark 927, 9052 Ghent, Belgium

[3] The Plant Chemetics Laboratory, Department of Plant Sciences, University of Oxford, Oxford OX1 3RB, UK

[4] Department of Biomolecular Medicine, Ghent University, B9000 Ghent, Belgium

[5] VIB Center for Medical Biotechnology, B9000 Ghent, Belgium

* Corresponding author

Author contributions:

ADFF and SS wrote the manuscript with the help FVB, KG and RVdH. ADFF designed the figures. SS, FVB, KG and RVdH corrected the final version of the manuscript.

Aim and context

This review tries to compile the different tools that allow detection of protease activity and its visualization methods. In the manuscript the mechanisms of the reporters, the premises used for their engineering and their advantages and weaknesses are addressed. Additionally, possible alternatives to surpass obstacles during experimental setups and design of proteolytic reporters are considered. We believe that our manuscript will help plant scientists to understand the possibilities available to detect protease activity *in vivo* and it will help in the selection of the most suitable method for each case study.

Abstract:

Proteases are enzymes that cleave peptide bonds of other proteins. Their omnipresence and diverse activities make them important players in protein homeostasis and turnover of the total cell proteome as well as in signal transduction in plant stress response and development. To fully understand protease function, it is of paramount importance to assess when and where a specific protease is active. Here, we review the existing methods to detect *in vivo* protease activity by means of imaging chemical activity-based probes and genetically encoded sensors. We focus on the diverse fluorescent and luminescent sensors at the researcher's disposal and evaluate the potential of imaging techniques to deliver *in vivo* spatiotemporal detail of protease activity. We predict that in the coming years, revised techniques will help to elucidate plant protease activity, functions and hence expand the current status of the field.

Keywords: activity-based probes, *in vivo* imaging, fluorescent, luminescent, protease, plants, proteolysis, sensor, reporter.

Introduction

Proteases exert a tight control on cellular functions by breaking the polypeptide chain of their substrate proteins. Substrate cleavage is the result of the recognition between the target amino acid sequence of a substrate and the binding pocket of a protease. This action can lead to changes in the localization, biomolecular interactions, turnover or enzymatic activity of their substrates. The importance of proteases across biological kingdoms and viruses cannot be overstated as they intervene in most developmental processes and responses to environmental cues (Turk, 2006; van der Hoorn, 2008). Proteases are also quite numerous, for example *Arabidopsis thaliana* has a reviewed number around 600 proteases, totalling to 2% of the protein coding genes, similar in number and family conservation to other plant species like rice and poplar (García-Lorenzo *et al.*, 2006; van der Hoorn, 2008; Lallemand *et al.*, 2015).

Proteases are classified based on the catalytic amino acid inside their binding pocket, being cysteine, serine, aspartic acid, threonine or glutamic acid, whereas metalloproteases use a coordinated metal ion to catalyze peptide bond hydrolysis. In cysteine, serine and threonine proteases, the peptide bond is broken after nucleophilic attack of their respective catalytic amino acid aided by a molecule of water on the carbonyl group of the peptide bond. For metallo-proteases, aspartic and glutamic acid proteases, cleavage occurs by activation of a water molecule which then performs the nucleophilic attack. Asparagine peptide lyases, are also capable of peptide bond cleavage by an elimination reaction, and their classification as proteases is still under debate given the differences in their catalytic mechanism (Rawlings *et al.*, 2011). To the date, such enzymes have not been identified in plants.

The actual protease recognition site in the substrate is generally defined by the amino acids surrounding the scissile peptide bond and these amino acids are denoted as P and P' ($P_4P_3P_2P_1{\downarrow}P'_1P'_2P'_3P'_4$) where the downward arrow indicates the peptide bond cleavage (Schechter and Berger, 1968). Proteases can show narrow substrate selectivity, like thrombin and its minimal substrate LVPR↓GS (Gallwitz *et al.*, 2012), to have a broader substrate selectivity such as trypsin that cleaves after arginine and lysine (Olsen *et al.*, 2004). Many proteases have a certain degree of permissibility in their substrate sites such as the Tobacco Etch Virus protease, which recognizes the

optimal substrate sequence ENLYFQ↓S but also recognizes EXLYXQ↓X sequence where X denotes amino acids with similar properties to those found in the optimal substrate sequence (Boulware *et al.*, 2010; Sandersjöö *et al.*, 2017). Considering protease specificity and activity redundancy over a common substrate are of main importance when designing a probe for *in vivo* applications. In general, most part of the substrates should be as specific as possible for a single protease and in some particular cases the detection of a processed substrate can confirm entrance in specific biological processes. This is well exemplified by caspase-3 and caspase-7 which are cysteine proteases that cleave after aspartic acid (D) and are considered triggers of apoptosis.

To understand the *in vivo* relevance of protease activities, it is important to know their spatiotemporal activity profiles. Gene expression and protein distribution patterns in plant tissues can certainly offer some clues to the role of a particular protease. However, their activity is often strictly restrained following their synthesis. For most proteases, initial protein synthesis produces an inactive or zymogenic proteoform. Subsequently, a first layer of control is exerted by zymogen activation during which for instance an inhibitory propeptide is catalytically removed, as it is the case for subtilisin-like serine proteases or subtilases in plants (Meyer *et al.*, 2016). Other proteases lack clear inhibitory propeptide signals and instead require cleavage of internal sites for their activation. One example is the cleavage of the linker region between the p20 and p10 domains of metacaspases and caspases (Salvesen *et al.*, 2016). Additional internal cleavages or maturation of proteases can further affect their activity (Gu *et al.*, 2012). Protease activation can also occur following a switch in the microenvironment such as altered pH, elevated calcium levels or other activating conditions, such as proximity induced multimerization (Lam and Zhang, 2012). Post-translational modifications (PTMs) may further determine a protease's activation state. For example, the catalytic cysteine of Arabidopsis metacaspase 9 (AtMC9) is reversibly regulated by S-nitrosylation that blocks its activity (Belenghi *et al.*, 2007). Other proteases in Arabidopsis such as ClpP1, Deg2 and ClpC2 were also found to be S-nitrosylated, but the importance of this modification is not clear (Romero-Puertas *et al.*, 2008; Hu *et al.*, 2015). Proteases are also regulated by proteinaceous inhibitors (protease inhibitors or PIs) that bind the catalytic site, thereby blocking the capability to process their substrates (Grosse-Holz and van der Hoorn, 2016). Such interactions can be

reversible, as it occurs for certain cysteine proteases and their inhibitors, cystatins (Benchabane *et al.*, 2010), or irreversible, for example SERPIN1 that traps both AtMC9 and RD21A (Vercammen *et al.*, 2006; Lampl *et al.*, 2013; Grosse-Holz and van der Hoorn, 2016). PIs themselves are subject to regulation under various stress or developmental conditions, and a given protease can be regulated by multiple PIs (Rustgi *et al.*, 2018), thereby further increasing the repertoire to fine-tune regulation of protease activities. Lastly, proteases localize to all sub-compartments in cells (van Wijk, 2015) and protease activity can depend on relocation of proteases in cells. For example, phytaspase delocalizes from the extracellular to the intracellular space (Chichkova *et al.*, 2010) and vacuolar processing enzymes first need to be released into the cytosol during programmed cell death (Hatsugai *et al.*, 2015).

Researchers studying proteases often use peptidic probes that report proteolytic activity. Ideally, such probes are used *in vivo* to evaluate protease activities in their natural cell environment, while delivering spatiotemporal resolution. Over the years, increasing knowledge of protease substrate specificity has allowed to develop better tools for studying protease activity. Here, we review the use of chemical probes and genetically encoded sensors in plants and future expectations. We then widen the horizon towards techniques available for *in vivo* spatiotemporal protease activity detection in the mammalian protease field, which are expected to be implemented in plant cells.

CHEMICAL PROBES

Chemical activity-based probes

Activity-based probes (ABPs) are chemical probes capable of detecting the catalytic enzymatic site of an enzyme (Heal *et al.*, 2011). ABP can be synthesized from a known chemical inhibitor which is generally termed warhead, followed by a linker that can resemble an amino acidic sequence allowing a certain degree of specificity in the design and finally a tag that can be used with different purposes such as pull-down or imaging. These chemical probes mimic a substrate cleavage/scissile bond and lock the proteolytic mechanism in a covalent intermediate stage (Sanman and Bogyo, 2014). Inherently, ABPs only exist for proteases that react with their substrate through a covalent intermediate like for cysteine, serine and threonine proteases, but not for

aspartic, glutamic acid and metallo-proteases (van der Hoorn, 2008). Probes for the latter classes are photo-affinity probes based on reversible inhibitors and require UV exposure to establish a covalent bond (Li *et al.*, 2000). ABPs entail a reporter tag to facilitate detection. Fluorescent reporter tags are practical tools for cell biology studies, usually permit cell entry, but do not report on the identity of the labelled proteases. By contrast, biotinylated probes can be readily used to purify and identify the labelled proteases by means of mass spectrometry, but the entry of biotinylated probes into the cell is often problematic. To overcome this issue, chemical moieties called "minitags" can be coupled to e.g. fluorophores or biotin using a bio-orthogonal chemical coupling reaction, such as click chemistry between alkynes (C≡C) and azides (N_3) (Speers and Cravatt, 2004; Kaschani *et al.*, 2009b). Minitagged probes are efficient to cross cellular membranes, reason why they are frequently used for *in vivo* labelling. However, minitags detection usually involves a copper-catalysed coupling reaction in the protein extraction (*ex vivo*). Other two-step labelling processes include the Staudinger-Bertozzi ligation and the Diels-Alder ligation (Verdoes and Verhelst, 2016). Using specific coupling protocols and reagents, it is possible to perform the coupling reaction *in situ* like in fixed tissues (Paper *et al.*, 2018), or even *in vivo* (Chang *et al.*, 2010).

Fluorescently labelled proteins can be visualised in protein gels after electrophoresis (profiling) or by microscopic imaging (Fig. 1A). Although nearly all fluorescent ABPs can be used for *in vivo* labelling and profiling, imaging makes little sense if the probes are not sufficiently specific. For instance, active serine proteases can be efficiently profiled *in vivo* with fluorophosphonate probes (Liu *et al.*, 1999; Kaschani *et al.*, 2009a). However, these probes label over 50 different proteins, making life-cell imaging with such probes not very informative due to redundancy of activity over one ABP. In fact, more selective and specific probes are better suited for activity-based imaging. For instance, three fluorescent ABPs targeting the proteasome revealed quick labelling from 1 minute for subunits β2 and β5, while subunit β1 was labelled in a timeframe of around one hour (Kolodziejek *et al.*, 2011).The probes were designed based on known proteasome inhibitors from vinyl sulfones, epoxomicin and syringolin A. Another probe designed for detection of vacuolar processing enzyme probes displayed fluorescence in the vacuole of plant cell cultures and leaves while this signal was absent in vacuolar processing enzyme null mutants (Misas-Villamil *et al.*, 2013).

An important technical limitation of fluorescently labelled probes is that they fluoresce also when not bound to a target and may thus cause background signals that obscure detection of labelled proteins. This problem is often solved by washing out non-reacted probes or by using low probe doses such that eventually all probes are immobilized on their targets. An elegant alternative solution are quenched probes, which have been described for various cysteine proteases but were mostly used in mammalian systems (Blum *et al.*, 2005; Edgington *et al.*, 2009; Verdoes *et al.*, 2012; Edgington *et al.*, 2013). These probes contain a quenching group that suppresses the fluorophore emission of light until the probe interacts with the protease (Edgington-Mitchell *et al.*, 2017).
Although activity-based imaging also has limitations because these probes inactivate the labelled proteases, and that some labelling time (up to one hour) is needed to achieve labelling, the great advantage of these probes is that after imaging, the labelled proteins can also be detected upon separation of extracts on protein gels to identity the labelled proteins.

Chemical protease substrate probes

Together with ABPs, the utilization of fluorescent chemical reagents that react to protease activity can be of interest for the plant field. For example, DEVD-NucView488 (Fig. 1B) can be used to visualize caspase by nuclei labelling (Cen *et al.*, 2008). It contains a DNA intercalating reagent similar to thiazole orange that react with a negatively charged group of DEVD hiding fluorescence. This motif is cleaved by caspases after the second aspartic acid (DEVD↓) leading to subsequent DNA binding of the reagent and thereby inducing detectable fluorescence. When indole-3-carbinol is applied to trigger apoptosis to cell cultures, the nuclei of these cells are thereby fluorescently labelled. In plants, caspase-3 and -7-like activities have been detected in the cytosol, generative cell and vegetative nuclei of *Papaver roheas* pollen (Bosch and Franklin-Tong, 2007, 2008) using CR-$(DEVD)_2$. In this probe, two DEVD peptides were added to cresyl violet, which quenches its fluorescence (Fig. 1C). DEVDases and proteases with caspase-3/7-like activities release the DEVD peptides, uncovering the compound's fluorescence. With this probe, caspase-3-like activity was detected in the vegetative nucleus, cytosol and generative cells of pollen tubes of poppy plants during induction of self-incompatibility. Despite the fact that other proteases are known to be involved in pollen programmed cell death (PCD) in Arabidopsis (Zhang *et al.*, 2014), the actual poppy plant proteases involved remain unidentified. These probes

were further developed to include infrared fluorophores that can be used to image in deeper tissues of live animals and are now developed to image malignant tumours during surgery (Blum *et al.*, 2007; Edgington *et al.*, 2009; Verdoes *et al.*, 2012).

DIRECT-FLUORESCENT SENSORS OF PROTEOLYTIC ACTIVITY

Most of the existing proteolytic sensors are based on modifications of eGFP (Fig. 2A and 2B) and its derived fluorescent proteins. The use of these proteins can be advantageous due to their modifications in emission and excitation parameters. It is also a point to consider that other proteins can outperform eGFP in specific environments like super folding versions of GFP, Cherry2, or split mNeonGreen2 (Feng *et al.*, 2017) and that researchers can profit of them to engineer protease activity sensors in the future.

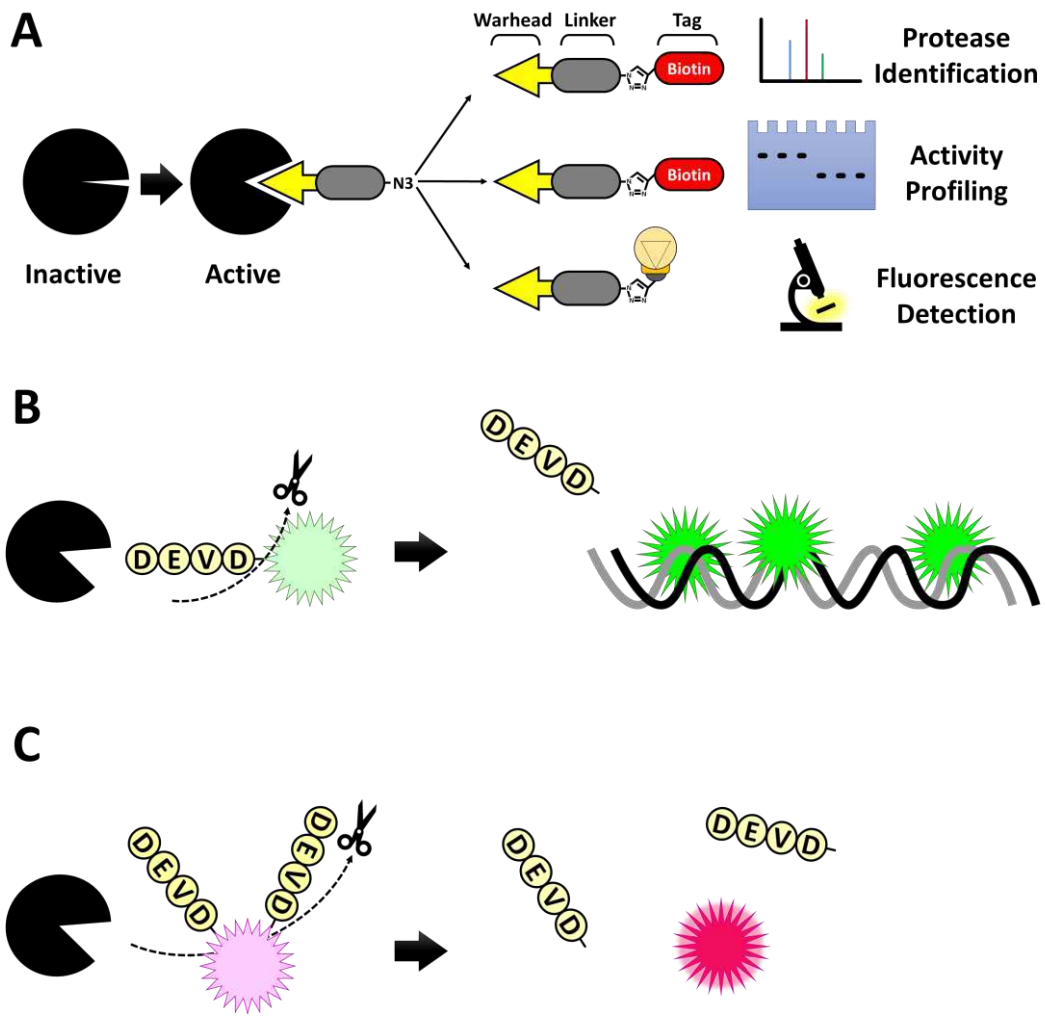


Figure 1. Representation of chemical probes used for protease detection. A) Different uses of ABPs and several possibilities to apply for protease detection combined with MS, gel detection of active

proteases and imaging techniques using fluorescence microscopes. **B)** DEVD-NucView488 mode of action. A carboxyl derivative of thiazole orange is used to synthesize the probe that contains an acetylated tetrapeptide which hampers the dye reaction with DNA. Once it is released in can report signal, corresponding to caspase-3 activity. **C)** Representation of Rationale for CR-$(DEVD)_2$ for the detection of caspase-3 like activity. The figure depicts quenched fluorescence of cresyl violet (CR) by two DEVD caspase-3 substrates which are bound to. After processing by activated caspases, the peptides are released and cresyl violet turns fluorescent serving as a readout of protease activity.

FRET sensors

Förster Resonance Energy Transfer (FRET) applications are nowadays employed in almost every field of cell biology and have the longest history as genetic protease reporters (Weiss, 2000). FRET is based on energy transfer between two fluorescent proteins or particles in a spatial contiguity and orientation. Additionally, the emission wavelength range of one particle (donor) overlaps with the excitation of the second particle (acceptor). Upon excitation of the donor, it transmits part of its energy to the acceptor, which is then excited and emits fluorescence. Perturbations in the distance between the donor and acceptor results in changes in the fluorescence properties of both components. FRET sensors for proteases include a sequence that resembles the cleavage motif of a protease, keeping both fluorophores proximal until cleavage occurs. Once the proximal position of the fluorescent component is lost, the acceptor signal decreases and the donor signal enhances (Fig. 2C). FRET proteolytic biosensors include detection of Xa protease (Mitra *et al.*, 1996), thrombin (Zhang, 2004), trypsin (Heim and Tsien, 1996), caspase-1(Mahajan *et al.*, 1999), caspase-3 (Xu *et al.*, 1998; Luo *et al.*, 2001) and caspase-8 (Luo *et al.*, 2003), anthrax lethal factor protease (Kimura *et al.*, 2007), matrix metalloprotease (MMP) 1 (Ouyang *et al.*, 2008), MMP-2 (Yang *et al.*, 2007) and MMP-9 (Stawarski *et al.*, 2014); ADAM 17 protease (Chapnick *et al.*, 2015), neutrophil elastase (Schulenburg *et al.*, 2016), Hepatitis C Virus (HCV) NS3 protease (Sabariegos *et al.*, 2009), calpain (Vanderklish *et al.*, 2000; Stockholm *et al.*, 2005) and MALT1 ((Schairer *et al.*, 2020)).

FRET reporters are made of two fluorescent proteins with complementing features in emission and absorbance. Some improvements were made aimed at increasing FRET efficiency. For example, two mutations in the surface between the interaction of a YFP variant and a Cyan Fluorescent Protein (CFP) increases FRET 4-fold compared to the original versions of these proteins (Vinkenborg *et al.*, 2007). More recently, novel

FRET sensors included the use of weak interactor peptides at the fluorescent proteins, bringing them spatially together and enhancing energy transfer (Grunberg *et al.*, 2013). This set of sensors was named helper-interaction FRET and proteolytic sensors were arranged to recognise caspase activity using LDEVD as linker between mTurquoise2, and mCitrine, which are improved versions of CFP and YFP respectively. The position of the weak interactor peptides was tested at the N- and C-termini of the sensor or included in structural loops of the fluorescent proteins. Upon caspase-3 activity, the linker is cleaved, and despite the helpers, the proximity is lost and the protease acceptor loses its fluorescence in higher ratios than common FRET sensors that did not include the interaction helpers.

In plants, FRET sensors are routinely used for the detection of protein-protein interactions and intracellular signals such as calcium levels, abscisic acid, pH levels or ATP concentration and have been extensively reviewed in (Grossmann *et al.*, 2018). However, reports on successful use of genetically encoded sensors for plant proteases are scarce. Zhang and colleagues used FRET based sensor with an amino acid linker, separating the two chromophores, containing a DEVD sequence for the detection of caspase-3-like activity in Arabidopsis protoplasts when exposed to UV-C light (Zhang *et al.*, 2009). Exposure to UV-C light for less than half an hour resulted in a reduced FRET signal with a maximum at 1 hour after treatment, indicating a relatively high activation by caspase-3-like proteases. The reduction in signal observed was not detected in the negative control reporter that used the DEVG linker sequence that is not recognized by caspase-3-like protease. Caspase-3-like activity is a recurrent topic in the field of plant proteases although there are no genuine caspase genes found in plant genomes, and though inhibitors of caspases generally block plant PCD (Sueldo and van der Hoorn, 2017). A FRET based biosensor was also used in plant protoplasts for the detection of a latent peptidase activity of the ubiquitin-activated peptidase DA1 (Dong *et al.*, 2017). The approach used eGFP and mCherry flanking the whole sequence of DA1 substrate, Big Brother instead of the conventional YFP/CFP pair. An increase in the green fluorescence lifetime, which in this case corresponded to the donor eGFP, was shown when DA1 was co-expressed, indicating that mCherry part is distant and that FRET decreased.

Apart from reporting protease activity *in vivo* and *in vitro,* FRET reporters can serve as markers of certain pathways like apoptosis and to screen for inhibitors of enzymatic activities (Jones *et al.*, 2000; Zhang, 2004). Another FRET application is the

generation of randomized substrate libraries to detect protease specificity (Fretwell *et al.*, 2008) as a way of designing fluorescent probes that are as specific as possible for the protease of interest.

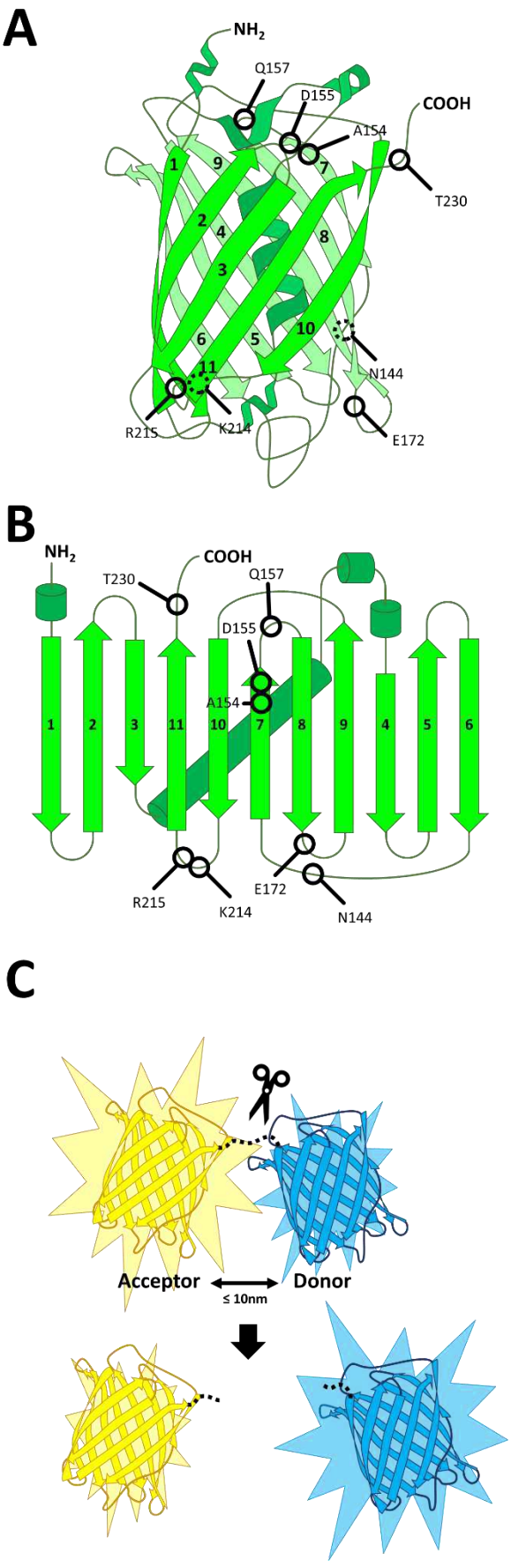


Figure 2. Fluorescent proteins at the rescue of proteolytic activity. A) Schematic representation of the β-barrel fold of eGFP. β-barrel numbers are indicated over the structure. Important residues including in different experiments for the development of proteolysis-sensitive version of eGFP are indicated with a circle. When the specific residues are behind a strand in the spatial dimension the circle is displaced with a discontinuous line. Asparagine at position 144 (N144), glutamine at position 157 (Q157) and glutamic acid at position 172 (E172) were used for the grafting experiments including processing sites in the loop 7, 8 and 9 of eGFP respectively. A154 and D155 are commonly used for splitting GFP in both C- and N-terminal and used in BiFC assays and VC3AI and derivate proteolytic sensors. Lysine 214 (K214) is generally used to delimitate GFP1-10. Its complementing part GFP11 starts at arginine at site 215 (R215) and lacks the last 8 amino acids being a threonine (T230) the last of them as used in ZipGFP and ProGFP. NH_2 and COOH indicates the N- and C-terminal ends of eGFP. **B)** Cartoon representing the secondary structure of GFP. The corresponding amino to the design of protease sensitive eGFP are also indicated in this cartoon. **C)** Example of a FRET sensor using CFP and YFP. Both proteins are part of a single molecule joined by a linker containing a cleavable peptide. Generally, only the donor fluorescent protein is excited which uses energy to transmit to the acceptor. When proteins are not in a spatial proximity, this energy transfer is lost, the donor fluorescence increases and the acceptor fluorescence decreases as indicated by the background stars. The maximal indicated distance to emit fluorescence is generally accepted as 10 nm of distance between fluorescent proteins as indicated in the figure.

Fluorescence complementation

The newest generation of genetic fluorescent protease reporters use fluorescence complementation. This mechanism relies on split versions of fluorescent proteins with null or decreased fluorescence emission, which in proximity can complement and reconstitute fluorescence (Ohad *et al.*, 2007). Although this technique has been mainly exploited for the detection of protein-protein interaction such as bimolecular fluorescence complementation (BiFC) it can be tuned for the detection of proteolytic activity. Here we include designs where both parts are designed as separated segments to recompose fluorescence and rearrangements in the amino acid composition that disturb protein structure or destabilizes chromophore formation. To detect protease activity, it is necessary that these reporters include a protease recognition sequence that partially reconstitutes or enhances fluorescence after processing (Fig. 3). Some of these sensors are conceived as individual sensors (Fig. 3A-C and 3E), others have a stable additional fluorescent signal, independent of proteolysis, to obtain a ratiometric signal (Fig. 3D). Structural loops are candidate regions in which protease cleavage sites can be introduced, though probe design has to be carefully considered and success can depend on both the protease and the cleavage site used. For instance, eGFP loops 7, 8 and 9 were modified to include a motif that is known to be cleaved by trypsin (Chen *et al.*, 2009b). The design of the sequences, including protein grafting at several locations shows a better response when incubated with the protease. For example, fluorescent increased after incubation with trypsin *in vitro* and after trypsin induction with caerulein in pancreatic cell cultures. The same design showed positive results when studying caspase-3 activity by crafting DEVD-motifs, overall indicating that previous mentioned glutamic acid at loop 9 might be well suited to develop new protease reporters upon implementing minimal substrates (Chen *et al.*, 2009a) (Fig. 3A).

More recently, Callahan and colleagues modified a split GFP protein reporter system (Kamiyama *et al.*, 2016) by partially caging a GFP11 β-strand that has a high affinity for the remaining GFP1-10 β-strands (Callahan *et al.*, 2010). In this reporter named Pro-GFP, the GFP11 β-strand replaces the exposed loop of the protease inhibitor eglin c with a proteolytic site at the C-terminal part of this β-strand. Both inhibitory domains are highly stable and remain together even after cleavage however, cleavage

concedes enough flexibility for GFP11 to complement GFP1-10 (Fig. 3B). This sensor was also tailored for the detection of thrombin, caspase-3 and HIV proteolytic activity, showing a 6-fold increasing in signal when the proteases were present.

VC3AI is a single protein biosensor for caspases 3 and 7 (Zhang *et al.*, 2013). VC3AI stands for Venus Caspase-3 Activity Indicator shaped by a circularly permuted YFP protein with a F46L mutation called Venus. In its design, the original design of N- and C-halves of Venus are inverted and ligated using a typical caspase cleavage site. VC3AI is flanked by inteins (Mills *et al.*, 2014), which are peptides that can self-excise leading to association of the remaining adjacent sequences, in this case both ends of the permuted sensor which is then circularized. The structure of the circular protein constricts the Venus moieties in a format with decreased fluorescence, which, after cleavage, is loosened up and shows higher levels of fluorescence (Fig. 3C). Detection of apoptotic events in tumour cell culture showed a 10-fold fluorescence increase within 8 hours after induction of apoptosis. Here, DEVD processing was also observed in non-apoptotic cells, which may lead to exploring new functions of caspases apart from cell death, and exemplifies the power of generating new biological hypotheses with these sorts of probes. Versions of this sensor using blue, green and red fluorescent proteins, termed respectively, CC3AI, GC3AI and RC3AI, also successfully reconstituted fluorescence upon cell death induction. This extended colour palette opens diverse possibilities to combine with several cell dyes or markers indicating cell viability and to exploit the intrinsic characteristics of each fluorescent protein. In a further study, such different probes for detecting caspase-3 activity in fruit flies were reported, thereby increasing the applications of the caspase reporter toolkit (Schott *et al.*, 2017).

iProteases are a group of sensors in which the two domains of an Infrared Fluorescent Protein are truncated and fused both N- and C-terminal to self-complementing split eGFP which here acts as the ratiometric signal (To *et al.*, 2015). The domains in the truncated version keep a crucial cysteine distant enough from the site of chromophore formation, that is mediated by biliverdin. This cysteine displacement avoids thioether formation between biliverdin in the binding site and the cysteine, and therefore blocks fluorescence. Using a linker holding specific substrates for TEVp, caspase-3/7 and HCV protease, respective iSensors (iTEV, iCasper and iHCV) were obtained, showing how a single sensor design can be turned into a potent detector of dynamics for

different proteases. The analysis of the cleaved purified sensor showed an increase in fluorescence ten seconds after biliverdin addition, detecting immediate response which can be used for quick readout of proteolysis when biliverdin is available. Results *in vivo* indicated detection of apoptosis events in Drosophila in a frame time of 1 hour after staurosporine addition, which is a protein kinase inhibitor which induces apoptosis including caspase activation. The presence of a ratiometric permanent green signal allowed to detect transfected cells independently of protease activity.

Another ratiometric sensor, dubbed ZipGFP was developed for detection of proteases action, where eGFP was first split and interaction of the separate parts was blocked by zipping together the C- and N-terminal ends of β-strands 1 to 10 and β-strand 11 with heterodimers of E5 and K5 coils (To *et al.*, 2016). Ratiometry was achieved by N-terminal fusion of eGFP β-strands 1-10 to mCherry with a T2A peptide for self-cleavage by ribosomal skipping (Szymczak *et al.*, 2004), producing equimolar amount of protease reporter and mCherry (Fig. 3D). Cleavage of TEVp and caspase-3 substrate sites liberated the E5/K5 heterodimers allowing the cavity of the main structure to be complemented both *in vitro* and during apoptosis of zebrafish embryos. ZipGFP outperformed previously described FRET sensors given a 10-fold increase of eGFP signal, which rarely could be detected by energy transfer (To *et al.*, 2016).

Single fluorescently labelled sensors

Various sensors use translational fusions to proteins or protein domains that hamper GFP folding, and therefore fluorescence, or reduce lifetime of the sensor due to their instability in a particular cellular environment. For instance, Caspase Activatable GFP sensors (CA-GFP) are based on this concept (Nicholls *et al.*, 2011). Here, eGFP is fused to a 27 amino acid stretch of the monomer of the homotetrameric influenza M2 proton channel protein, bridged by a DEVD sequence. The M2 domain also tetramerized in CA-GFP, thereby hindering correct maturation of the eGFP chromophore and reducing its fluorescence (Fig. 3E). Caspase cleavage freed this quenching domain, thereby enhancing fluorescence 45-fold in bacteria, while only 3-fold in mammalian cell culture. These results illustrate that protease reporters can be highly system-dependent and that they need to be tested and optimized for each organism or cellular context. Interestingly, N-terminal fusions of the M2 domain did not quench eGFP fluorescence while C-terminal fusion gave 40-fold increases for caspase-3/7 (Nicholls and Hardy, 2013). YFP, Cerulean and mNeptune versions were

later developed with varying results, probably because of the variable spatial conformation and protein maturation of the different fluorophores (Wu *et al.*, 2013). Similarly, the fusion of GFP to a C-terminal bacterial degron ssrA (AANDENYNYALAA) through different substrates of TEVp showed preference for ENLYFQ↓G over other minimal substrates in *Escherichia coli* (Kostallas and Samuelson, 2010; Kostallas *et al.*, 2011). Here, ssrA is an optimized peptide that induces fast protein turnover of the sensor via the bacterial degradation system shaped by the caseinolytic proteinase XP complex. This knowledge about GFP structure was exploited to permute the last β-strand at the beginning of the sensor that includes proteolytic cleavable sites at the loops between the original β-strands 6 and 7 or β-strands 7 and 8. The circularly permuted GFP reported 50 times faster degradation of Lon bacterial substrates *in vivo* than native GFP, possibly this is due to the decreased thermal stability of the permuted sensors (Wohlever *et al.*, 2013).

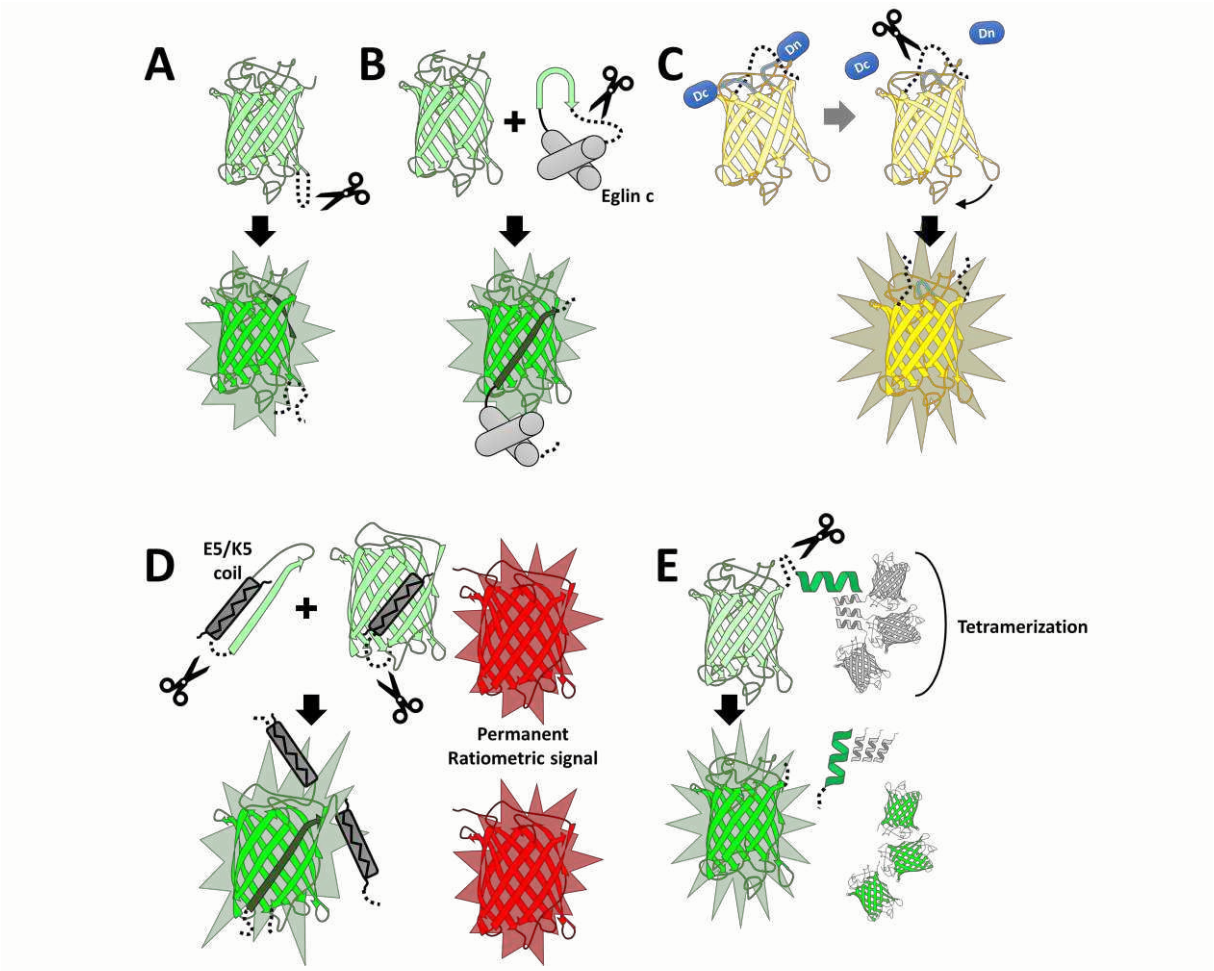


Figure 3. A palette of genetically encoded fluorescent reporters for the detection of proteolytic activity. Representation of the different mechanisms employed by genetically encoded sensors of protease activity based on fluorescent proteins. Fluorescent proteins are depicted in the colour of their

maxima spectra emission and cleavage sites are indicated by discontinuous thick black lines with scissor indicating cleavage and by open dotted lines after processing. **A)** Representation of a grafted version of eGFP**.** By addition of a cleavable recognition site in a eGFP loop after a glutamic acid in position 172, one of the β-strands of eGFP is misplaced affecting the overall final protein conformation. The position of the amino acids indispensable for chromophore formation is in this manner sufficiently distant to lower eGFP efficiency. Once the loop constringency is loosened up by protease activity, the β-strand can relocate to its natural position. Note that other two amino acids in previous loops were used for this experiment as indicated in . **B)** The Pro-GFP sensor. Pro-GFP is based on the enclosure of the β-11 strand of eGFP using a proteolytic resistant protein domain of eglin c containing two alpha-helices (indicated as grey cylinders) and a cleavable site. While the protein containing β-strands 1-10 is not capable of producing fluorescence, complementation reconstitutes eGFP conformation and fluorescence. **C)** Venus Caspase 3 Activity Indicator (VC3AI). VC3AI is based on the constriction and modification of the structure of a YFP variant. The fluorescent protein is permuted by generating new N- and C-terminal ends at A154 and D55 and the original termini are linked using a protease recognition site showed by the black dotted line. To circularize the sensor Dc and Dn fragment of *Npu* DnaE intein are added to the new termini which can self-release in a process known as protein splicing and lead to circularization of the sensor (indicated by the horizontal grey arrow). In this conformation the fluorescent protein is not capable to efficiently reassemble with itself or other permuted circularize proteins. Once that the linker is processed complementation of the molecule occurs and the fluorescence is reconstituted as indicated by the black arrow. **D)** ZipGFP ratiometric protease detector. ZipGFP contains the β-strands 1-10 of eGFP linked by a 2A peptide to mCherry. Those sequences are efficiently processed by ribosomal skipping of the 2A-like peptide resulting in a constitutive and continuous ratiometric red signal. The β-strands 1-10 and 11 of eGFP contain additional extensions incorporating binding peptides that knot the structure impeding unspecific reconstitution by both parts. The peptide position can be relaxed by cleavage of one of the linkers re-establishing GFP as a full functional reporter. **E)** Caspase Activatable GFP (CA-GFP). C-terminal fusion of eGFP to the 28 amino acids of the intermembrane domain of influenza matrix protein 2 leads to tetramerization of the sensor hampering maturation of the chromophore. Upon proteolysis, the multimeric dependence of GFP is broken and four molecules of GFP can mature and emit fluorescence.

Proteolysis-mediated delocalization sensors

This class of sensors exploits protein localization (and re-localization) to read out cleavage by proteases. The simplest design comes from the observation of changes in cellular compartmentalization using single fluorescent fusion proteins. However, some minimal conditions are required: the fluorescent tag needs to be attached to the mobile component and this tag may not interfere with specific signals or transit peptides present in the protein addressing correct localization. In the cell content, some transcriptional effectors are targeted to other location than nucleus in a state

that requires to undergo PTMs for transport to the nucleus where they can exert their function (Andréasson *et al.*, 2006; Iwata *et al.*, 2017). In Arabidopsis, NAC Transmembrane Motif 1 is processed by an intramembrane protease, likely phytocalpain, loses its transmembrane domain and relocates to the nucleus (Kim *et al.*, 2006). The processing was visualized by GFP fused to whole NAC factor and in the truncated versions of the protein, mimicking cleavage site but blocked when the full protein tagged is incubated with specific phytocalpain inhibitors.

Dual-tagged substrate proteins provide information for both parts of a cleaved construct. An example in *Arabidopsis thaliana* is mCherry/GFP-tagged NAC017 at its N- and C-site respectively. Proteolysis occurs after application of antimycin A, an inducer of mitochondrial retrograde signalling. Initially, both GFP/mCherry reside at the endoplasmic reticulum membrane, but the mCherry signal moves to the nucleus after treatment, while GFP remains at the original position. Although no additional experiments to determine the identity of the responsible protease have been performed, the readout can be used as retrograde signalling markers when antimycin A is used (Ng *et al.*, 2013).

A clear example of de-localization based reporters are caspase biosensors containing an N-terminal Nuclear Export signal (NES) sequence followed by a caspase recognition site, YFP and a C-terminal Nuclear Localization Signal (NLS) C-terminal sequence (NES-DEVD-YFP-NLS), being the NES detachable by caspases (Tang *et al.*, 2012). YFP is present in the cytosol and nucleus in normal conditions but turns mainly nuclear after treatment with apoptotic inducers. Something to take into consideration with these sensors is that they only work when the protease and the sensor can meet each other in the same subcellular space, in this case cytosol and/or nucleus. An elegant application of dual fluorescent sensors is the demonstration of elusive separase activity in cohesin cleavage, necessary for sister chromatid disengagement. A cohesin domain was doubly tagged with mCherry and eGFP at its N- and C-terminus respectively. Additionally, to ensure co-localization with separase to its active cell cycle phase, the sensor was cloned in frame with a centromeric located protein (CENP-B) or histone marker for chromosomal localization (H2B). Separase processing of its substrate led to chromosomal red labelling, while green signal diffused to the rest of the cell. This sensor helped to elucidate how separase regulation

occurs in HeLa cells and to detect protease activity in a specific cell cycle phase in a concrete and controlled subcellular location (Shindo *et al.*, 2012).

A successful reporter denoted apoptotic processes in different organisms such as Drosophila and chicken embryos (Bardet *et al.*, 2008). This reporter named apoliner, depends on membrane anchoring using mCD8 followed by mRFP, a BIR1 domain and an NLS-eGFP. Caspase activity was detected both by fluorescence microscopy and immunoblots in a period shorter than one hour in fruitfly cell lines after induction of apoptosis. Initially, in healthy cells membrane co-localization of mRFP and eGFP was found, while the latter transited to the nucleus once caspases were activated. At the final stage of apoptosis, eGFP was likely degraded by additional proteases and mRFP remained as a unique signal at membranes. Apoliner showed no impact on developmental cell death measured by the TUNEL assay in Drosophila and could detect caspase activity in chick embryos faster (6 to7 hours) than by immunohistochemistry (17 hours). Another similar probe used reporters as molecules that change localization after protease cleavage (Kim *et al.*, 2013). Here, both single- and dual-colour fluorescent reporters were used that determined hepatitis C virus NS3 protease activity by miss-localization of one or both of their components. In an appealing approach for the production of multiple proteins in *Escherichia coli* and mammalian cells, a single open reading frame alternating proteins with TEVp cleavage sites showed efficient delivery of fluorescently tagged sub-products (Chen *et al.*, 2010). A correct cleavage and quantification of the different proteins in bacteria was detected on immunoblot. One of the mammalian cell lines included diverse localization proteins to nuclei, membranes and cytosol with masked target peptides which, after processing, label this location with the different fluorescent protein markers, showing flexibility for designing delocalization sensors.

TRANSCRIPTIONAL REPORTERS

Protease cleavage of membrane-bound transcriptional activators can result in reporter gene activation. Examples of setups relying on this mechanism include CaspaseTracker and CasExpress (Tang *et al.*, 2015; Ding *et al.*, 2016) that study caspase activity during apoptosis and its evasion by anastasis. The term anastasis refers to the Greek word for resurrection. In molecular biology, it indicates the mechanism by which cells experiencing caspase activation are capable to exit

apoptosis (Sun *et al.*, 2017). Although caspase activation has been thought to be a point of no return in apoptosis, anastasis is showing the potential of organisms to evade cell death. Both systems exploit the fusion of a membrane-bound mCD8 protein to the yeast transcription activator protein galactose transcription factor (GAL4) through a caspase-3-like recognition sequence (DQVD) and a BIR1 domain for enhanced caspase recognition. Membrane tethering by mCD8 prevents GAL4 to translocate to the nucleus which only occurs upon cleavage at the DQVD site. Once in the nucleus, GAL4 binds to the transcription enhancer Upstream Activation Sequence (UAS) and induces both nuclear targeted RFP expression and expression of a recombinase named flippase. A third component of the sensor, is a cassette that contains a transcriptional stop sequence flanked by recombinase sites followed by an initial untranscribed GFP. Transcription arrest is reverted through recombinase activity, excising the fragment of DNA which is subsequently inherited by the cell-line lineage (Fig. 4A). This system presents a captivating feature by which transient caspase activity and developmental caspase activity can be discriminated. Here, RFP expression reflects relatively fast apoptotic activity, GFP expression accounts for caspase activities during physiological events and development. The main difference between both systems is that with CaspaseTracker GFP is expressed in the cytosol, while with CasExpress, both signals are found in the nucleus.

Splitting transcriptional inductors using a cleavable linker between their DNA binding site and the transcriptional effector also allows to detect proteolysis (Smith and Kohorn, 1991). Recently, a series of new transcriptional reporters of activity has been developed to screen for Hepatitis C Virus NS3 protease inhibitors (Fig. 4B and C) (Tague *et al.*, 2018). Similarly to other viral proteases, NS3 can cleave itself, which can be exploited for conditional degradation (Lin *et al.*, 2008; Chung *et al.*, 2015). To study drug-conditioning proteolysis, Tague and colleagues split Gal4 into its DNA binding domain and its transcriptional activation and translationally fused to a NS3 protease through a linker containing the protease minimal substrates at both sides. In normal conditions the protease can free itself by cleavage, and the spatial contiguity of the components are lost. On the contrary, when the protease activity is inactivated by a specific inhibitor, the whole complex remained intact and moves together to the nucleus, where it recognizes the UAS region leading to H2B-Citrine expression which resulted in nuclear fluorescence (Fig. 4B). Another system uses sequestering of the Gal4 protein by a translational fusion with a protein that localizes in the membrane and

NS3 protease as dock. The system is robust enough to keep Gal4 away from the nuclei if there is presence of inhibitor, but on the contrary to the previous system, proteolysis leads to fully active Gal4 to translocate and induce expression of the reporting signal (Fig. 4C). Those are very elegant examples of both on-to off and off-to on induction and repression of a reporting system based on proteolysis and controlled by inhibitors. In the same study, an inactive Cas9 version serving as location system and transcriptional activation domain was also used. Here, by changing the guide RNA design, any reporter or product of an endogenous gene can be used as readout. In addition, an “inhibitor-off” system was developed by sequestering a full transcriptional activator into the membrane. This system is switched on upon protease activation, allowing the transcription factor to relocate to the nucleus.

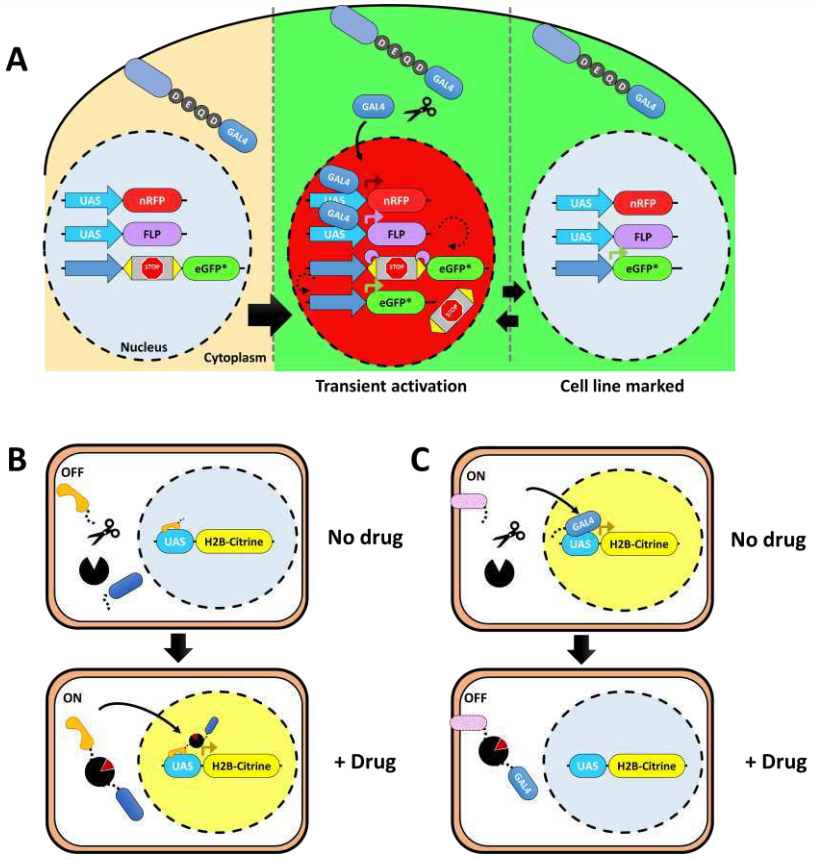


Figure 4. Transcriptional based reporters for the detection of proteolysis in a cell type. A) Scheme of the functioning mechanism of CaspaseTracker. A transcriptional activator Gal4 containing a cleavage site for caspase-3, DEQD, is translationally fused to a membrane anchored protein, which blocks unspecific transcription. Proteases cleaving the site liberate the activator, which then translocalizes to the nucleus, binds the UAS sequence and transiently expresses nuclear RFP and a recombinase such as flippase. The last part of the reporters contains a constitutive promoter followed by flippase recognition targets (yellow triangles) containing signals for transcription termination and an external eGFP coding sequence which initially is not transcribed. Once the recombinase is expressed,

it excises the DNA fragment bordering the flippase recognition targets sites allowing permanent cytosolic GFP expression in the cell and in the cell lineage originated from this cell. **B)** An off-to-on proteolytic drug-inducible transcriptional reporter where both parts of Gal4 necessary for transcription are divided by adding a NS3 protease. Under normal conditions, the activation domain is cleaved off and even if the binding domain can bind to the reporter gene, histone H2B labelled with Citrine, the signal is not present. When a NS3 inhibitor is present the continuity of the structure of the protein is maintained, shifting all together to nuclei position and inducing gene expression. **C)** Example of a proteolytic on-to-off drug-repressed transcriptional system, where Gal4 is confined to the membrane by a peptide with a NS3 protease in between. The system is initially active by action of NS3 which liberates Gal4 to report its activity. Once the drug is present, Gal4 is maintained in the membrane and the signal vanishes.

BIOLUMINESCENT SENSORS

Besides fluorescence means to detect protease activities, luminescent reporters can also be used. Bioluminescence resonance energy transfer (BRET) are based in the same principle than FRET, but generally the protein acting as donor is a luciferase which needs a substrate such as coelenterazine and ATP or oxygen to emit luminescence (Subramanian *et al.*, 2006). Depending on the range of emission of the luciferase different fluorescent proteins can be used as acceptors (Fig. 5A). In plants BRET has been used to detect processing of several versions of ubiquitin-like proteins AtAtg8 (a-i) by the Arabidopsis cysteine protease AtAtg4 in a study linking proteolysis, trafficking and autophagy (Woo *et al.*, 2014). The different versions of Atg8 were cloned into a reporter consisting of a N-terminal fusion to Citrine and a C-terminal fusion to an optimized version of Renilla luciferase (Woo and von Arnim, 2008). Interestingly, BRET ratios showed substantial differences between wild-type plants and double *Atg4a4b* mutants. BRET independent Citrine signal was detected in wild-type vacuoles, mainly as autophagosomes, indicating that the cleaved sensor behaves as the naturally occurring protein whereas the punctuated signal is lost in the mutant lines, likely due to the absence of proteolytic processing, and is detected in the cytosol. Furthermore, luciferases retain activity in native gels after addition of luciferin allowing direct detection of cleaved reporter fragments containing intact luciferase. Other BRET sensors were developed for detecting caspase-3 activity using click beetle green luciferase and tandem dimer Tomato (Gammon *et al.*, 2009). One of the newest sensors uses NanoLuc (Hall *et al.*, 2012), an engineered luciferase with optimized characteristics, with mNeonGreen for the individual detection of caspase-3, -8 and -9

(den Hamer *et al.*, 2017). A 10-fold decrease of BRET signal *in vitro* after staurosporine addition was reported.

A different design was followed for creating iGLuc, initially consisting of pro-interleukin fused to Gaussia luciferase (Bartok *et al.*, 2013). Similar to the CA-GFP probe, multimerization of the protein via pro-interleukin inactivated luciferase luminesce, which increased more than 500-fold upon caspase-1 activation (Fig. 5B). iGluc could also be used when the cleavage site was changed for a caspase-3 or a TEVp recognition site and addition of its respective proteases (Bartok *et al.*, 2013). This design also allowed the generation of specific and functional luciferase reporters for *in vitro* detection and *in vivo* imaging of other proteases such as the enterovirus 3C protease (Zhang *et al.*, 2017).

Similar to the fluorescence complementation probes, structural destabilization of luciferase reporters that can be reconstituted by proteolysis was also explored. By cyclic permuting the N- and C-terminal parts of firefly luciferase, luminescence is decreased. Both regions can be linked together by a protease cleavage site that allows enough structural flexibility after cleavage to enhance luciferase activity (Fig. 5C). This model was demonstrated for caspase-3/7/8, enterokinase and TEVp among others (Fan *et al.*, 2008). Another luciferase reporter benefits from a split firefly luciferase reporter system by addition of coil-coiled domains that hamper luciferase reconstitution by spatial limitation. The linker region can be detached by caspase-3 or TEVp cleavage, both luciferase parts can then come together and perform mono-oxygenation of luciferin, producing luminescence (Shekhawat *et al.*, 2009). This was also used for reporting proteolysis by applying the split β-lactamase system, which shows potential applications for other split reporter versions.

One of the main drawbacks of luminescent sensors is the necessity to deliver/provide co-factors for the emission of luminescence. The addition of supplementary substances might modify plant endogenous responses and requires proper controls. Moreover, detection of luminescence can decrease over time, affecting experiments where long time tracking is necessary. On the other hand, luminescence-based sensors are more sensitive than fluorescent ones and generally display very low background signals, which improves quantitation.

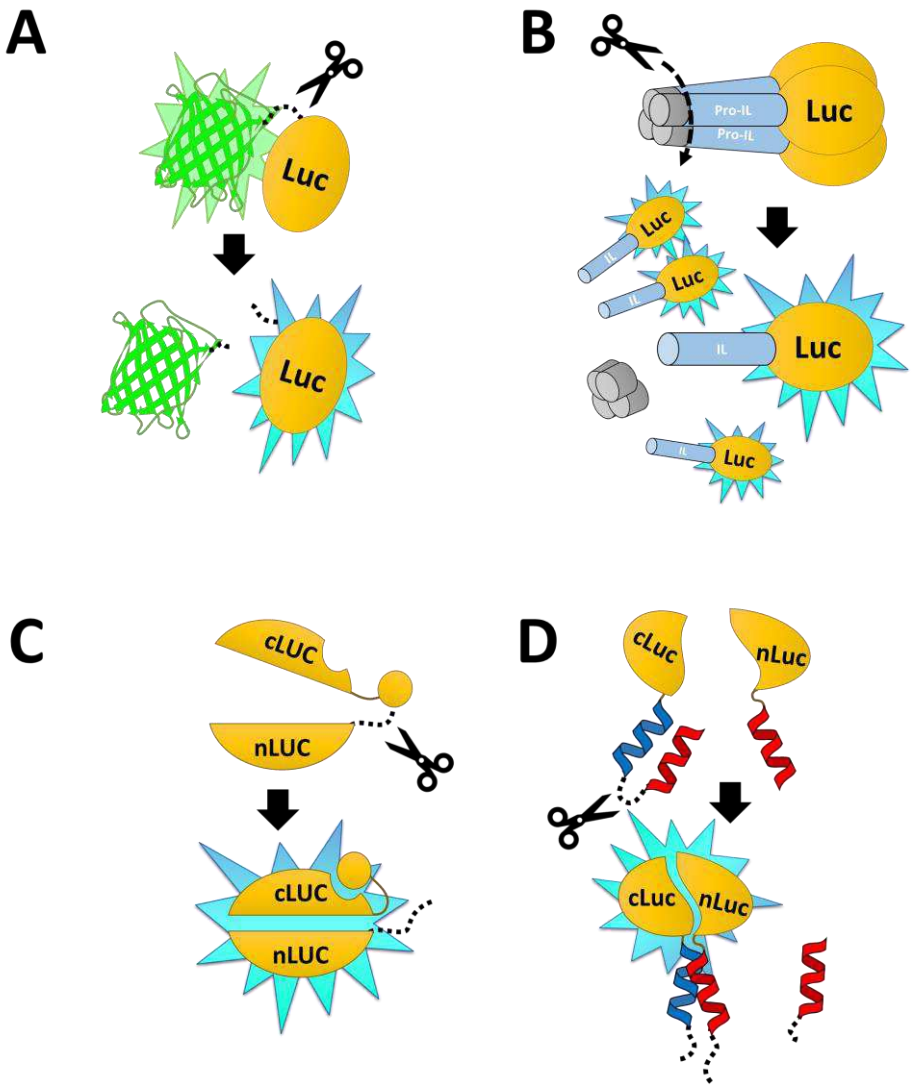


Figure 5. Luminescent reporters for the detection of protease activity. A) Schematic drawing to depict BRET rationale. The different BRET systems are based on the proximity transfer of energy of luciferases, usually blue light to a proximal fluorescence protein, in this case eGFP. The ratio between emission and absorption at different wavelengths serves as a readout for processing of the linker peptide containing a minimal substrate. **B)** Representation of iGLuc system of Gaussia luciferase based on the multimerization of pro-interleukins diminishing luciferase action. The action of protease serves as element to allow release of the conformation of the multimers resulting in a luminescence enhancement. **C)** Drawing of the cyclic permuted Luciferase (CP-Luc) used for the detection of proteolysis. N- and C-terminal domains have been swapped and connected using a linker containing a site for proteolysis. Keeping the luciferase domains separate avoids enzymatic activity when co-factors are added, but it is enhanced after freeing the components. **D)** Mechanism of the auto-inhibited coil-coiled reporter Luciferase. In this system, complementation of the component parts is blocked by coil-coiled element addition to the split luciferase structures, blocking the interaction of the halves. When this appendages are proteolytically removed, both parts of firefly luciferase can reconstitute the structure, which in presence of luciferin and ATP, will emit luminescence.

CONCLUSIONS AND PERSPECTIVES

In the last decade, the implementation and development of novel protease activity probes for use in plants has led to significant biological insight on processes ranging from development and senescence to biotic stress (Morimoto and van der Hoorn, 2016). Some of these studies leveraged the potential of imaging spatiotemporal protease activity, as addressed in the beginning of this review, together with an increased use of chemicals and dyes for imaging. ABPs in the plant field have mainly been used for the identification of active proteases *in vivo* and *in vitro*. The diversity of established protocols in plants have greatly aided in the pipeline for protease identification using mass spectrometry. However, not many probes have been used for imaging, mostly because they label many proteases from the same class. One of the main drawbacks of chemical probes is that ABPs work as inhibitors of proteolytic activity, thus inactivating enzymatic activity and potentially blocking downstream processes and disturbing cell biology. Possibly though, only a part of the total pool of ABP-targeted proteases is labelled, allowing the other fraction of the pool to perform their biological role. In addition, the membrane-permeability and bioavailability might be limited in some cases, although many fluorescent probes can be used for labelling *in vivo*. Therefore, to study *in vivo* processes one should use the minimal amount of ABP that permits the tracking and visualization of the target protease while minimizing the effect on plant physiology.

More recently, the development of proteinaceous probes to visualize protease activity has evolved. Genetically encoded probes include fluorescent, luminescent and transcriptional detectors of proteolysis. Genetic reporters are relatively cheap and easy to tune for specific protease requirements by standard cloning techniques available in most molecular biology labs. They also allow versatility when specific tissues are the area of interest where a protease is expressed. Genetic probes do not inhibit proteases, but might compete for genuine substrates. Therefore, it is important to test that such reporters do not have a significant influence on the system to which they are applied. For instance, adding a substrate of caspases may speed up apoptosis generating internal bias in the experiment. Different genetic reporters allow fast detection of proteolysis without additional steps needed for labelling or washing of reagents. In terms of speed, FRET reporters are well-suited for real time and fast processes, because their readout depends on protein proximity, which is lost

immediately upon processing. Fluorescent complementation follows a gradual activation, but reporting the activity is delayed in time. Luminescent sensors have as an advantage that they allow very precise measurements, but, on the other hand, they require expensive co-factors. Lastly, transcriptional sensors are interesting tools that mark activity from parental cells in which proteolysis occurs. Transcriptional sensors also have the capacity to amplification the signal reporting proteolysis. This property is advantageous for proteases that are lowly abundant or with reduced activity and therefore their activity might not be detectable using sensors only reporting direct cleavage. However, they are without any doubt the slowest to report activity due to their mechanism and, in general, they require translocation of the transcriptional particle, transcription of the reporter protein following folding and maturation of the newly synthesized protein.

Genetic protease reporters have been underexplored in plant research, maybe due to the lack of known protease activation mechanisms and the poor number of validated substrates. Minimal knowledge on protease specificity is generally required to design such tools, meaning the capacity to identify real or synthetic substrates by some of the existing proteomic techniques used in N-terminomics (Demir *et al.*, 2018). However, these arguments also hold true for chemical probes. Nevertheless, probes could be a useful tool to determine specificity of the cleavage motifs in cases where N-terminomics are not suitable to identify specific peptides. For instance, some peptides resulting from proteases activity can contain physical and chemical properties which made them undetectable using mass spectrometry, but which can be inserted in sensors to determine action of proteases with the proper controls. A considerable effort will be required to establish trusted protocols and protease-specific probes in plant research. Except for the limited availability of biliverdin in all plants cellular compartments (Kohchi *et al.*, 2001) which would limit the use of iProteases sensors (To *et al.*, 2015), there seem no significant biological barriers to adapt existing genetic probes to plant research. In mammalian studies, caspases were mostly targeted by various genetic probes, primary because of their importance in disease and inflammation, but perhaps also because they are a prime example of an inducible protease with a well-known activation mechanism with very distinctive cleavage profiles and therefore easy to investigate (Salvesen *et al.*, 2016). Apart from reconfirming previous observations, genetic protease probes could serve to discover

new biology, as exemplified in the case of anastasis (Sun *et al.*, 2017). As plant proteases are generally less studied than mammalian proteases, this provides a clear incentive for translation of the concepts of genetic probes to plants and the discovery of new biology.

Box

Considerations when working with or translating chemical and genetic probes to plants:

- Occam's razor principle states that the simplest solution is most likely the best. The most uncomplicated sensors and the least number of additional components introduced in the experiments might help to produce results. Also probes with less sophisticated equipment requirements for readout will facilitate the work of the researchers (excitation/emission).
- Probe design. There is not a clear rule on how long a linker can be, this might depend on the probe and on the nature of the amino acids and the final spatial conformation of its ends. Some examples in this review indicate that it is possible to detect cleavage of some full proteins labelled at the C- and N-termini as is the case for DA1 and separase substrates using FRET or delocalization sensors. Additionally, there exists a lot of variability between organisms and efficiency on the reporters. We suggest to test reporters with known proteases, such as TEVp, can help to optimize the system. Furthermore, we encourage to use uncleavable linker sequences containing flexible linkers such as repetitions of glycine-glycine-serine or known proteolysis resistant sequences that can have mutagenized version of key residues in the recognition sequence as negative controls. Additional controls using mutant lines where proteases are knock-out or are inactive is highly recommended.
- The selection of the fluorophores for a concrete application need to be considered (Evers *et al.*, 2006; Bajar *et al.*, 2016). Far red probes developed specifically for imaging deeper in mammalian tissues are not particularly needed for plant protease research. For example, using red shifted chemicals or proteins in the aerial parts of plants might overlap with chlorophyll auto-fluorescence, thereby complicating imaging.
- Cell permeability. Chemical probes are very diverse and presence of hydrogen bond donors and acceptors can reduce membrane permeability (Verdoes and Verhelst, 2016). This is exemplified with some biotinylated and carboxylate containing probes. This problem can be circumvented by conjugation of cell-penetrating peptides or addition of a two-step labelling protocol. Carboxylic acid can also be masked in the shape of ethyl ester, which increases its permeability.

- Subcellular localization. Genetic probes can be targeted via targeting signals, whereas chemicals are generally untargeted. Targeting might be a way to increase a probe's specificity. Most genetic probes so far have been applied in the cytosol, although there should be no problem targeting them to organelles. A point of concern might be the effect of pH variation for example in the secretory pathway (decreasing pH along trafficking from the ER to the extracellular space or vacuole).
- Kinetics. Looking at fast processes is probably most efficient using FRET probes. Slower folding of complementation type probes or incubation times for chemical labelling better serve a static and macroscopic overview of protease activity.
- Availability of protease substrate specificities is a bottleneck for both chemical (that frequently employ peptide chemical bonds in their structure) and genetic probes. However, these are increasingly becoming available for plant proteases (Demir *et al.*, 2018). Protease specificity of genetic probes might benefit here from the added information on P' site signatures, which are mostly not present in chemical probes because of their make-up having a functional group (warhead) in the P1' position.

Acknowledgements

S.S. is supported by a grant from the Flemish funding agency FWO (#12M3418N). Ghent University Special Research Fund (BOF 01J11311) to F.V.B. and K.G. The authors wish to thank Annick Bleys for help in preparing the manuscript.

Sensor	Readout	Mechanism	Protease(s) targeted	Uses	Substrates	Maximal Substrate coverage	Response	References
Activity Based Probes (ABP)	Fluorescence	Catalytic site binding	+++	*in vitro/ in vivo*	+++	P4-P1'	Fast (1)	(Morimoto and van der Hoorn, 2016)
DEVD-NucView488	Fluorescence	Cleavage activation	Casp-3	*in vitro/ in vivo*	DEVD	P4-P1	Fast	(Cen *et al.*, 2008)
CR-(DEVD)2	Fluorescence	Cleavage activation	Casp-3	*in vitro/ in vivo*	DEVD	P4-P1	Fast	(Bosch and Franklin-Tong, 2007)
FRET	rFluorescence	Cleavage reduces signal	+++	*in vitro/ in vivo*	+++	P5-P5' (2)	Fast	+++
Grafted GFP	Fluorescence	Fluorescence Complementation	Trypsin, Casp-3	*in vitro/ in vivo*	32 aa linker, DEVD	NA	Medium	(Chen *et al.*, 2009a; Chen *et al.*, 2009b)
Pro-GFP	Fluorescence	Fluorescence Complementation	Thrombin, Casp-3, HIV protease	*in vivo*	+++	+++	Medium	(Callahan *et al.*, 2010)
VC3AI	Fluorescence	Fluorescence Complementation	Casp-3	*in vitro/ in vivo*	DEVD	P4-P1'	Medium	(Zhang *et al.*, 2013)
iProteases	rFluorescence	Fluorescence Complementation	Casp-3, TEVp, HCV	*in vitro/ in vivo*	+++	9-15 aa consensus sequence	Fast-Medium	(To *et al.*, 2015)
ZipGFP	rFluorescence	Fluorescence Complementation	Casp-3, TEVp	*in vitro/ in vivo*	DEVD/ENLYFQ	+++	Medium	(To *et al.*, 2016)
Caspase Activatable-GFP (CA-GFP)	Fluorescence	Fluorescence Gain	Casp-3, Casp-7, TEVp	*in vitro/ in vivo*	+++	P6-P4'	Medium	(Nicholls *et al.*, 2011; Nicholls and Hardy, 2013; Wu *et al.*, 2013)
GFP-ssrA	Fluorescence	Fluorescence Gain	TEVp	*in vitro/ in vivo*	ENLYFQX variants	P6-P1'	Medium	(Kostallas and Samuelson, 2010; Kostallas *et al.*, 2011)
NES-DEVD-YFP-NLS	Fluorescence change	Miss-localization	Casp-3	*in vivo*	DEVD	P4-P1	Fast	(Tang *et al.*, 2012)
Separase sensor	Fluorescence change	Miss-localization	Separase	*in vivo*	Cohesin	Full protein	Fast	(Shindo *et al.*, 2012)
Apoliner	Fluorescence change	Miss-localization	Casp-3	*in vivo*	DEVD	P4-P1	Medium-Late	(Bardet *et al.*, 2008)
Caspase Tracker	Fluorescence	Fluorescence transcription	Casp-3	*in vivo*	DQVD	P4-P1'	Late	(Tang *et al.*, 2015; Tang *et al.*, 2018)

CasExpress	Fluorescence	Fluorescence transcription	Casp-3	*in vivo*	DQVD	P4-P1	Late	(Ding *et al.*, 2016)
Chemogenetic probes	Fluorescence	Fluorescence transcription	HCV	*in vivo*	NA	NA	Late	(Tague *et al.*, 2018)
BRET	rLuminescent	Cleavage reduces signal	+++	*in vitro/ in vivo*	+++	+++	Fast	(Gammon *et al.*, 2009; Hall *et al.*, 2012; Woo *et al.*, 2014)
iGluc	Luminescent	Fluorescence Gain	+++	*in vitro/ in vivo*	+++	NA	Medium	(Bartok *et al.*, 2013)
CP-Luc	Luminescent	Luminescence Complementation	+++	*in vitro/ in vivo*	+++	NA	Medium	(Fan *et al.*, 2008)
Coiled-coil Luciferase	Luminescent	Luminescence Complementation	Casp-3, TEVp	*in vitro/ in vivo*	DEVD/ENLYFQ	P6-P1'	Medium	(Shekhawat *et al.*, 2009)

Table 1. Main protease reporters mentioned in this review and their characteristics

(1) Dependent on cell permeability.

(2) Linkage length highly variable depending on design.

+++ Indicates more than 3 counts on this cell value.

Aa: amino acids.

rFluorescence: ratiometric Fluorescence.

Casp: Caspase

NA: Not available

REFERENCES

Andréasson C, Heessen S, Ljungdahl PO. 2006. Regulation of transcription factor latency by receptor-activated proteolysis. *Genes & Development* **20,** 1563-1568.

Bajar BT, Wang ES, Zhang S, Lin MZ, Chu J. 2016. A guide to fluorescent protein FRET pairs. *Sensors* **16,** 1488.

Bardet P-L, Kolahgar G, Mynett A, Miguel-Aliaga I, Briscoe J, Meier P, Vincent J-P. 2008. A fluorescent reporter of caspase activity for live imaging. *Proceedings of the National Academy of Sciences of the United States of America* **105,** 13901-13905.

Bartok E, Bauernfeind F, Khaminets MG, Jakobs C, Monks B, Fitzgerald KA, Latz E, Hornung V. 2013. iGLuc: a luciferase-based inflammasome and protease activity reporter. *Nature Methods* **10,** 147-154.

Belenghi B, Romero-Puertas MC, Vercammen D, Brackenier A, Inzé D, Delledonne M, Van Breusegem F. 2007. Metacaspase activity of *Arabidopsis thaliana* is regulated by *S*-nitrosylation of a critical cysteine residue. *Journal of Biological Chemistry* **282,** 1352-1358.

Benchabane M, Schlüter U, Vorster J, Goulet M-C, Michaud D. 2010. Plant cystatins. *Biochimie* **92,** 1657-1666.

Blum G, Mullins SR, Keren K, Fonovič M, Jedeszko C, Rice MJ, Sloane BF, Bogyo M. 2005. Dynamic imaging of protease activity with fluorescently quenched activity-based probes. *Nature Chemical Biology* **1,** 203-209.

Blum G, von Degenfeld G, Merchant MJ, Blau HM, Bogyo M. 2007. Noninvasive optical imaging of cysteine protease activity using fluorescently quenched activity-based probes. *Nature Chemical Biology* **3,** 668-677.

Bosch M, Franklin-Tong VE. 2007. Temporal and spatial activation of caspase-like enzymes induced by self-incompatibility in *Papaver* pollen. *Proceedings of the National Academy of Sciences of the United States of America* **104,** 18327-18332.

Bosch M, Franklin-Tong VE. 2008. Self-incompatibility in *Papaver*: signalling to trigger PCD in incompatible pollen. *Journal of Experimental Botany* **59,** 481-490.

Boulware KT, Jabaiah A, Daugherty PS. 2010. Evolutionary optimization of peptide substrates for proteases that exhibit rapid hydrolysis kinetics. *Biotechnology and Bioengineering* **106,** 339-346.

Callahan BP, Stanger MJ, Belfort M. 2010. Protease activation of split green fluorescent protein. *ChemBioChem* **11,** 2259-2263.

Cen H, Mao F, Aronchik I, Fuentes RJ, Firestone GL. 2008. DEVD-NucView488: a novel class of enzyme substrates for real-time detection of caspase-3 activity in live cells. *FASEB Journal* **22,** 2243-2252.

Chang PV, Prescher JA, Sletten EM, Baskin JM, Miller IA, Agard NJ, Lo A, Bertozzi CR. 2010. Copper-free click chemistry in living animals. *Proceedings of the National Academy of Sciences of the United States of America* **107,** 1821-1826.

Chapnick DA, Bunker E, Liu X. 2015. A biosensor for the activity of the "sheddase" TACE (ADAM17) reveals novel and cell type-specific mechanisms of TACE activation. *Science Signaling* **8,** rs1.

Chen N, Huang Y, Yang L, Liu R, Yang JJ. 2009a. Designing caspase-3 sensors for imaging of apoptosis in living cells. *Chemistry* **15,** 9311-9314.

Chen N, Zou J, Wang S, Ye Y, Huang Y, Gadda G, Yang JJ. 2009b. Designing protease sensors for real-time imaging of trypsin activation in pancreatic cancer cells. *Biochemistry* **48,** 3519-3526.

Chen X, Pham E, Truong K. 2010. TEV protease-facilitated stoichiometric delivery of multiple genes using a single expression vector. *Protein Science* **19,** 2379-2388.

Chichkova NV, Shaw J, Galiullina RA*, et al.* 2010. Phytaspase, a relocalisable cell death promoting plant protease with caspase specificity. *EMBO Journal* **29,** 1149-1161.

Chung HK, Jacobs CL, Huo Y, Yang J, Krumm SA, Plemper RK, Tsien RY, Lin MZ. 2015. Tunable and reversible drug control of protein production via a self-excising degron. *Nature Chemical Biology* **11,** 713-720.

Demir F, Niedermaier S, Villamor JG, Huesgen PF. 2018. Quantitative proteomics in plant protease substrate identification. *New Phytologist* **218,** 936-943.

den Hamer A, Dierickx P, Arts R, de Vries JSPM, Brunsveld L, Merkx M. 2017. Bright bioluminescent BRET sensor proteins for measuring intracellular caspase activity. *ACS Sensors* **2,** 729-734.

Ding AX, Sun G, Argaw YG, Wong JO, Easwaran S, Montell DJ. 2016. CasExpress reveals widespread and diverse patterns of cell survival of caspase-3 activation during development in vivo. *eLife* **5,** e10936.

Dong H, Dumenil J, Lu F-H*, et al.* 2017. Ubiquitylation activates a peptidase that promotes cleavage and destabilization of its activating E3 ligases and diverse growth regulatory proteins to limit cell proliferation in *Arabidopsis*. *Genes & Development* **31,** 197-208.

Edgington-Mitchell LE, Bogyo M, Verdoes M. 2017. Live cell imaging and profiling of cysteine cathepsin activity using a quenched activity-based probe. *Methods in Molecular Biology* **1491,** 145-159.

Edgington LE, Berger AB, Blum G, Albrow VE, Paulick MG, Lineberry N, Bogyo M. 2009. Noninvasive optical imaging of apoptosis by caspase-targeted activity-based probes. *Nature Medicine* **15,** 967-973.

Edgington LE, Verdoes M, Ortega A, Withana NP, Lee J, Syed S, Bachmann MH, Blum G, Bogyo M. 2013. Functional imaging of legumain in cancer using a new quenched activity-based probe. *Journal of the American Chemical Society* **135,** 174-182.

Evers TH, van Dongen EMW, Faesen AC, Meijer EW, Merkx M. 2006. Quantitative understanding of the energy transfer between fluorescent proteins connected via flexible peptide linkers. *Biochemistry* **45,** 13183-13192.

Fan F, Binkowski BF, Butler BL, Stecha PF, Lewis MK, Wood KV. 2008. Novel genetically encoded biosensors using firefly luciferase. *ACS Chemical Biology* **3,** 346-351.

Feng S, Sekine S, Pessino V, Li H, Leonetti MD, Huang B. 2017. Improved split fluorescent proteins for endogenous protein labeling. *Nature Communications* **8,** 370.

Fretwell JF, SM KI, Cummings JM, Selby TL. 2008. Characterization of a randomized FRET library for protease specificity determination. *Molecular BioSystems* **4,** 862-870.

Gallwitz M, Enoksson M, Thorpe M, Hellman L. 2012. The extended cleavage specificity of human thrombin. *PLoS ONE* **7,** e31756.

Gammon ST, Villalobos VM, Roshal M, Samrakandi M, Piwnica-Worms D. 2009. Rational design of novel red-shifted BRET pairs: platforms for real-time single-chain protease biosensors. *Biotechnology Progress* **25,** 559-569.

García-Lorenzo M, Sjödin A, Jansson S, Funk C. 2006. Protease gene families in *Populus* and Arabidopsis. *BMC Plant Biology* **6,** 30.

Grosse-Holz FM, van der Hoorn RAL. 2016. Juggling jobs: roles and mechanisms of multifunctional protease inhibitors in plants. *New Phytologist* **210,** 794-807.

Grossmann G, Krebs M, Maizel A, Stahl Y, Vermeer JEM, Ott T. 2018. Green light for quantitative live-cell imaging in plants. *Journal of Cell Science* **131,** jcs209270.

Grunberg R, Burnier JV, Ferrar T*, et al.* 2013. Engineering of weak helper interactions for high-efficiency FRET probes. *Nat Methods* **10,** 1021-1027.

Gu C, Shabab M, Strasser R, Wolters PJ, Shindo T, Niemer M, Kaschani F, Mach L, van der Hoorn RAL. 2012. Post-translational regulation and trafficking of the granulin-containing protease RD21 of *Arabidopsis thaliana*. *PLoS ONE* **7,** e32422.

Hall MP, Unch J, Binkowski BF*, et al.* 2012. Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS Chemical Biology* **7,** 1848-1857.

Hatsugai N, Yamada K, Goto-Yamada S, Hara-Nishimura I. 2015. Vacuolar processing enzyme in plant programmed cell death. *Frontiers in Plant Science* **6,** 234.

Heal WP, Dang THT, Tate EW. 2011. Activity-based probes: discovering new biology and new drug targets. *Chemical Society Reviews* **40,** 246-257.

Heim R, Tsien RY. 1996. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Current Biology* **6,** 178-182.

Hu J, Huang X, Chen L, Sun X, Lu C, Zhang L, Wang Y, Zuo J. 2015. Site-specific nitrosoproteomic identification of endogenously *S*-nitrosylated proteins in Arabidopsis. *Plant Physiology* **167,** 1731-1746.

Iwata Y, Ashida M, Hasegawa C, Tabara K, Mishiba KI, Koizumi N. 2017. Activation of the Arabidopsis membrane-bound transcription factor bZIP28 is mediated by site-2 protease, but not site-1 protease. *Plant Journal* **91,** 408-415.

Jones J, Heim R, Hare E, Stack J, Pollok BA. 2000. Development and application of a GFP-FRET intracellular caspase assay for drug screening. *Journal of Biomolecular Screening* **5,** 307-317.

Kamiyama D, Sekine S, Barsi-Rhyne B*, et al.* 2016. Versatile protein tagging in cells with split fluorescent protein. *Nature Communications* **7,** 11046.

Kaschani F, Gu C, Niessen S, Hoover H, Cravatt BF, van der Hoorn RAL. 2009a. Diversity of serine hydrolase activities of unchallenged and *Botrytis*-infected *Arabidopsis thaliana*. *Molecular & Cellular Proteomics* **8,** 1082-1093.

Kaschani F, Verhelst SHL, van Swieten PF*, et al.* 2009b. Minitags for small molecules: detecting targets of reactive small molecules in living plant tissues using 'click chemistry'. *Plant Journal* **57,** 373-385.

Kim JH, Lee MJ, Hwang I, Hwang HJ. 2013. In-cell protease assay systems based on trans-localizing molecular beacon proteins using HCV protease as a model system. *PLoS ONE* **8,** e59710.

Kim Y-S, Kim S-G, Park J-E, Park H-Y, Lim M-H, Chua N-H, Park C-M. 2006. A membrane-bound NAC transcription factor regulates cell division in *Arabidopsis*. *Plant Cell* **18,** 3132-3144.

Kimura RH, Steenblock ER, Camarero JA. 2007. Development of a cell-based fluorescence resonance energy transfer reporter for *Bacillus anthracis* lethal factor protease. *Analytical Biochemistry* **369,** 60-70.

Kohchi T, Mukougawa K, Frankenberg N, Masuda M, Yokota A, Lagarias JC. 2001. The Arabidopsis *HY2* gene encodes phytochromobilin synthase, a ferredoxin-dependent biliverdin reductase. *Plant Cell* **13,** 425-436.

Kolodziejek I, Misas-Villamil JC, Kaschani F*, et al.* 2011. Proteasome activity imaging and profiling characterizes bacterial effector syringolin A. *Plant Physiology* **155,** 477-489.

Kostallas G, Löfdahl P-A, Samuelson P. 2011. Substrate profiling of tobacco etch virus protease using a novel fluorescence-assisted whole-cell assay. *PLoS ONE* **6,** e16136.

Kostallas G, Samuelson P. 2010. Novel fluorescence-assisted whole-cell assay for engineering and characterization of proteases and their substrates. *Applied and Environmental Microbiology* **76,** 7500-7508.

Lallemand J, Bouché F, Desiron C, Stautemas J, de Lemos Esteves F, Périlleux C, Tocquin P. 2015. Extracellular peptidase hunting for improvement of protein production in plant cells and roots. *Frontiers in Plant Science* **6,** 37.

Lam E, Zhang Y. 2012. Regulating the reapers: activating metacaspases for programmed cell death. *Trends in Plant Science* **17,** 487-494.

Lampl N, Alkan N, Davydov O, Fluhr R. 2013. Set-point control of RD21 protease activity by AtSerpin1 controls cell death in Arabidopsis. *Plant Journal* **74,** 498-510.

Li Y-M, Xu M, Lai M-T*, et al.* 2000. Photoactivated γ-secretase inhibitors directed to the active site covalently label presenilin 1. *Nature* **405,** 689-694.

Lin MZ, Glenn JS, Tsien RY. 2008. A drug-controllable tag for visualizing newly synthesized proteins in cells and whole animals. *Proceedings of the National Academy of Sciences of the United States of America* **105,** 7744-7749.

Liu Y, Patricelli MP, Cravatt BF. 1999. Activity-based protein profiling: the serine hydrolases. *Proceedings of the National Academy of Sciences of the United States of America* **96,** 14694-14699.

Luo KQ, Yu VC, Pu Y, Chang DC. 2001. Application of the fluorescence resonance energy transfer method for studying the dynamics of caspase-3 activation during UV-induced apoptosis in living HeLa cells. *Biochemical and Biophysical Research Communications* **283,** 1054-1060.

Luo KQ, Yu VC, Pu Y, Chang DC. 2003. Measuring dynamics of caspase-8 activation in a single living HeLa cell during TNF*α*-induced apoptosis. *Biochemical and Biophysical Research Communications* **304,** 217-222.

Mahajan NP, Harrison-Shostak DC, Michaux J, Herman B. 1999. Novel mutant green fluorescent protein protease substrates reveal the activation of specific caspases during apoptosis. *Chemistry & Biology* **6,** 401-409.

Meyer M, Leptihn S, Welz M, Schaller A. 2016. Functional characterization of propeptides in plant subtilases as intramolecular chaperones and inhibitors of the mature protease. *Journal of Biological Chemistry* **291,** 19449-19461.

Mills KV, Johnson MA, Perler FB. 2014. Protein splicing: how inteins escape from precursor proteins. *Journal of Biological Chemistry* **289,** 14498-14505.

Misas-Villamil JC, Toenges G, Kolodziejek I, Sadaghiani AM, Kaschani F, Colby T, Bogyo M, van der Hoorn RAL. 2013. Activity profiling of vacuolar processing enzymes reveals a role for VPE during oomycete infection. *Plant Journal* **73,** 689-700.

Mitra RD, Silva CM, Youvan DC. 1996. Fluorescence resonance energy transfer between blue-emitting and red-shifted excitation derivatives of the green fluorescent protein. *Gene* **173,** 13-17.

Morimoto K, van der Hoorn RAL. 2016. The increasing impact of activity-based protein profiling in plant science. *Plant & Cell Physiology* **57,** 446-461.

Ng S, Ivanova A, Duncan O*, et al.* 2013. A membrane-bound NAC transcription factor, ANAC017, mediates mitochondrial retrograde signaling in *Arabidopsis*. *Plant Cell* **25,** 3450-3471.

Nicholls SB, Chu J, Abbruzzese G, Tremblay KD, Hardy JA. 2011. Mechanism of a genetically encoded dark-to-bright reporter for caspase activity. *Journal of Biological Chemistry* **286,** 24977-27986.

Nicholls SB, Hardy JA. 2013. Structural basis of fluorescence quenching in caspase activatable-GFP. *Protein Science* **22,** 247-257.

Ohad N, Shichrur K, Yalovsky S. 2007. The analysis of protein-protein interactions in plants by bimolecular fluorescence complementation. *Plant Physiology* **145,** 1090-1099.

Olsen JV, Ong S-E, Mann M. 2004. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Molecular & Cellular Proteomics* **3,** 608-614.

Ouyang M, Lu S, Li X-Y, Xu J, Seong J, Giepmans BNG, Shyy JY-J, Weiss SJ, Wang Y. 2008. Visualization of polarized membrane type 1 matrix metalloproteinase activity in live cells by fluorescence resonance energy transfer imaging. *Journal of Biological Chemistry* **283,** 17740-17748.

Paper JM, Mukherjee T, Schrick K. 2018. Bioorthogonal click chemistry for fluorescence imaging of choline phospholipids in plants. *Plant Methods* **14,** 31.

Rawlings ND, Barrett AJ, Bateman A. 2011. Asparagine peptide lyases: a seventh catalytic type of proteolytic enzymes. *Journal of Biological Chemistry* **286,** 38321-38328.

Romero-Puertas MC, Campostrini N, Mattè A, Righetti PG, Perazzolli M, Zolla L, Roepstorff P, Delledonne M. 2008. Proteomic analysis of S-nitrosylated proteins in *Arabidopsis thaliana* undergoing hypersensitive response. *Proteomics* **8,** 1459-1469.

Rustgi S, Boex-Fontvieille E, Reinbothe C, von Wettstein D, Reinbothe S. 2018. The complex world of plant protease inhibitors: insights into a Kunitz-type cysteine protease inhibitor of *Arabidopsis thaliana*. *Communicative and Integrative Biology* **11,** e1368599.

Sabariegos R, Picazo F, Domingo B, Franco S, Martinez M-A, Llopis J. 2009. Fluorescence resonance energy transfer-based assay for characterization of hepatitis C virus NS3-4A protease activity in live cells. *Antimicrobial Agents and Chemotherapy* **53,** 728-734.

Salvesen GS, Hempel A, Coll NS. 2016. Protease signaling in animal and plant-regulated cell death. *FEBS Journal* **283,** 2577-2598.

Sandersjöö L, Jonsson A, Löfblom J. 2017. Protease substrate profiling using bacterial display of self-blocking affinity proteins and flow-cytometric sorting. *Biotechnology Journal* **12,** 1600365.

Sanman LE, Bogyo M. 2014. Activity-based profiling of proteases. *Annual Review of Biochemistry* **83,** 249-273.

Schairer R, Hall G, Zhang M*, et al.* 2020. Allosteric activation of MALT1 by its ubiquitin-binding Ig3 domain. *Proc Natl Acad Sci U S A* **117,** 3093-3102.

Schechter I, Berger A. 1968. On the active site of proteases. III. Mapping the active site of papain; specific peptide inhibitors of papain. *Biochemical and Biophysical Research Communications* **32,** 898-902.

Schott S, Ambrosini A, Barbaste A, Benassayag C, Gracia M, Proag A, Rayer M, Monier B, Suzanne M. 2017. A fluorescent toolkit for spatiotemporal tracking of apoptotic cells in living *Drosophila* tissues. *Development* **144,** 3840-3846.

Schulenburg C, Faccio G, Jankowska D, Maniura-Weber K, Richter M. 2016. A FRET-based biosensor for the detection of neutrophil elastase. *Analyst* **141,** 1645-1648.

Shekhawat SS, Porter JR, Sriprasad A, Ghosh I. 2009. An autoinhibited coiled-coil design strategy for split-protein protease sensors. *Journal of the American Chemical Society* **131,** 15284-15290.

Shindo N, Kumada K, Hirota T. 2012. Separase sensor reveals dual roles for separase coordinating cohesin cleavage and cdk1 inhibition. *Developmental Cell* **23,** 112-123.

Smith TA, Kohorn BD. 1991. Direct selection for sequences encoding proteases of known specificity. *Proceedings of the National Academy of Sciences of the United States of America* **88,** 5159-5162.

Speers AE, Cravatt BF. 2004. Profiling enzyme activities in vivo using click chemistry methods. *Chemistry & Biology* **11,** 535-546.

Stawarski M, Rutkowska-Wlodarczyk I, Zeug A, Bijata M, Madej H, Kaczmarek L, Wlodarczyk J. 2014. Genetically encoded FRET-based biosensor for imaging MMP-9 activity. *Biomaterials* **35,** 1402-1410.

Stockholm D, Bartoli M, Sillon G, Bourg N, Davoust J, Richard I. 2005. Imaging calpain protease activity by multiphoton FRET in living mice. *Journal of Molecular Biology* **346,** 215-222.

Subramanian C, Woo J, Cai X, Xu X, Servick S, Johnson CH, Nebenführ A, von Arnim AG. 2006. A suite of tools and application notes for *in vivo* protein interaction assays using bioluminescence resonance energy transfer (BRET). *Plant Journal* **48,** 138-152.

Sueldo DJ, van der Hoorn RAL. 2017. Plant life needs cell death, but does plant cell death need Cys proteases? *FEBS Journal* **284,** 1577-1585.

Sun G, Guzman E, Balasanyan V, Conner CM, Wong K, Zhou HR, Kosik KS, Montell DJ. 2017. A molecular signature for anastasis, recovery from the brink of apoptotic cell death. *Journal of Cell Biology* **216,** 3355-3368.

Szymczak AL, Workman CJ, Wang Y, Vignali KM, Dilioglou S, Vanin EF, Vignali DAA. 2004. Correction of multi-gene deficiency *in vivo* using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nature Biotechnology* **22,** 589-594.

Tague EP, Dotson HL, Tunney SN, Sloas DC, Ngo JT. 2018. Chemogenetic control of gene expression and cell signaling with antiviral drugs. *Nature Methods* **15,** 519-522.

Tang HL, Tang HM, Fung MC, Hardwick JM. 2015. *In vivo* CaspaseTracker biosensor system for detecting anastasis and non-apoptotic caspase activity. *Scientific Reports* **5,** 9015.

Tang HL, Tang HM, Mak KH*, et al.* 2012. Cell survival, DNA damage, and oncogenic transformation after a transient and reversible apoptotic response. *Molecular Biology of the Cell* **23,** 2240-2252.

Tang HM, Fung MC, Tang HL. 2018. Detecting anastasis *in vivo* by CaspaseTracker biosensor. *Journal of Visualized Experiments* **(132),** 54107.

To T-L, Piggott BJ, Makhijani K, Yu D, Jan YN, Shu X. 2015. Rationally designed fluorogenic protease reporter visualizes spatiotemporal dynamics of apoptosis in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **112,** 3338-3343.

To T-L, Schepis A, Ruiz-González R, Zhang Q, Yu D, Dong Z, Coughlin SR, Shu X. 2016. Rational design of a GFP-based fluorogenic caspase reporter for imaging apoptosis in vivo. *Cell Chemical Biology* **23,** 875-882.

Turk B. 2006. Targeting proteases: successes, failures and future prospects. *Nature Reviews Drug Discovery* **5,** 785-799.

van der Hoorn RAL. 2008. Plant proteases: from phenotypes to molecular mechanisms. *Annual Review of Plant Biology* **59,** 191-223.

van Wijk KJ. 2015. Protein maturation and proteolysis in plant plastids, mitochondria, and peroxisomes. *Annual Review of Plant Biology* **66,** 75-111.

Vanderklish PW, Krushel LA, Holst BH, Gally JA, Crossin KL, Edelman GM. 2000. Marking synaptic activity in dendritic spines with a calpain substrate exhibiting fluorescence resonance energy transfer. *Proceedings of the National Academy of Sciences of the United States of America* **97,** 2253-2258.

Vercammen D, Belenghi B, van de Cotte B*, et al.* 2006. Serpin1 of *Arabidopsis thaliana* is a suicide inhibitor for metacaspase 9. *Journal of Molecular Biology* **364,** 625-636.

Verdoes M, Edgington LE, Scheeren FA, Leyva M, Blum G, Weiskopf K, Bachmann MH, Ellman JA, Bogyo M. 2012. A nonpeptidic cathepsin S activity-based probe for noninvasive optical imaging of tumor-associated macrophages. *Chemistry & Biology* **19,** 619-628.

Verdoes M, Verhelst SHL. 2016. Detection of protease activity in cells and animals. *Biochimica et Biophysica Acta - Proteins and Proteomics* **1864,** 130-142.

Vinkenborg JL, Evers TH, Reulen SWA, Meijer EW, Merkx M. 2007. Enhanced sensitivity of FRET-based protease sensors by redesign of the GFP dimerization interface. *ChemBioChem* **8,** 1119-1121.

Weiss S. 2000. Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy. *Nature Structural Biology* **7,** 724-729.

Wohlever ML, Nager AR, Baker TA, Sauer RT. 2013. Engineering fluorescent protein substrates for the AAA+ Lon protease. *Protein Engineering, Design and Selection* **26,** 299-305.

Woo J, Park E, Dinesh-Kumar SP. 2014. Differential processing of *Arabidopsis* ubiquitin-like Atg8 autophagy proteins by Atg4 cysteine proteases. *Proceedings of the National Academy of Sciences of the United States of America* **111,** 863-868.

Woo J, von Arnim AG. 2008. Mutational optimization of the coelenterazine-dependent luciferase from *Renilla*. *Plant Methods* **4,** 23.

Wu P, Nicholls SB, Hardy JA. 2013. A tunable, modular approach to fluorescent protease-activated reporters. *Biophysical Journal* **104,** 1605-1614.

Xu X, Gerard ALV, Huang BCB, Anderson DC, Payan DG, Luo Y. 1998. Detection of programmed cell death using fluorescence energy transfer. *Nucleic Acids Research* **26,** 2034-2035.

Yang J, Zhang Z, Lin J, Lu J, Liu B-f, Zeng S, Luo Q. 2007. Detection of MMP activity in living cells by a genetically encoded surface-displayed FRET sensor. *Biochimica et Biophysica Acta - Molecular Cell Research* **1773,** 400-407.

Zhang B. 2004. Design of FRET-based GFP probes for detection of protease inhibitors. *Biochemical and Biophysical Research Communications* **323,** 674-678.

Zhang D, Liu D, Lv X, Wang Y, Xun Z, Liu Z, Li F, Lu H. 2014. The cysteine protease CEP1, a key executor involved in tapetal programmed cell death, regulates pollen development in *Arabidopsis*. *Plant Cell* **26,** 2939-2961.

Zhang J, Wang X, Cui W*, et al.* 2013. Visualization of caspase-3-like activity in cells using a genetically encoded fluorescent biosensor activated by protein cleavage. *Nature Communications* **4,** 2157.

Zhang L, Xu Q, Xing D, Gao C, Xiong H. 2009. Real-time detection of caspase-3-like protease activation in vivo using fluorescence resonance energy transfer during plant programmed cell death induced by ultraviolet C overexposure. *Plant Physiology* **150,** 1773-1783.

Zhang Y, Ke X, Zheng C, Liu Y, Xie L, Zheng Z, Wang H. 2017. Development of a luciferase-based biosensor to assess enterovirus 71 3C protease activity in living cells. *Scientific Reports* **7,** 10385.

Research Chapters

Chapter 4

"Damage on plants activates Ca^{2+}-dependent metacaspases for release of immunomodulatory peptides".

Tim Hander[1#], Álvaro Daniel. Fernández-Fernández[4,5#], Robert P. Kumpf[4,5], Patrick Willems[4,5,6,7], Hendrik Schatowitz[1], Debbie Rombaut[4,5], An Staes[6,7], Jonah Nolf[4,5], Robin Pottie[4,5], Panfeng Yao[4,5], Amanda Gonçalves[2], Benjamin Pavie[2], Thomas Boller[1], Kris Gevaert[6,7], Frank Van Breusegem[4,5], Sebastian Bartels[1,3], Simon Stael[4,5,6,7*]

[1] Zürich-Basel Plant Science Center, Department of Environmental Sciences, Botany, University of Basel, CH-4056 Basel, Switzerland.

[2] VIB BioImaging Core Gent, VIB-UGent Center for Inflammation Research (IRC), 9052 Ghent, Belgium

[3] Department of Medicine II, University Hospital Freiburg - Faculty of Medicine, University of Freiburg, Freiburg 79106, Germany.

[4] Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium

[5] VIB-UGent Center for Plant Systems Biology, 9052 Ghent, Belgium

[6] Department of Biomolecular Medicine, Ghent University, 9000 Ghent, Belgium

[7] VIB-UGent Center for Medical Biotechnology, 9000 Ghent, Belgium

[#]These authors contributed equally to this work.

Author contributions:

T.H., A.D.F.F, S.B., and S.S. conceived and designed the analysis. T.H., A.D.F.F, R.P.K., H.S., D.R., A.S., J.N., P.Y., R.P., A.G., B.P., and S.S. performed the experimental work. T.H. and S.S. wrote the manuscript. T.B., K.H., F.V.B., and S.B. revised the manuscript and were involved in the discussion of the work. Competing interests: The authors declare no competing interests. Data and material availability: All data to support the conclusions of this manuscript are included in the main text and supporting online material. The full MS data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifiers PXD010816 and PXD005740. All materials are available on request, including chemical compounds as supplies permit, subject to a standard material transfer agreement.

Aim and context

This research chapter focuses on the processing of PROPEP1 by AtMC4, in wounding stresses in plants. We characterize the local calcium dynamics in wounding and evaluate the different regions using laser ablation in Arabidopsis roots. The dependence on calcium levels for AtMC4 self-processing and subsequent cleavage of of PROPEP1 is addressed. Moreover it is shown the requirements of an arginine at the position preceding metacaspases processing for the release of the mature Pep1. This study shows the role of metacaspases in processing the elicitor peptide Pep1 and likely other PROPEP family members. Our observations suggests that type-II metacaspases can work as modulators of stress by translating calcium signals through proteolysis to peptide signalling.

Abstract:

Physical damage to cells leads to the release of immunomodulatory peptides to elicit a wound defense response in the surrounding tissue. In *Arabidopsis thaliana*, the plant elicitor peptide 1 (Pep1) is processed from its protein precursor, PRECURSOR OF PEP1 (PROPEP1). Here, we demonstrate that upon damage, both at tissue and single-cell level, the cysteine protease METACASPASE4 (MC4) is instantly and spatiotemporally activated by binding high levels of Ca^{2+} and is necessary and sufficient for Pep1 maturation. Cytosol-localized PROPEP1 and MC4 react only after loss of plasma membrane integrity and prolonged extracellular Ca^{2+} entry. Our results reveal that a robust mechanism consisting of conserved molecular components links the intracellular and Ca^{2+}-dependent activation of a specific cysteine protease with the maturation of damage-induced wound defense signals.

Keywords: metacaspases, proteomics, *Arabidopsis thaliana*, N-terminomics, proteases, protease inhibitors, degradomics, wounding.

INTRODUCTION

Physical trauma of plants can be provoked, for example, by insect chewing, animal feeding and trampling, or weather damage, thereby creating entry sites for various pathogens. From damaged plant cells, a multitude of cellular constituents that act as signaling factors are released to alert and activate a defense response in the surrounding tissues against invading pathogens. These signaling factors, or damage-associated molecular patterns (DAMPs), include extracellular ATP and DNA, cell wall-derived molecules, peptides, and glutamate (Heil and Land, 2014; Toyota et al., 2018).

In *Arabidopsis thaliana*, the plant elicitor peptides (Peps) were proposed to act as DAMPs based on their intracellular origin and ability to elicit defense responses when applied as synthetic peptides (Huffaker et al., 2006). The eight Arabidopsis Peps are embedded in the C-terminus of their respective precursor proteins, the PROPEPs (Bartels and Boller, 2015). After the supposed proteolytic maturation of PROPEPS, the Peps are perceived by two plasma membrane-localized Leucine-Rich Repeat Receptor-Like Kinases (LRR-RLKs) named PEP RECEPTOR 1 (PEPR1) and PEPR2 (Krol et al., 2010; Yamaguchi et al., 2010). Upon Pep perception, PEPRs interact with the coreceptor BRI1-ASSOCIATED KINASE1 (BAK1) (Yamaguchi et al., 2010; Tang et al., 2015) to induce a typical innate immunity-like response leading to an oxidative burst, ethylene (a gaseous plant stress-related hormone) production, and pathogen defense-related gene expression (Bartels et al., 2013). In the absence of both PEPRs, the pathogen and wound defense responses are reduced, increasing pathogen sensitivity and insect herbivory (Tintor et al., 2013; Ross et al., 2014; Klauser et al., 2015; Yamada et al., 2016). Although the roles of Peps as wound-immunomodulatory peptides are conserved across diverse plant families, including in the monocot Zea mays (maize) (Huffaker et al., 2013), the mechanism by which Peps are released from their precursors has been elusive. Here, we demonstrate that the cysteine protease MC4 is necessary and sufficient for cleavage and release of Pep1 from PROPEP1 in an immediate, spatiotemporally controlled, Ca^{2+}-dependent manner.

RESULTS

Damage induces instant PROPEP1 processing

To assess the Pep maturation mechanism, we established two experimental setups in which the PROPEP cleavage kinetics after plant tissue damage were monitored and perturbed. Whole seedlings of transgenic plant lines expressing PROPEP1-YFP fusion proteins (Bartels et al., 2013) were damaged with serrated forceps or by thawing of snap-frozen and crushed tissue (Fig. 1A). In both experimental setups, we evaluated the PROPEP1-YFP cleavage kinetics through immunoblot analysis. Prompt handling of the samples (see Materials and methods) was essential to prevent cleavage during or after protein extraction. Upon thawing of the frozen tissue samples the PROPEP1-YFP band was processed into a band with the approximate molecular weight of a Pep1-YFP fusion protein (PEP1-YFP), peaking at 5 minutes, followed by general protein degradation (Fig. 1C and D, fig. S1A). Accumulation of PEP1-YFP was evident after 30 seconds (Fig. 1D, fig. S1A). In the forceps-damaged samples, PEP1-YFP cleavage products accumulated more evenly over time compared to thawed tissue powder (Fig. 1B).

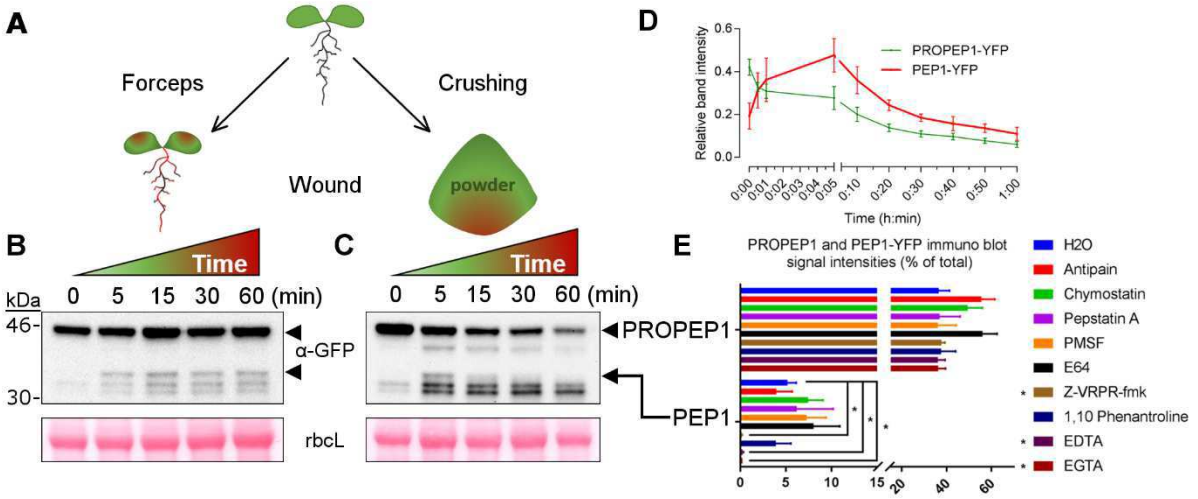


Fig. 1 Rapid PROPEP1 processing to Pep1 in damaged plants. (A) Visual representation of wound model experiments. (B and-C) PROPEP1-YFP (top arrowhead in the immunoblot) is processed to PEP1-YFP (bottom arrowhead) after incubation of forceps-wounded seedlings (B) or crushed tissue powder of whole seedlings (C) at room temperature for the indicated time (min = minutes). Ponceau S stain of the rubisco large subunit (rbcL) indicates protein load. (D) Immunoblot band intensity quantification of PROPEP1-YFP and PEP1-YFP over time in crushed tissue powder. Time points represent the mean of five biological replicates ± standard error on the mean (SEM). (E). PROPEP1-YFP and PEP1-YFP band intensity, 5 minutes after transfer of crushed tissue powder to room

temperature in the presence of water or the indicated protease inhibitor. Mean band intensities are shown for three biological replicates + SEM. Statistics are described in the Material and methods.

Metacaspase activity correlates with Pep1 cleavage specificity

To identify the protease activities necessary to process PROPEP1, we used a pharmacological approach in which whole seedlings were infiltrated with a range of protease inhibitors and ion-chelators 10 minutes prior to the damage. Whereas the majority of inhibitors did not affect PROPEP1-YFP cleavage (fig. S1C), PEP1-YFP fragment formation was 96% reduced on average by infiltration of either the ion chelators EGTA and EDTA or the metacaspase inhibitor Z-VRPR-fmk (Vercammen et al., 2006) (Fig. 1E, fig. S1C).

Metacaspases are a family of cysteine-dependent proteases that together with the metazoan paracaspases, such as mucosa-associated lymphoid tissue lymphoma translocation protein 1 (MALT1), were suggested to behave as caspase-like proteins (Uren et al., 2000). This view has been challenged, because unlike caspases that cleave C-terminally to aspartic acid, metacaspases cleave their protein substrates C-terminally to the basic amino acids arginine (R) and lysine (K) embedded in a double basic (R/K)x(R/K) cleavage signature (Vercammen et al., 2004). Deviation from this pattern occurs, for instance, for MC9 substrates (Tsiatsiani et al., 2013; Wrzaczek et al., 2015), but the cleavage takes place after R or K. PROPEP1 also lacks a double basic motif. However, mutation of the conserved arginine (R69 in Arabidopsis PROPEP1) that precedes the demonstrated or predicted mature Peps ((Huffaker et al., 2006); Fig. 2A) to glutamate (E) or alanine (A) blocked the degradation of the PROPEP1-YFP protein to PEP1-YFP *in vivo* (Fig. 2B and C). An R-specific cleavage occurred, because cleavage was impaired similarly by a mutation into the homotypic amino acid K (Fig. 2B).

Damage activates MC4

The Arabidopsis genome encodes nine metacaspases, classified according to their domain organization and biochemical characteristics. Type-I metacaspases, MC1, MC2, and MC3, contain a proline/glutamine-rich repeat domain and a zinc finger motif in their N-terminal prodomain. MC1 and MC2 are involved in a plant-specific programmed cell death (PCD), typical for a hypersensitive response(Coll et al., 2010),

but their activation mechanism and proteolytic requirements remain unknown (Tsiatsiani et al., 2011). Type-II metacaspases, MC4 to MC9, lack the N-terminal prodomain and most require Ca^{2+} concentrations in the low millimolar range and a pH optimum of 7.5 for activation in vitro and autocatalytic cleavage to p20 and p10 domains (designated according to the caspase nomenclature for their approximate molecular mass in kilodaltons). Crystal structures for type-I metacaspases from Trypanosoma brucei and Saccharomyces cerevisiae revealed that Ca^{2+} binding to four aspartate residues induced a conformational shift and was necessary for auto-activation (McLuskey et al., 2012; Wong et al., 2012). Based on homology to Type-I metacaspase sequences, point mutation and domain swap experiments revealed similar conserved aspartate residues important for the activity in tomato (*Solanum lycopersicum*), tobacco (*Nicotiana tabacum*), and Arabidopsis type-II metacaspases (Wen et al., 2013; Acosta-Maspons et al., 2014; Fortin and Lam, 2018). In contrast, MC9 does not depend on Ca^{2+} and functions optimally at pH 5.5 (Vercammen et al., 2004; Watanabe and Lam, 2011a). Considering the ubiquitous expression pattern of Arabidopsis PROPEP1 and MC4 (Watanabe and Lam, 2011b; Bartels et al., 2013), which would be a fitting prerequisite for a general wound response regulator, we wondered whether damage activates MC4. With a MC4-specific antibody that recognizes both the zymogen (zMC4) and the activated p20 domain (Watanabe and Lam, 2011a, b) (fig. S1E and F), autocatalytic MC4 processing was monitored during damage (Fig. 2D; quantified in Fig. 2I and K). In Ler plants, a MC4-derived band at an approximate size of 30 kDa, which we named p20*, decreased in intensity (Fig. 2D and I), most probably because the p20-p10 linker region C-terminal of p20* is trimmed, resulting in a p20 size of approximately 26 kDa. This observation is in agreement with N-terminal peptides (Table 1) (Willems et al., 2019)and endogenous peptides from peptidomics (table S1) identified by mass spectrometry and indicative of cleavage events upstream of the initial autocatalytic cleavage at K255 (Watanabe and Lam, 2011b). Previously, trimmed p20 protein fragments were also identified *in vitro* and found to be active against a chemical test substrate, biotin-FPR-cmk (Watanabe and Lam, 2011a). Therefore, we used the differential accumulation of the p20* and p20 bands in time as a proxy for MC4 activation. MC4 was activated with kinetics similar to those of PROPEP1 cleavage (fig. S1B). Both events were blocked by the addition of EDTA and EGTA (fig. S1D). Z-VRPR-fmk did not inhibit p20 accumulation (fig. S1D), but inhibited PROPEP1 cleavage (fig. S1C), indicating a sequence of events: Ca^{2+}

binds zMC4 to induce autocatalytic cleavage (p20 accumulation is inhibited by EDTA and EGTA (fig. S1D)) and subsequently, active MC4 can cleave substrates, such as PROPEP1, or inhibitory substrate analogs, such as Z-VRPR-fmk.

MC4 cleaves PROPEP1 at R69

To assay direct MC4-dependent maturation of Pep1, we tested *in vitro* whether purified recombinant His-tagged MC4 could process a PROPEP1-glutathione S-transferase (GST) fusion protein. MC4 efficiently cleaved PROPEP1-GST within the nanomolar range (Fig. 2E). Similarly to the *in vivo* situation, mutated PROPEP1R69A/E-GST cleavage sites were not cleaved by MC4 *in vitro* (Fig. 2E). The homotypic R69K mutation was cleaved, albeit less efficiently: the PROPEP1-GST band decreased and PEP1-GST increased 50%, whereas for the R69K mutation they decreased and increased 30%, respectively, at a concentration of 62.5 nM recombinant MC4 (rMC4; Fig. 2G). Despite the occurrence of extra cleavage sites downstream of R69, based on the size and intensity of the cleavage products (Fig. 2F and G), R69 of PROPEP1 remains seemingly the preferred MC4 cleavage site, both *in vivo* and *in vitro*.

Other PROPEP1-YFP cleavage products, such as the double band just below PEP1-YFP, can be distinguished by immunoblots (Fig. 1B and C). Anti-GFP immunoprecipitation combined with mass spectrometry revealed these bands (bands 4 and 5 in fig. S2A and F) as cleavage products downstream of R69 (table S2). The bands are mostly present under steady state condition (time point 0 in the PROPEP1-YFP immunoblots) and become more abundant after damage (fig. S2G). At least during forceps-induced damage, accumulation of band 4 might depend on MC4 activity, because it accumulated less in a MC4 deficient mutant mc4(-/-) (fig. S2G), but is not necessarily a direct target of MC4; because in crushed seedlings, it accumulated irrespective of MC4 (fig. S2C). MC4 may initially cleave at R69, followed by downstream proteolysis of PEP1-YFP by MC4 (Fig. 2E-G) or other undefined proteases. Furthermore, different cleavages occur in crushed when compared to forceps-treated seedlings (fig. S2C), such as a peptide (band 2 and 7 in fig. S2A and F) that was detected by endogenous peptidomics (table S1) and can be distinguished by a human influenza hemagglutinin (HA) tag C-terminally to YFP (fig. S2B and E).

In mc4(-/-), the PEP1-YFP and PROPEP1-YFP levels did not increase and concurrently decrease, respectively, after 5 minutes (quantified in Fig. 2H and J) and

at least up to 1 hour after damage (fig. S2C and G). As additional control, PEP1-YFP maturation was not affected in a MC9-deficient mutant (fig. S2D and G). Our results demonstrate that upon physical damage, PROPEP1 is cleaved *in vitro* and *in vivo* by Ca^{2+}-dependent MC4 activity at R69 into the mature Pep1 immunomodulatory signaling peptide.

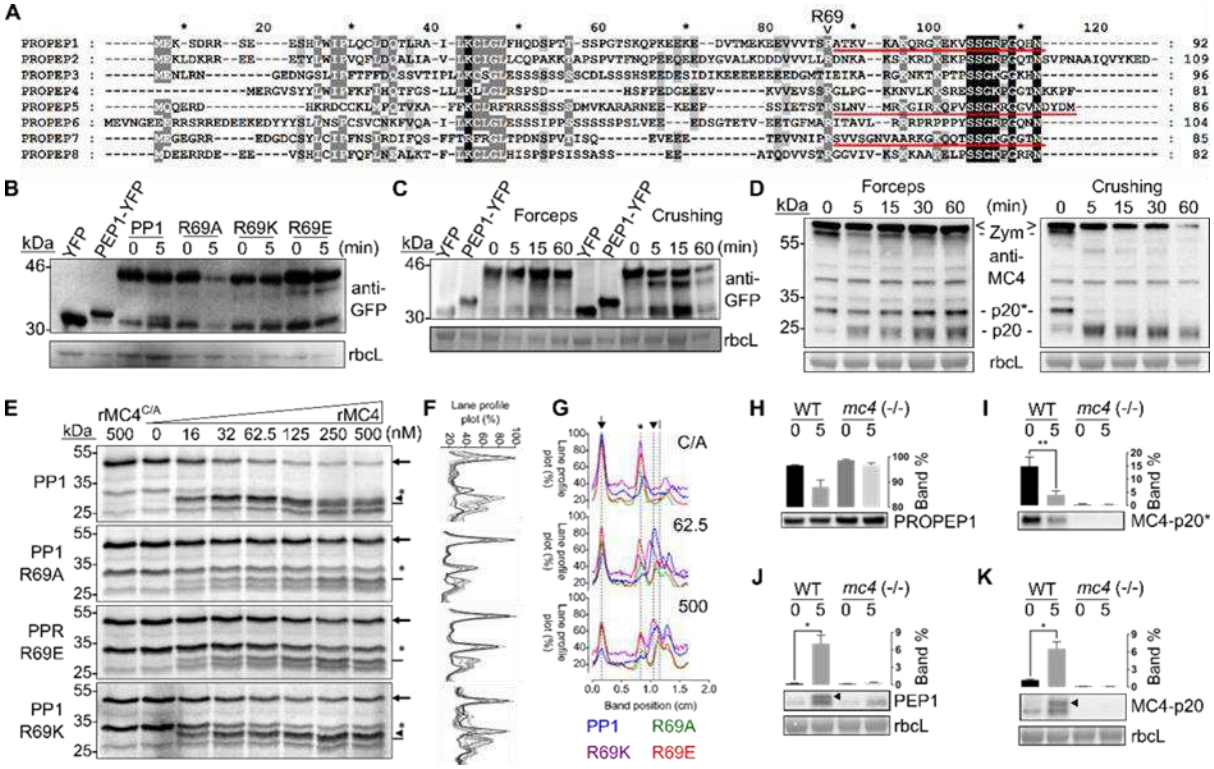


Fig. 2 Activation of MC4 upon damage and cleavage of PROPEP1 behind R69 to release Pep1 *in vitro* and *in vivo*. (A) Multiple sequence alignment of the eight Arabidopsis PROPEPs. Previously identified Peps (by mass spectrometry) are underlined in red. Arginine 69 = R69. (B) Replacement of PROPEP1 (PP1) R69 with A, K, or E impaired forceps-induced accumulation of PEP1-YFP *in vivo*. R69A mutation seemed to destabilize PROPEP1-YFP, leading to its accelerated degradation after wounding. (C) R69E mutation impaired PEP1-YFP accumulation at least up to 1 hour after damage by forceps (left blot part) and crushing (right blot part). Extracts from Arabidopsis lines expressing YFP and PEP1-YFP were loaded to indicate their respective band heights. (D) Anti-MC4 immunoblots of both wound model experiments for the indicated time points. The top bands indicated with (<) and (>) are unspecific (see fig. S1E and F for more information on anti-MC4 specificity). (E) *In vitro* TNT®-protease assay of PROPEP1 (PP1) and mutant versions incubated with increasing amounts of recombinant MC4 (rMC4) or its inactive mutant rMC4C/A (active site cysteine mutated to alanine). The arrow and arrowhead represent PROPEP1 fused to glutathione S-transferase (PROPEP1-GST) and PEP1-GST, respectively. At increased rMC4 concentrations extended cleavage occurs at a downstream site, potentially the lysine (K)-rich region between R69 and the conserved Pep motif SSG-(R/K)x1-G-x2-N, resulting in a close-by lower band indicated with a dash (-). The band indicated with an asterisk (*) is most probably an alternative initiation product from a downstream methionine (M59), because it cleaved

in the wild type as well, but not in the mutant versions R69A and R69E. (F) Relative band intensity profiles of the images in (E) overlaid for the different rMC4 concentrations per construct. Black trace is the rMC4 C/A lane and light gray the 500 nM rMC4 lane. (G) Intensity profile overlay relative to the first peak for the indicated concentrations and constructs. (H to K) Quantification of mean immunoblot band intensities for three biological replicates + SEM. (H) PROPEP1-YFP, (I) PEP1-YFP (◄), (J) MC4 p20*, and (K) MC4 p20 subunit (◄), 5 minutes after damage by forceps. Statistics are described in the Material and methods. rbcL indicates protein load.

Wounding triggers distinct Ca^{2+} patterns in root epidermal cells

Ca^{2+} concentrations in the low millimolar range are needed to activate MC4 *in vitro* (Vercammen et al., 2004). We showed that the Ca^{2+} chelators EGTA or EDTA impaired MC4 activation and PROPEP1 cleavage *in vivo* after tissue damage. Ca^{2+} functions as a secondary messenger in signal transduction networks in all eukaryotic organisms and many stimuli can trigger increases in cytosolic Ca^{2+} concentrations or so-called [Ca^{2+}]cyt spikes (Stael et al., 2012). However, cleavage of PROPEP after every [Ca^{2+}]cyt spike seems counter-productive. In an effort to link the spatiotemporal dynamics of [Ca^{2+}]cyt fluxes, metacaspase activation, and Pep maturation, we focused on roots. Pep1 triggers a stronger defense gene expression in the root than typical bacterial or fungal elicitors (Poncini et al., 2017). Although roots have not often been studied for wounding studies, they are constantly in contact with beneficial, but also potentially harmful, microorganisms as well as with invasive organisms, such as nematodes, and thus are relevant for tissue damage investigations. Furthermore, Pep1 maturation occurs similarly in root tissues (fig. S3A, B, E and F) as in leaf tissues (fig. S3C and G) and whole seedlings (Fig. 1B and C), but, in contrast to leaf tissues, MC4 is not exclusively responsible for Pep1 maturation, because PEP1-YFP accumulation is only halved in mc4 (-/-) root tissues (fig. S3D). Because Z-VRPR-fmk reduces cleavage better than mc4(-/-) (fig. S3H), residual Pep1 maturation in roots is most probably due to redundancy with other (Ca^{2+}-dependent type-II) metacaspases, or with undefined proteases with similar inhibition characteristics.

Cell damage was inflicted by multiphoton laser wounding to obtain the highest possible spatial and temporal resolution during imaging [Ca^{2+}]cyt with a Yellow Cameleon Ca^{2+} probe (YC3.60-NES) in epidermal transition zone root cells (Fig. 3A), revealing detailed spatiotemporal [Ca^{2+}]cyt differences after damage (Fig. 3B; movies S1 and S2). [Ca^{2+}]cyt responses could be grouped according to their duration and intensity in

four zones in and surrounding the wound site (Fig. 3D and E). Propidium iodide (PI) entry into the cell, as a proxy for plasma membrane integrity loss, marked the damaged cells of zone 1 (Fig. 3B) that accumulated long-lasting high $[Ca^{2+}]$cyt (Fig. 3E). Zones 2-4 surrounding the damaged cells displayed $[Ca^{2+}]$cyt spikes that returned to resting $[Ca^{2+}]$cyt. The cells closest to the wound experienced the largest transient spikes and a $[Ca^{2+}]$cyt 'wave' traveled outward beyond zone 4 (note that zone 4 also peaks slightly later than zones 2 and 3) with an average speed of 1.63e-6 ± 0.4 m/s. Similar $[Ca^{2+}]$cyt dynamics were observed in Arabidopsis leaves following mechanical damage (Costa et al., 2017; Behera et al., 2018) and in laser-ablated embryonic cells or larvae of Drosophila melanogaster (Razzell et al., 2013; Shannon et al., 2017). Thanks to the subtle damage inflicted by the multiphoton laser, we could obtain a similar or higher cellular resolution and enable the imaging of cells inside the wound (Zone 1). $[Ca^{2+}]$cyt peaks were reduced in zones 2-4 by addition of 1 mM EGTA (Fig. 3E; movie S2) and increasing concentrations of the alternative Ca^{2+} chelator BAPTA (fig. S4A and B), hinting at an extracellular source of Ca^{2+}. Zone-1 cells showed a residual, but considerable $[Ca^{2+}]$cyt increase in the presence of 1 mM EGTA (Fig. 3E) or 0.5 mM BAPTA (fig. S4A). This might be indicative of Ca^{2+} release from internal stores that is chelated when EGTA or BAPTA enters the damaged cells or of such substantial (extracellular or intracellular) amounts that are not immediately chelated by either compound.

Pep1 maturation correlates with Ca^{2+} patterns

In independent Arabidopsis lines expressing p35S::PROPEP1-YFP, the fluorescent signal localized to the cytosolic face of the vacuolar membrane (tonoplast) in normal conditions (Fig. 3C-1st vertical panel, fig. S5A and D) (8). When laser wounding, as explained for the calcium reporters above, was applied in the p35S::PROPEP1-YFP lines, the YFP fluorescence delocalized to the cytosol only in damaged cells accumulating PI (zone 1) (Fig. 3C-2nd to 5th vertical panel and F; movies S3 and S6). Two types of cellular events occurred in zone 1 cells: (i) the vacuole almost immediately collapsed after delocalization and (ii) the vacuole and thereby the whole cell stay relatively intact after delocalization (Fig. 3C). Occasionally, cells accumulated some PI without YFP delocalization (# in Fig. 3C), hinting at a certain threshold of membrane integrity loss for PROPEP1-YFP processing. This was also apparent from the BAPTA dilution experiment (fig. S4), where addition of 0.25 mM BAPTA lowered,

but did not alter the shape of the [Ca^{2+}]cyt transient (fig. S4A), whereas YFP delocalization was inhibited (fig. S4C). Furthermore, addition of EGTA or Z-VRPR-fmk reduced YFP delocalization in the epidermal cells of the transition zone (Fig. 3F; fig. S5A; movies S4 and S5) and the maturation zone (fig. S5C and D). These results underscore the localized heterogeneity of [Ca^{2+}]cyt fluxes in zones 2 to 4 and the generation of passive near-millimolar [Ca^{2+}]cyt transients in zone-1 damaged cells. Furthermore, PROPEP1 processing, evidenced by YFP delocalization, correlates with [Ca^{2+}]cyt abundance in zone-1 cells and is localized to damaged cells only. Consequently, the number of wounded cells could convert a digital cellular output (generation or no generation of Pep1) into an analogue response, proportional to the damage.

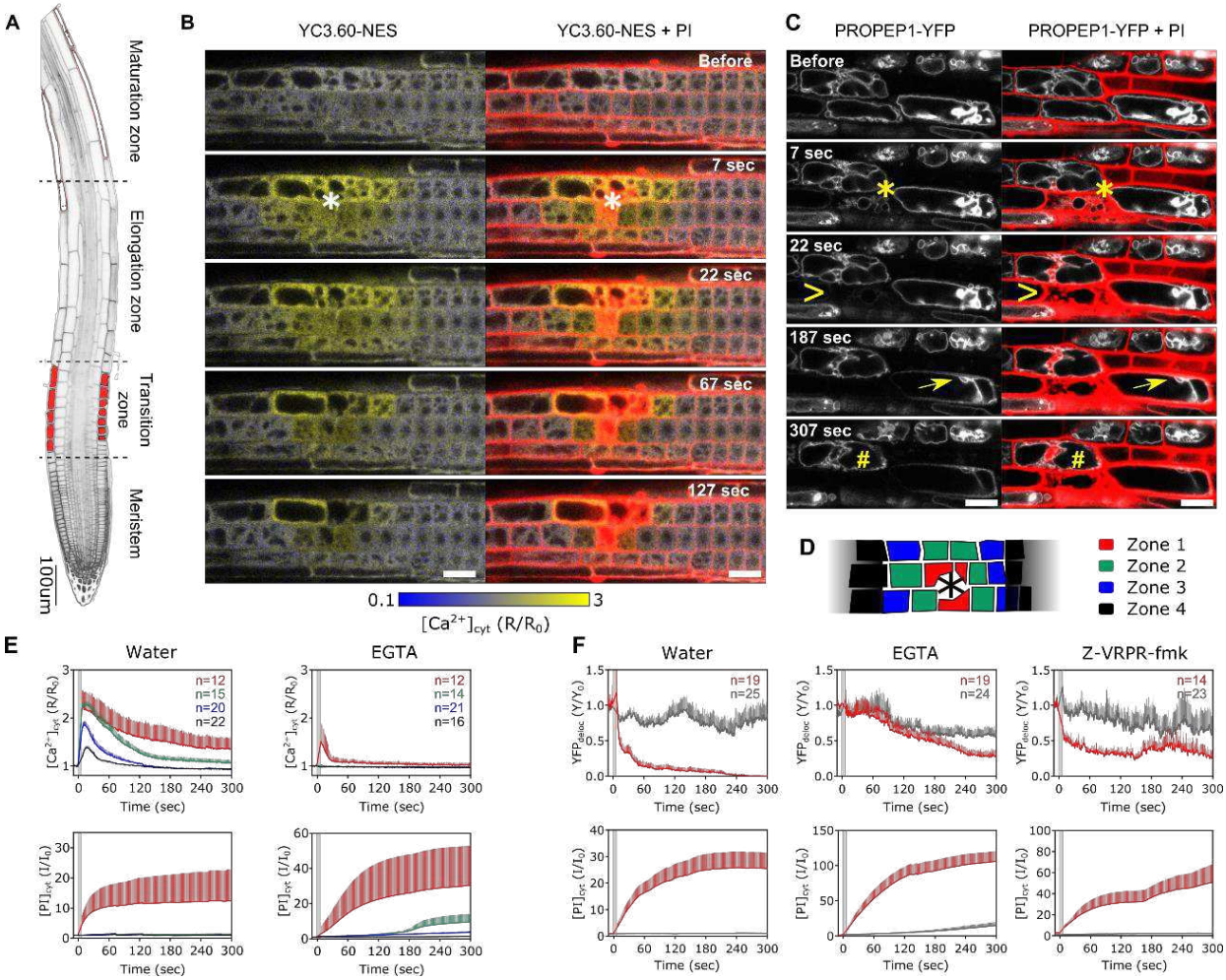


Fig. 3 Correlation of local [Ca^{2+}]cyt fluxes and PROPEP1 processing in laser-damaged root cells. (A) Arabidopsis root tip with the epidermal cells of the transition zone that were targeted by laser (filled in red). (B and C) Confocal microscopy time series of laser wounding (site indicated with an asterisk) of the YC3.60-NES Ca^{2+} probe and overlaid with PI stain (B), and PROPEP1-YFP delocalization from the

tonoplast (sharp edges) to the cytosol (more diffuse signal) over time (C). Damaged cells either collapsed immediately after delocalization (>) or stay relatively intact after delocalization (arrow). The hashtag (#) indicates intracellular PI accumulation without PROPEP1-YFP delocalization. (D) Hypothetical zones in a laser-inflicted root wound. Color gradient outside the zones symbolizes the spread of the $[Ca^{2+}]$cyt flux. (E) $[Ca^{2+}]$cyt flux (top) and accumulation of PI (bottom) over time for the zones colored according to (D) in pure water or in the presence of 1 mM EGTA. (F) Quantification of YFP delocalization in cells from zone 1 and zones 2, 3, and 4 combined in pure water or in the presence of 1 mM EGTA or 50 µM Z-VRPR-fmk. Curves are the mean of N cells +SD in (E) and + SEM in (F). Grey vertical bars, spanning 7 seconds in (E) and (F) indicate the average lag time between the laser shot and imaging start. Scale bar = 20 µm.

Subcellular retention constrains Pep1 signaling

PROPEP1, which lacks a canonical transmembrane domain, is seemingly constrained at the tonoplast through its N-terminus. Indeed, a double labelled fluorophore construct with mCitrine fused to the N-terminus of PROPEP1 and red fluorescent protein (RFP) fused to its C-terminus localized to the tonoplast, but only released the RFP moiety to the cytosol upon damage (Fig. 4A; movie S7). Furthermore, PEP1-YFP was present and PROPEP1-YFP was absent in the soluble phase, when seedling extracts were centrifuged to remove solid and membrane fractions (Fig. 4E). A GST-PEP1 fusion protein can bind the extracellular domains of PEPR1 *in vitro* (Yamaguchi et al., 2010), so the N-terminus of Pep1 does not need to be free to interact with its receptor per se. Additionally, we found that exogenous application of purified N-terminally truncated PROPEP1 (Δ39-PP1; fig. S6C) triggers root growth inhibition - a known negative effect of Pep1 application (Krol et al., 2010) - irrespective of cleavage by rMC4 (Fig. 4B and C).

The prevailing concept that proteases and substrates are sequestered in separate subcellular compartments and only mix upon cellular disintegration does not apply here. The vacuole, in which proteases are stored, changes shape, rounding up, and becomes immobile during PROPEP1-YFP delocalization, although it stays intact until after PROPEP1-YFP delocalization, when the vacuole can burst (Fig. 3C and 4A; fig. S5; movie S3 and S6).

The Pep1 signal is probably short-lived in cells in which rapid vacuolar collapse can cause further ‘trimming’ of the Pep1 peptide (Fig. 3C) as mirrored in the grinding damage investigated in Fig. 1C and D. Truncated versions of synthetic Pep1 (similar

to bands 4 and 5 in fig. S2) had previously been shown to lose their activity (Pearce et al., 2008). Conversely, prolonged Pep1 accumulation upon forceps application (Fig. 1B) is most probably caused by damaged cells with stable vacuoles (Fig. 3C) or the continuous, unsynchronized loss of membrane integrity during wounding (movie S6). Taken together, MC4-dependent Pep1 maturation seems to mainly facilitate release from the tonoplast membrane to promote Pep1 motility into the surrounding tissue. This passage needs to be guarded because Pep1 overload negatively affects plant growth (Fig. 4B and C) (Krol et al., 2010).

Conservation of Pep maturation

A family member of the subtilisin-like serine proteases or subtilases, called phytaspase (Beloshistov et al., 2018), processes the wound signaling peptide prosystemin (McGurl et al., 1992). Other subtilases are also relevant for development-related peptide hormone processing events (Ghorbani et al., 2016; Schardon et al., 2016). Whereas systemin is specific to the Solanaceae, PROPEP and metacaspase genes are found throughout the Angiosperms (Lori et al., 2015). We found that MC4 is able to cleave other Arabidopsis PROPEPs and the tomato ortholog of PROPEP1 *in vitro* as well (Fig. 4D; fig. S6A and B), suggesting that metacaspase function in Pep maturation is conserved across plants.

MODEL AND FUTURE PERSPECTIVES

Taken together, a model emerges in which MC4 and PROPEP1 both remain inactive in the cytosol until a loss of plasma membrane integrity in the damaged zone-1 cells leads to a prolonged increase in intracellular [Ca^{2+}] (Fig. 4F). Ca^{2+} mainly originates from the extracellular space and potentially from internal stores, such as the vacuole (Fig. 4F), and binds zMC4 to initiate autocatalytic cleavage, where active MC4 can cleave PROPEP1. The current approach characterized Pep1 delocalization and calcium levels in independent lines. Ideally, both calcium levels and PROPEP1 processing in the same line can be visualized indicating a better correspondence with the calcium fluxes and effect on the cleavage. By such, a better delineation of the calcium dynamics in the local wounding zones could be directly related to Pep maturation and delocalization. For instance, this approach, could be performed using

a C-terminal tagged PROPEP with a stable fluorescent protein (mScarlet or mNeonGreen) and single color calcium reporters that do not overlap with the tagged PROPEP1 like GCaMP or R-GECO variants (Keinath et al., 2015; Krogman et al., 2020). Pep1 is released from the tonoplast to the cytosol from where it can passively diffuse or be potentially actively secreted through the compromised plasma membrane to bind the extracellular domains of the BAK1-PEPR1/2 receptor kinase complex and to signal the surrounding intact cells of zones 2-4 to activate a defense response (Fig. 4F).

Sequence diversity of PROPEPs and PEPRs implies that PROPEPs evolve quickly. Although PEPR signaling seems to be preserved, its loss might go unnoticed in unstressed plants (Lori et al., 2015). This multi-component system may have been deactivated or even lost during plant breeding programs. Indeed, PROPEPs, PEPRs, and metacaspases may be useful targets for marker-assisted breeding or CRISPR-Cas mutagenesis to improve disease resistance of high-performance crop varieties. The ameliorative effect of Pep-mediated PEPR signaling against a variety of diseases caused by microbes and herbivores (Tintor et al., 2013; Klauser et al., 2015) should become useful as global warming leads to increased crop losses due to insect pests (Deutsch et al., 2018).

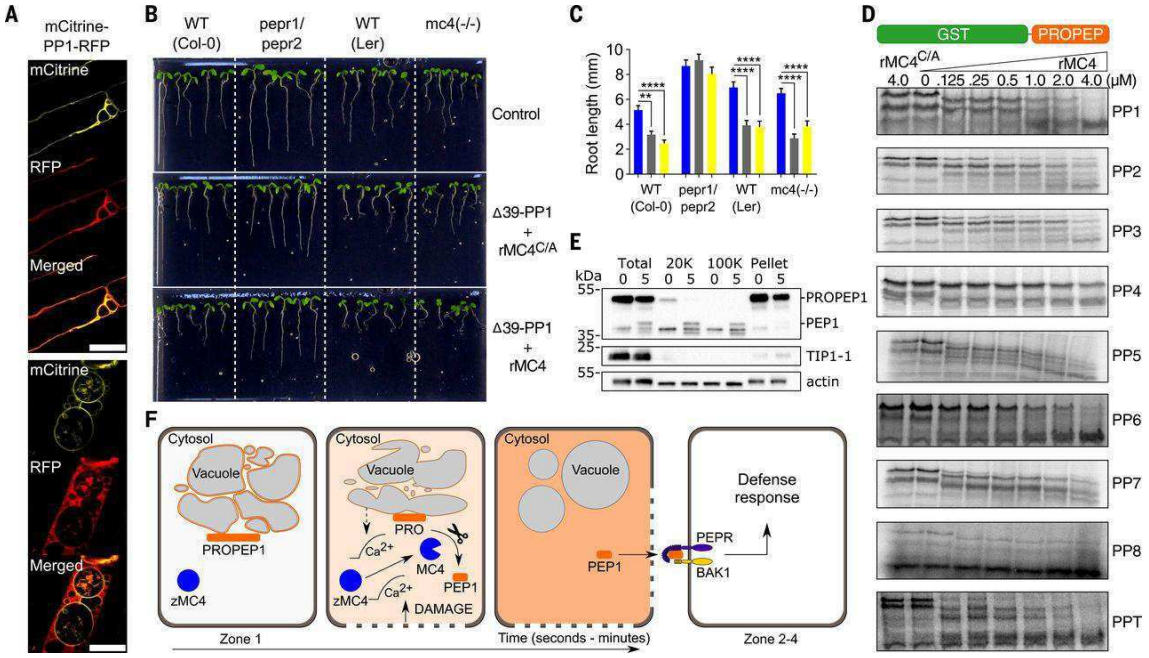


Fig. 4 Control, conservation and hypothetical model of Pep1 release. (A) A mCitrine-PROPEP1-RFP fusion protein in control (top) and laser-damaged root epidermal cells (bottom). Scale bar = 20 µm. (B) Arabidopsis seedlings 5 days after transfer to medium containing 25 nM of a N-terminally truncated

PROPEP1 protein (Δ39-PP1) treated with active rMC4 or inactive rMC4C/A. The double mutant deficient for the PEP receptors (pepr1/pepr2) served as negative control. (C) Quantification of root length in (B). Values are the average of three biological repeats with 45 seedlings in total and plotted as mean + SEM. Statistics are described in the supplementary Material and methods. Blue, control; gray, ΔN-PROPEP1 + rMC4C/A,; and yellow, ΔN-PROPEP1 + rMC4. (D) TNT®-protease assays of GST-PROPEP fusion proteins (Arabidopsis PP1 to PP8) and the tomato PROPEP1 ortholog (PPT). All assays were repeated at least twice and were run together with PP1 as a positive control. See also fig. S6A and B for interpretation of results. (E) Immunoblots for PROPEP1-YFP (anti-GFP), the aquaporin TONOPLAST INTRINSIC 1-1 (anti-TIP1-1), and actin (anti-actin) of total protein extracts from seedling at 0 and 5 minutes of forceps-induced damage, a soluble fraction after centrifugation at 20,000g (20K), 10,0000g (100K), and the pellet fraction from 100K. (F) A hypothetical model for damage-induced Pep1 maturation.

MATERIAL AND METHODS

Plant material and treatments

For the preparation of sterile seedlings, *Arabidopsis thaliana* (L.) Heyhn. seeds were surface sterilized with 70% ethanol and plated on half-strength Murashige and Skoog (½MS) medium supplemented with 1% (w/v) sucrose and 0.5% (w/v) Phytagel (Sigma-Aldrich), stratified for at least 2 days at 4°C, and then germinated at 21°C under continuous light (MLR-350 Plant growth chamber; Sanyo). After 5 days, individual seedlings were transferred to liquid ½MS medium with 1% (w/v) sucrose and grown for an additional 9 days. Seedlings treated with the protease inhibitors antipain (100 μM), chymostatin (100 μM), pepstatin A (1 μM), PMSF (1 mM), E64 (10 μM), 1,10-phenanthroline (20 mM), Z-VRPR-fmk (50 μM), EDTA (1 mM), ethylene glycol tetra-acetic acid (EGTA) (1 mM), and Protease Inhibitor Cocktail 1:100 (MFCD00677817) (Sigma-Aldrich) were vacuum infiltrated with the individual solutions 3 times for 2 minutes each, and incubated at room temperature (RT) for an additional 10 minutes before freezing. Because the cleavage event happens so fast, preinfiltration was necessary to deliver the inhibitors as close and as quickly as possible to the presumed proteolysis site (the cytoplasm) before damage ensues. In the case of cell-impermeable inhibitors, such as EGTA and EDTA, they are at least absorbed in the intercellular space and apoplast from where they can presumably enter the cell instantly upon damage.

For the *in vivo* wounding treatment, 8 seedlings were pooled, squeezed 5 times with serrated forceps, and incubated at RT for the indicated amount of time before freezing in liquid nitrogen and subsequent immunoblot analysis. Treated and untreated seedlings were frozen in liquid nitrogen and ground to powder with a mortar and pestle under constant supply of liquid nitrogen, because the use of automated homogenizers led to thawing and immediate PEP detection in the immunoblot.

To evaluate the potential difference of PROPEP1 processing in wild type (WT) and metacaspase mutants and between leaf (whole rosette) and root tissues, seedlings were grown for 14 days at 21°C under continuous light on ½MS medium supplemented with 0.7% (w/v) agar without sucrose. Roots were separated from leaf tissue with a scalpel, thereby minimizing unwanted tissue damage prior to the forceps treatment or grinding in liquid nitrogen. For inhibitor treatments of crushed root and leaf tissues, dimethyl sulfoxide (DMSO), 5 mM EGTA, and 50 µM Z-VRPR-fmk were not preinfiltrated, but added to the tissue in deionized water prior to thawing at RT. Slightly elevated cleavage in samples treated with EGTA and Z-VRPR-fmk (fig. S3H) could be due to a slight delay in absorption of the chemicals in the crushed tissue.

Tissue powder from liquid nitrogen-ground seedlings or tissues were stored at -80°C and for any application, other than incubation of ground tissue at RT, was immediately supplemented with 3× SDS loading buffer (0.5 M Tris, pH 6.8, 15% (w/v) glycerol, 0.3 M DTT, 5% (w/v) SDS, and bromophenol blue) preheated at 70°C. Transgenic Arabidopsis lines (Columbia-0 accession) expressing PROPEP1-YFP, PEP1-YFP, and YFP had been described previously (Bartels et al., 2013).

Cloning of constructs and generation of transgenic Arabidopsis lines

Mutated PROPEP sequences were prepared by site-specific mutagenesis of the original coding sequence in the plasmid pEarley101 (Huffaker et al., 2006; Tang et al., 2015). For the mCitrine-PROPEP1-RFP construct, mCitrine was ligated at a ScaI site to the N-terminus of PROPEP1 in pDONR207 and cloned by Gateway™ into the destination vector pB7RWG2 for an in-frame fusion of RFP to the C-terminal end. For comparison of the native and R69-mutated PROPEP1 sequences, Arabidopsis (Col-0) was transformed by Agrobacterium tumefaciens by means of the the floral dip method, whereas for comparison to METACASPASE 4-deficient plants, a

homozygous mc4 (-/-) mutant in the Arabidopsis Landsberg erecta accession (Ler-0), CSHL_GT7237 and a corresponding wild-type Ler-0 line were transformed with the pEarley101-PROPEP1 construct by the floral dip method and for comparison to MC9-deficient plants, a mc9 (-/-) mutant (Col-0; GABI_540H06) was used. These lines were selected for an expected single-insertion genetic segregation of the pEarley101-PROPEP1 construct at a 3:1 ratio on selective ½MS medium containing 10 mg/L of glufosinate-ammonium (Sigma-Aldrich). Equal amounts of PROPEP1-YFP fusion proteins between the selected Ler-0 and mc4 (-/-) lines and the Col-0 and mc9 (-/-) lines were confirmed by immunoblots with anti-GFP (data not shown). To re-confirm the genetic background of the different lines, a minimum of 10 seedlings per line were genotyped with primers spanning the MC4 and MC9 genomic loci and T-DNA-specific primers (fig. S2D, table S3).

Immunoblotting and band percentage quantification

Ground tissue was immediately supplemented with approximately an equal volume of 3× SDS loading buffer preheated at 70°C, heated for 5 minutes, and centrifuged for 5 minutes at 16,000g to remove cellular debris. Proteins were separated in 10% precast SDS polyacrylamide gels (Genscript) for PROPEP1-YFP and PEP1-YFP separation, or in 4-20% precast gradient SDS polyacrylamide gels (Genscript) for MC4 subunit separation. Analysis was done by semi-dry Western blotting to PVDF membranes and incubated with anti-GFP antibodies (mouse, 1:1000; #11814460001 Roche) or anti-AtMC4 (rabbit, 1:15000). Only non-overexposed immunoblots were analyzed. Bands were quantified for PROPEP1-YFP and PEP1-YFP as a percentage of the total signal per lane either with ImageJ (Schindelin et al., 2012) (Fig. 1D and E) or with Image Lab software (Bio-Rad) (Fig. 2H-K, fig. S2G, fig. S3D and H) and described as band percentage (band %) in all graphs. Briefly, each lane on a given immunoblot was represented as a pixel-intensity profile (lane profile) and peaks corresponding to the different bands were marked. The relative area under the curve for each peak corresponded to the band %. Quantification of a given band within a lane alleviated the need for comparison to a reference band or to the loading control (for band quantification across different lanes).

In vitro TNT®-protease assay

Unmodified and mutated PROPEP-coding sequences (CDS) were cloned by the Gateway™ method to pDEST15 (N-terminal GST tag) or pDEST24 (C-terminal GST tag). Addition of a GST tag was necessary to increase the size and amount of incorporated 35S-methionine and improve visualization of the fused protein product. Recombinant MC4 (rMC4) and mutated inactive rMC4C/A (alanine substitution of active-site cysteine at position 139) fused to a His-tag were expressed and purified from Escherichia coli as previously described (16) and stored in 50% (w/v) glycerol, 25 mM HEPES (pH 7.5). TNT-protease assays were carried out as described (Minina et al., 2014). Briefly, PROPEP CDSs were *in vitro* transcribed and translated (TNT® coupled transcription/translation system, Promega) in the presence of radiolabeled 35S-methionine and aliquots of this reaction were subsequently mixed and incubated with the indicated amounts of rMC4 or the inactive rMC4C/A protease for 30 minutes at 30°C in optimal rMC4 reaction buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 10% (w/v) glycerol, 50 mM CaCl2e,and 10 mM DTT). Proteolysis was stopped by the addition of Laemmli buffer supplemented with 50 mM EGTA to avoid aberrant SDS-PAGE electrophoresis due to high levels of Ca^{2+} ions in the samples. Samples were separated by SDS-PAGE and visualized with storage phosphor screens.

Laser wounding and microscopy

Seedlings were grown upright on ½MS plates with 0.7% (w/v) and without sucrose after 48 hours of stratification and 7 to 10 days of growth under a 16-hour light/8 hour dark regime. Slides were prepared in two ways for microscopy. At first, seedlings were transferred to microscopy slides in 150 µl of 0.01 mg/ml propidium iodide (PI)-containing deionized water and, when indicated, 50 µM Z-Val-Arg-Pro-DL-Arg-fluoromethylketone trifluoroacetate (Z-VRPR-fmk; Bachem), 1 mM EGTA, or a range of concentrations of 1,2-bis(o-aminophenoxy)ethane-N,N,N′,N′-tetraacetic acid (BAPTA) dissolved in DMSO. Microscopy slides were taped at one end as a spacer to avoid squeezing and damaging of the root tip after transfer, whereafter the roots were carefully covered with standard glass coverslips, allowing a qualitative study. To overcome the problem that roots tend to move out of focus with this method and to improve quantitation, roots were mounted in custom-made sample holders as

described (Himschoot et al., 2018). Precise focal regions were wounded with a Ti:Sa laser (MaiTai DeepSee multiphoton laser; SpectraPhysics) at an excitation wavelength of 900 nm at 70% power and for variable durations of 300 to 7000 milliseconds. Confocal images were acquired on a Zeiss LSM780 confocal microscope with a Plan-Apochromat 40×/1.4 oil immersion objective, argon laser at an excitation wavelength of 514 nm; and respective emission regions for yellow fluorescent protein (YFP) and PI. Intracellular calcium concentrations were measured ratiometrically with a Yellow Cameleon 3.60 probe fused to a nuclear export signal (YC3.60-NES) as described (Behera et al., 2018). PI was excited at a wavelength of 561 nm, as to not cross-excite the YFP moiety of YC3.6-NES.

For the mCitrine-PROPEP1-RFP construct, 7- to 9-day-old seedlings were imaged on a Leica SP5-II-Matrix confocal microscope, with mCitrine and mRFP excitation at 514 nm and 561 nm and emission in the 530-560 nm and 550-750 nm range, respectively. Images were acquired with a 63× water-immersion objective, 2 Hybrid Detectors (HyD), and a bright-field scan. Wounding was achieved by slightly squeezing the roots with forceps prior to imaging (Fig. 4A). The mCitrine-PROPEP1-RFP line was also imaged after laser wounding on the Zeiss LSM780 setup (movie S7).

Microscopy image analysis for the quantification of calcium transients (NES-YC3.60) and YFP delocalization (PROPEP1-YFP) during laser wounding

Images were analyzed with custom-made scripts for ImageJ (Schindelin et al., 2012) (supplementary materials, YFP_deloc.groovy). First, the images were registered on a single channel (for example, the PI channel) by means of an elastic registration algorithm implementation available in ImageJ (bUnwarpJ), whereafter the others channels were transformed accordingly. Second, the scripts allowed the selection of multiple regions of interest (ROIs), corresponding to the various cells encompassing the wound region. Third, several intensity features (Mean, Median, and Maximum Intensities) were extracted from the different channels per ROI over time (images were taken every second). In addition, for the YC3.60-NES probe, the ratio of the Venus channel over the CFP channel (R) divided by the average intensity of 10 seconds before laser wounding (R0) was used as a proxy for free cytosolic calcium concentration (R/R0 $[Ca^{2+}]$cyt).

For the PROPEP1-YFP probe, the tonoplast-to-cytosol delocalization of the YFP signal can be viewed as relatively few pixels with a high YFP signal intensity (tonoplast) moving in time to relatively more pixels with a lower YFP signal intensity (cytosol) in a given ROI (see fig. S5B for a visual explanation). For this quantification we divided the pixels in the upper third of a pixel intensity distribution by the lower two-thirds in that distribution per ROI over time. Because not all the cells express the PROPEP1-YFP probe equally, the pixel intensity distribution was normalized beforehand to the highest value in the first image in time as follows, with H, histogram of pixel intensities in the image, B, number of Bin in the Histogram H, and b, index of the bin holding the maximum intensity value of the first time point:

H[i]=number of pixel with intensity in the i-tier bin (1≤i≤B)

$$ratio = \frac{\sum_{i=R+1}^{b} H[i]}{\sum_{i=1}^{R} H[i]}$$

$$R = \left[\frac{2}{3} \times b\right]$$

The PI channel was quantified as PI signal intensity divided by the average intensity of 10 seconds before laser wounding (I/I0).

Ultracentrifugation for the biochemical separation of soluble and insoluble (membrane) fractions

Sterile Arabidopsis seedlings expressing the PROPEP1-YFP fusion protein were grown for 14 days at 21°C in continuous light on ½MS medium supplemented with 0.7% (w/v) agar (without sucrose). Seedlings were frozen in liquid nitrogen and ground to powder with a mortar and pestle under constant supply of liquid nitrogen. Approximately 1 gram of tissue powder was divided in two 15-ml tubes. One tube was thawed and left for 5 minutes at RT, while the other was mixed immediately with ice-cold extraction buffer (10 mM HEPES, pH7.5, 10 mM KCl, 2 mM MgCl2, 10 mM EGTA, 1 mM EDTA, and protease inhibitor cocktail (cOmplete™ ULTRA Tablets EDTA-free, #05892791001 Roche)). For 500 mg of tissue powder, after addition of 5 ml of

extraction buffer, the sample was vortexed and transferred to an ultracentrifugation tube (Ultra-clear tube, 13×51 mm, #344057 Beckman) and centrifuged (SW55 Ti swinging-bucket rotor, Beckman Coulter) for 10 minutes at 20,000g (20K) at 4°C. The supernatant was transferred to a new tube and centrifuged for an additional hour at 100,000g (100K) at 4°C. The 100K pellet was solubilized with 3× SDS loading buffer. Both 20K and 100K supernatant fractions were precipitated with acetone and the protein pellets were solubilized with 3× SDS loading buffer. A control sample taken before centrifugation containing total tissue powder was extracted in 3× SDS loading buffer. Samples were separated on NuPAGE™ 4-12% Bis-Tris Protein Gels, 1.0 mm, 10-well (Thermo Fisher Scientific) and transferred to PVDF membranes by semi-dry Western blotting (Trans-Blot Turbo Transfer System and transfer packs, #1704157, Bio-Rad). The following antibodies were used according to the manufacturers' instructions: anti-GFP (mouse, 1:1000 dilution; #11814460001 Roche), anti-actin (Rabbit, 1:1000 dilution; #AS132640 Agrisera), anti-TIP1-1 (Rabbit, 1:1000 dilution; #AS09482, Agrisera) and their corresponding secondary antibodies, anti-mouse IgG (1:10000 dilution, #NA931V, GE healthcare) and anti-rabbit IgG (1:10000 dilution, #NA934V, GE healthcare).

Immunoprecipitation and in-gel digest of PROPEP1 fusion protein bands to determine peptide coverage by mass spectrometry

Total seedling tissue powder was stored at -70°C. Immunoprecipitation was carried out according to the manufacturer's protocol (GFP-trap_MA, Chromotek) with some adjustments. Approximately 1 g of seedling powder was homogenized in 1.5 ml of extraction buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1% NP-40, and Complete protease inhibitor cocktail (Roche)). The tissue lysate was centrifuged for 10 minutes at 20,000g and 120 µl of GFP-trap_MA beads slurry was added to the supernatant and incubated for 2 hours on a rotating wheel at 4°C. The beads were washed 3 times with 1 ml of wash buffer (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, and 0.5% NP-40) and the protein was eluted with 30 µl of 2× Laemmli sample buffer. The sample was loaded on 1.0-mm, 10-well, NuPAGE™ 4-12% Bis-Tris protein gels (Thermo Fisher Scientific) and protein bands were visualized with a mass spectrometry-compatible Pierce™ Silver Stain Kit (Thermo Fisher Scientific). Bands corresponding to the relative sizes of PROPEP1-YFP and lower molecular mass-processed forms, were cut

out of the gel and prepared by a standard in-gel digestion protocol for mass spectrometry analysis. Proteins were digested overnight in the gel bands soaked with either trypsin (at 37°C) or chymotrypsin (at 25°C) in 50 mM ammonium bicarbonate, 10% (v/v) acetonitrile (ACN). The following morning, samples were acidified with 0.5% trifluoroacetic acid (TFA) final volume and dried in a Speed-vac.

The immunoprecipitation experiment was done twice and samples were analyzed by liquid chrromatrography-tandem mass spectrometry on both a Q-Exactive HF and Orbitrap Elite mass spectrometer (Thermo Fisher Scientific). Generated tandem mass spectra from raw data files were extracted in a MGF file format with RawConverter. MGF files were searched against a concatenated target-decoy database of the representative Araport11 proteome supplemented with the PROPEP1-YFP protein sequence with the SearchGUI toolkit (version 3.3.3). SearchGUI was configured to run the tandem mass spectra with the identification search engines X!Tandem, MS-GF+, MyriMatch, Comet, and OMSSA. Non-default tandem mass spectrum identification settings were semi-specific digestion for either trypsin or chymotrypsin (both without P rule) and variable modifications included protein N-terminal acetylation, methionine oxidation, N-terminal pyroglutamate, and cysteine propionamidation. Fragment ion tolerance were set to 0.5 Da and 0.01 Da for the Orbitrap Velos and Q-Exactive HF spectrometers, respectively. There were no fixed modifications, because no cysteine alkylation was performed. Search identification output files were processed by the companion tool PeptideShaker (version 1.16.25) and all default reports (.txt) and identification files (.mzid) were exported. Peptides corresponding to PROPEP1-YFP have been summarized (table S2). The full mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD010816.

Endogenously generated peptides extraction and mass spectrometry analysis (peptidomics)

Sterile Arabidopsis seedlings expressing the PROPEP1-YFP fusion protein were grown on vertical ½MS plates without sucrose in two densely seeded rows (two plates) for 12 days with a 16-hour light/8-hour dark regime at 21°C. Seeds had been stratified for 2 days at 4°C. Roots were separated from the rosette with scissors and the tissue

was immediately ground in liquid nitrogen. Root tissue was unfrozen for 5 minutes at RT and proteins were extracted by sonication in 400 µl of urea lysis buffer (20 mM HEPES, pH 8.0, 8 M urea, 1 mM sodium orthovanadate, 2.5 mM sodium pyrophosphate, 1 mM β-glycerophosphate, 10 mM EGTA, 5 mM EDTA, and cOmplete ULTRA protease inhibitor cocktail tablets (Roche)). The lysates were cleared by centrifugation at 20,000g at 15ºC for 15 minutes and transferred to new tubes. Samples were acidified with 0.5% TFA and again cleared by centrifugation. Native peptides were enriched by reversed-phase chromatography on a SampliQ C18 column (Agilent Technologies). Columns were pre-wetted with 70% ACN and 0.1% TFA and elution was done with 50% ACN and 0.1% TFA whereafter the eluate was dried by Speed-vac. No trypsin or any other protease was used to preserve the structure of the endogenously generated peptides.

Samples were dissolved in 2% ACN and 0.1% TFA and subjected to mass spectrometry on a Q Exactive Orbitrap mass spectrometer operated as previously described (Stes et al., 2014). From the tandem mass spectrometry data, Mascot Generic Files (mgf) were created with the Mascot Distiller software (version 2.5.1.0, Matrix Science). Peak lists were then searched with the Mascot search engine and the Mascot Daemon interface (version 2.5.1, Matrix Science). Spectra were searched against the TAIR10 database concatenated with the PROPEP1-YFP protein sequence. Variable modifications were set to pyroglutamate formation of the amino-terminal glutamine and methionine oxidation. Mass tolerance on precursor ions was set to ±10 ppm (with Mascot's C13 option set to 1) and on fragment ions set to 20 mmu. The instrument setting was on ESI-quadruple (QUAD). As similarly to trypsin, metacaspase cleaves C-terminally to arginine and lysine, the enzyme was set to trypsin/P, allowing eight missed cleavages, and the cleavage was also allowed when lysine or arginine were followed by proline. Only peptides that were ranked first and scored above the threshold score, set at 99% confidence, were withheld. Peptides matching the above criteria have been summarized (table S1). The full mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD005740.

GST-TEV-Δ39-PP1, GST, rMC4 and rMC4C/A protein purification

A N-terminally truncated version of PROPEP1 lacking the first 39 amino acids (starting at a serine at position 40) was subcloned from full-length PROPEP1 and an upstream cleavable tag for the Tobacco Etch Virus protease (TEVp) was added at the N-terminal part by means of iProof polymerase (Bio-Rad) with the Gateway-compatible primers attB1-TEVrs-S40PROPEP1-FW and attb2-PROPEP1_s-RV (containing a stop codon) (see primer list in table S3). Truncation was necessary to improve protein purification from E. coli, because a full-length GST-TEV-PROPEP1 fusion protein was difficult to obtain in soluble form (data not shown). The PCR fragment was assembled into pDONR221 by BP reaction and correct insertion was confirmed by Sanger sequencing. The resulting entry clone was N-terminally fused to GST by LR reaction in pDEST15 and named GST-TEV-Δ39-PP1. The construct was transfected into BL21 (DE3) E. coli cells and protein expression was induced at OD600 = 0.8 at 28°C overnight with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). E. coli cells were pelleted and sonicated in GST extraction buffer (50 mM HEPESs-KOH, pH 7.5, and 300 mM NaCl) on ice. Proteins were purified by gravity flow with glutathione Sepharose resin (#17075601) and PD10 columns according to the manufacturer's protocol at 4℃ (GE Healthcare). Columns were washed with GST extraction buffer and the protein was eluted with 50 mM HEPES-KOH, pH 7.5, 300 mM NaCl, and 10 mM reduced glutathione (#78259, Thermo Fisher Scientific). Free GST protein (as control for the root growth experiment) was purified under the same conditions. Purified GST and GST-TEV-Δ39-PP1 proteins were stored at -70°C in elution buffer supplemented with 10% (v/v) glycerol.

E. coli cells for the purification of TEVp (BL21 DE3 pRK793 plasmid), rMC4, and rMC4C/A (BL21 DE3 (pLysE) pDEST17-MC4 and pDEST17-MC4C/A plasmid, respectively) were grown to OD600 = 0.6 and protein expression was induced overnight with 0.2 mM IPTG at 20°C. Cultures were pelleted and sonicated in HIS extraction buffer (50 mM HEPES-KOH, pH 7.5, 300 mM NaCl, and 20 mM imidazole). Protein was bound to Nickel Sepharose 6 Fast Flow agarose beads (#17531801, GE Healthcare) by gravity flow at 4°C, washed with 50 mM HEPES-KOH, pH 7.5, 300 mM NaCl, and 40 mM imidazole, and eluted with 50 mM HEPES-KOH, pH 7.5, 300 mM NaCl, and 400 mM imidazole. Proteins were subjected to size exclusion chromatography and eluted in 50 mM HEPES-KOH, pH 7.5, 300 mM NaCl, and 10% (v/v) glycerol.

Root growth inhibition analysis

Protein concentrations of GST-TEV-Δ39-PP1, GST, TEVp, rMC4, and rMC4C/A were determined by A280 on a NanoDrop spectrophotometer. The GST-TEV-Δ39-PP1 protein was incubated with TEVp at a concentration ratio of 1:100 overnight at 30°C and subsequently with rMC4 or rMC4C/A at a concentration ratio of 1:50 for 1 hour at 30°C. As a control treatment, GST was similarly incubated with TEVp and rMC4C/A. The reaction buffer consisted of 50 mM HEPES-KOH, pH7.5, 120 mM NaCl, 10% (v/v) glycerol, 10 mM DTT, and 50 mM CaCl2.

Seeds of Arabidopsis pepr1pepr2, mc4 (-/-), and their respective WT background accessions (Col-0 and Ler-0) were stratified at 4°C for 3 days, grown for another 3 days on vertical plates containing ½MS media with 0.7% (w/v) agar (without sucrose), and then transferred to plates containing 25 nM of each treatment mixed in the solid ½MS media. Treatments were done on the same plate per biological repeat and were changed from top to bottom on the 4-well plates (Nunc™ Nunclon 4-Well, #167063, Thermo Fisher Scientific) to avoid a positional effect on the root growth between treatments over the biological repeats. Root pictures were taken 5 days after transfer to the treatment and root length (mm per pixel) was measured with ImageJ (43).

R visualization

Peptide coverage plots for the PROPEP1-YFP fusion protein and fragments were generated by plotting the number of validated peptide spectrum matches (PSMs) per protein position (fig. S2F). From the exported default peptides reported, peptides with a confidence > 0.9 were used to cumulate the number of validated PSMs for each position of the peptide. Per sample, the number of validated PSMs was displayed with a heat-color scale ranging from 2 PSMs (grey) to the maximum of validated PSMs of that sample (red).

Statistics

A one-way analysis of variance (ANOVA) was applied to the log-transformed PEP1 and PROPEP1 intensities with inhibitor as main effect for the data presented (Fig. 1E,

fig. S3H). This transformation was necessary to stabilize the variance. The interest was in the difference in outcome of each inhibitor compared to the water treatment. P value adjustment was done with the Dunnett's method. R software (www.cran.org, R version 3.5.1) was applied for the analysis with the lm function. Post-hoc analysis was done with the emmeans package and adjusted P values were displayed on the bar charts in case of significance.

A two-way ANOVA was applied to the log-transformed PEP1, PROPEP1, p20, and p20* intensities with genotype and time as fixed effects and their interaction term for the data presented (Fig. 2H-K, fig. S3D). The log transformation was necessary to stabilize the variance. The interest was in the difference in outcome over time between both genotypes. R software (www.cran.org, R version 3.5.1) was applied for the analysis with the lm function. Post-linear analysis was done by means of the capabilities of the emmeans function and adjusted P values were displayed on the bar charts in case of significance.

A two-way ANOVA was applied to the root length data with genotype and treatment as fixed effects and their interaction term for the data presented (Fig. 4C). Tukey’s multiple comparison test was used to query significant differences between treatment in a given genotype. GraphPad Prism version 7.03 was applied for the analysis.

Acknowledgements

We thank Eric Lam (Rutgers University, New Jersey, USA), Eugenia Russinova (VIB-UGent Center for Plant Systems Biology, Ghent, Belgium) and Alex Costa (University of Milan, Milan, Italy) for kindly providing the anti-MC4 antibody, the mutant pepr1pepr2 seeds, and the YC3.60-NES seeds, respectively, Veronique Storme (VIB-UGent Center for Plant Systems Biology, Ghent, Belgium) for advice on statistics, the VIB BioImaging Core for excellent technical support, and Martine De Cock for excellent help in preparing the manuscript. Funding: This work was supported by the Swiss National Science Foundation (Grant 31003A_127563 to T.B.), the Research Foundation-Flanders (grant G.0C37.14N to K.G. and FWO14/PDO/166 to S.S.), the Ghent University Special Research Fund (grant 01J11311 to F.V.B.), and a CLEM grant from Minister Ingrid Lieten (Belgium) for the acquisition of the Zeiss LSM780 microscope.

Position		P4-P1	Peptide	Workflow	Publication
181		KFLR	Nt-SKVEGAIESR	COFRADIC, TAILS	(Zhang et al., 2018), (Willems et al., 2017)
189		GAIE	Nt-SRGFHIGGNKKDE	COFRADIC	(Willems et al., 2017)
191		IESR	Nt-GFHIGGNKKDEDEAEEIETK	COFRADIC	(Tsiatsiani et al., 2013)
			Nt-GFHIGGNKK		(Willems et al., 2017)
			Nt-GFHIGGNKKDE		
211		IETK	Nt-EIELEDGETIH	TAILS	(Kohler et al., 2015)
			EIELEDGETIHAK	Peptidomics	This study (table S1)
			EIELEDGETIHAKDK		
215		EIEL	Nt-EDGETIHAK	COFRADIC	(Willems et al., 2017)
			Nt-EDGETIHAKDK		
226		AKDK	Nt-SLPLQTLIDILKQQTGNDNIE	ChaFRADIC	(Venne et al., 2015)

Table 1 N-terminal peptides in between the active site cysteine (position 139) and initial autocatalytic cleavage site K225 of MC4.

PTM-viewer (Willems et al., 2019) and public available proteomics data were queried for MC4 N-terminal peptides upstream of K225 (peptide in position 226 was also identified). P4-P1 = four amino acids preceding the putative cleavage site. Peptide = N-terminal labeled peptides (Nt) or endogenous peptides. Workflow = protocol used to isolate the peptides: COmbined FRactional Diagonal Chromatography (COFRADIC), Terminal Amine Isotopic Labeling of Substrates (TAILS), Charge-based FRactional DIagonal Chromatography (ChaFRADIC).

SUPPLEMENTARY MATERIALS

Supplementary material can be found online at https://osf.io/6v2p5/.This data includes:

Figs. S1 to S6

Table S3

Captions for Movies S1 to S7

Other Suplementary Materials for this includes:

Tables S1 and S2

Movies S1 to S7

Microscopy data analysis script

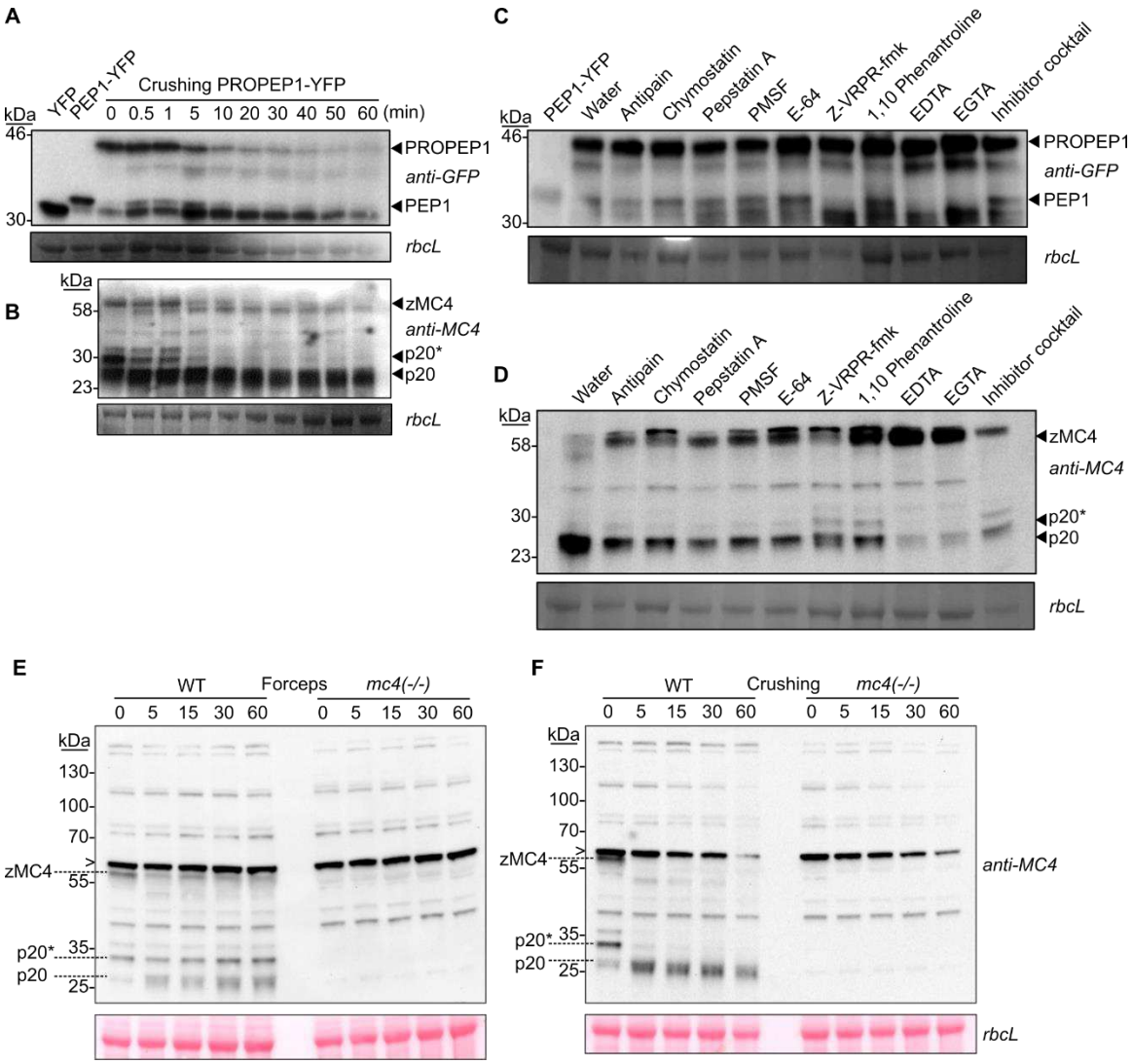


Fig. S1. Processing of PROPEP1-YFP correlates with AtMC4 processing and is inhibited by metacaspase inhibitors. (A) Anti-GFP immunoblot of PROPEP1-YFP processing in crushed tissue powder. Extracts from PEP1-YFP and YFP expressing seedlings were loaded as size controls. (B) Anti-MC4 immunoblot displaying processing of zymogenic MC4 (zMC4) into p20* and p20 subunits in crushed tissue powder. (C) Anti-GFP immunoblot showing extracts from PROPEP1-YFP expressing seedlings that were vacuum infiltrated with antipain (100µM), chymostatin (100µM), pepstatin A (1µM), PMSF (1mM), E-64 (10µM), 1,10-phenanthroline (20mM), Z-VRPR-fmk (50µM), EDTA (1mM), EGTA (1mM), and Protease inhibitor cocktail (Sigma Aldrich). Crushed tissue powder was incubated for 5 minutes at room temperature prior to protein extraction. (D) Anti-MC4 immunoblot displaying the effect of the protease inhibitors on MC4 processing in protein extracts under the same

conditions as (C). (E-F) Anti-MC4 immunoblot on whole seedlings wounded by forceps (E) or crushed tissue powder (F) incubated at RT for the indicated time in wild type and MC4 deficient background (mc4(-/-)). The band indicated with (>) is present, amongst others, in mc4(-/-) and almost overlaps with zMC4. Ponceau S stain of the rubisco large subunit (rbcL) indicates protein load.

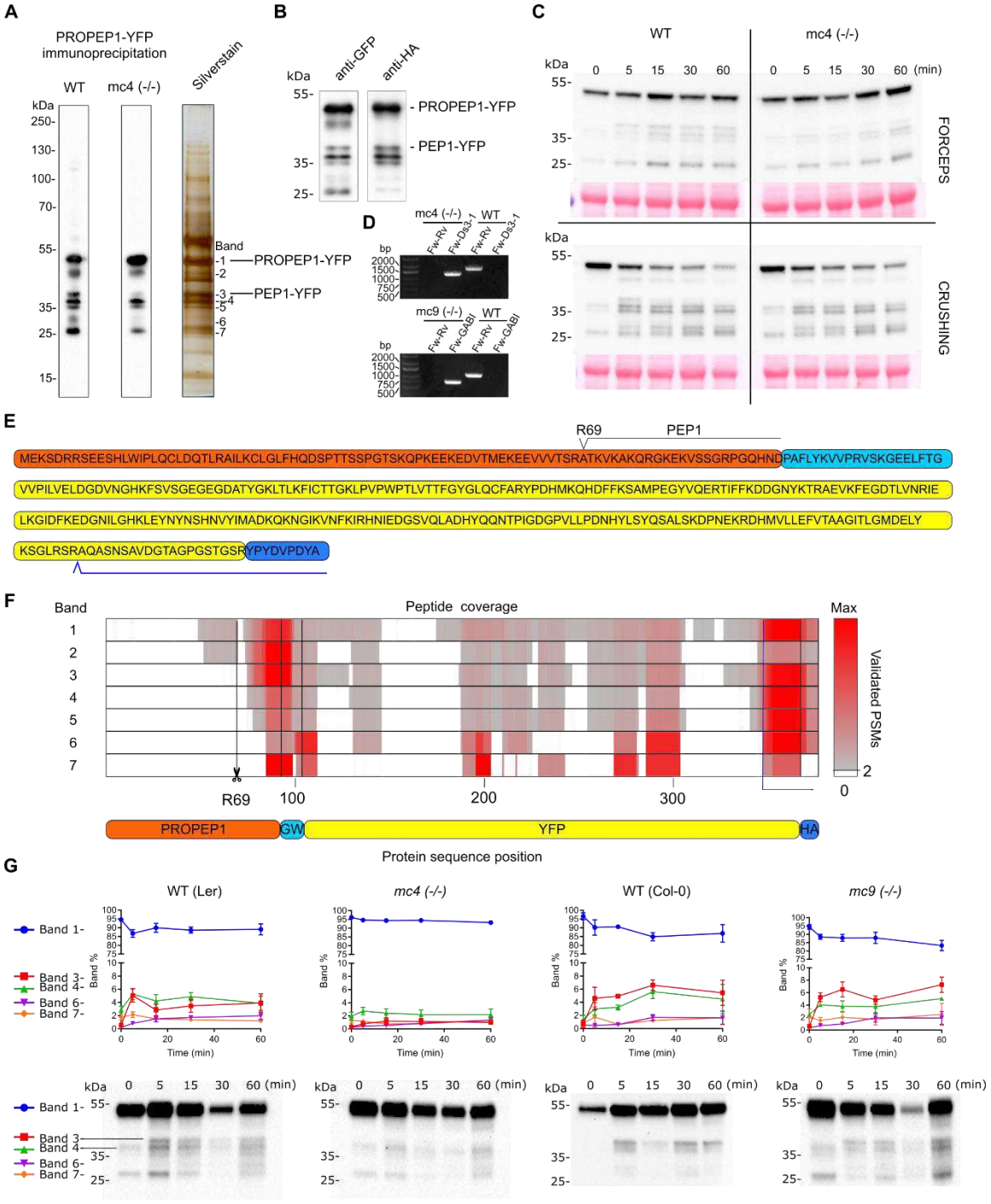


Fig. S2. PROPEP1-YFP immunoprecipitation reveals the identity of processed bands and cleavage is unaffected in mc9(-/-) mutant plants. (A) Anti-GFP immunoblot of immunoprecipitated PROPEP1-YFP and processed bands in wild type (WT) and mc4(-/-) from crushed tissue powder incubated for 5 minutes at room temperature. Bands indicated by their respective number were prepared for protein identification by mass spectrometry. (B) Comparison of anti-GFP and anti-HA immunoblot of PROPEP1-YFP bands (crushed seedling, 5 minutes at room temperature) revealed which bands have lost the C-terminus of the PROPEP1-YFP fusion protein. (C) Comparison of PROPEP1-YFP band pattern between wound model experiments in WT and mc4(-/-) seedlings. Note that the immunoblots of the WT samples were used in Fig. 1 and repeated here to clearly compare to mc4(-/-) samples. RbcL indicates protein load. (D) Genotyping of metacaspase deficient mutants. Comparison of PCR products with forward and reverse primers (Fw&Rv) and T-DNA specific primers (Ds3-1 for mc4 and GABI for mc9). (E) Amino acid sequence of the PROPEP1-YFP fusion protein. Orange = PROPEP1; light blue = gateway cloning site (GW); yellow = YFP; dark blue = human influenza hemagglutinin (HA) tag. An endogenous peptide from PROPEP1-YFP identified by peptidomics (table S1) is underlined in blue. (F) Peptide coverage plot of the bands indicated in (A). Experiment was repeated twice and peptides with two or more peptide-spectrum matches (PSMs) are displayed. Coloring from grey (2 PSMs) to red (relative to the maximum amount of PSMs per peptide in a given band) indicates peptide abundance. The position of the cleavage site R69 and the transitions between the parts of the PROPEP1-YFP fusion protein are indicated by black vertical lines. Position from the endogenous peptide is indicated by a blue vertical line. (G) PROPEP1-YFP cleavage band analysis in mc4(-/-) and mc9(-/-) seedlings and their respective WT backgrounds, Ler and Col-0, damaged by forceps for the indicated time points. Band numbers are indicated on example immunoblots under the graphs. Data points (band % of total per lane) are the mean of a minimum of three experiments (+/- SEM). Variations in protein load and immunoblot signal intensity are caused by the need for rapid sample handling, but do not affect band quantitation (see Material and methods for further explanation).

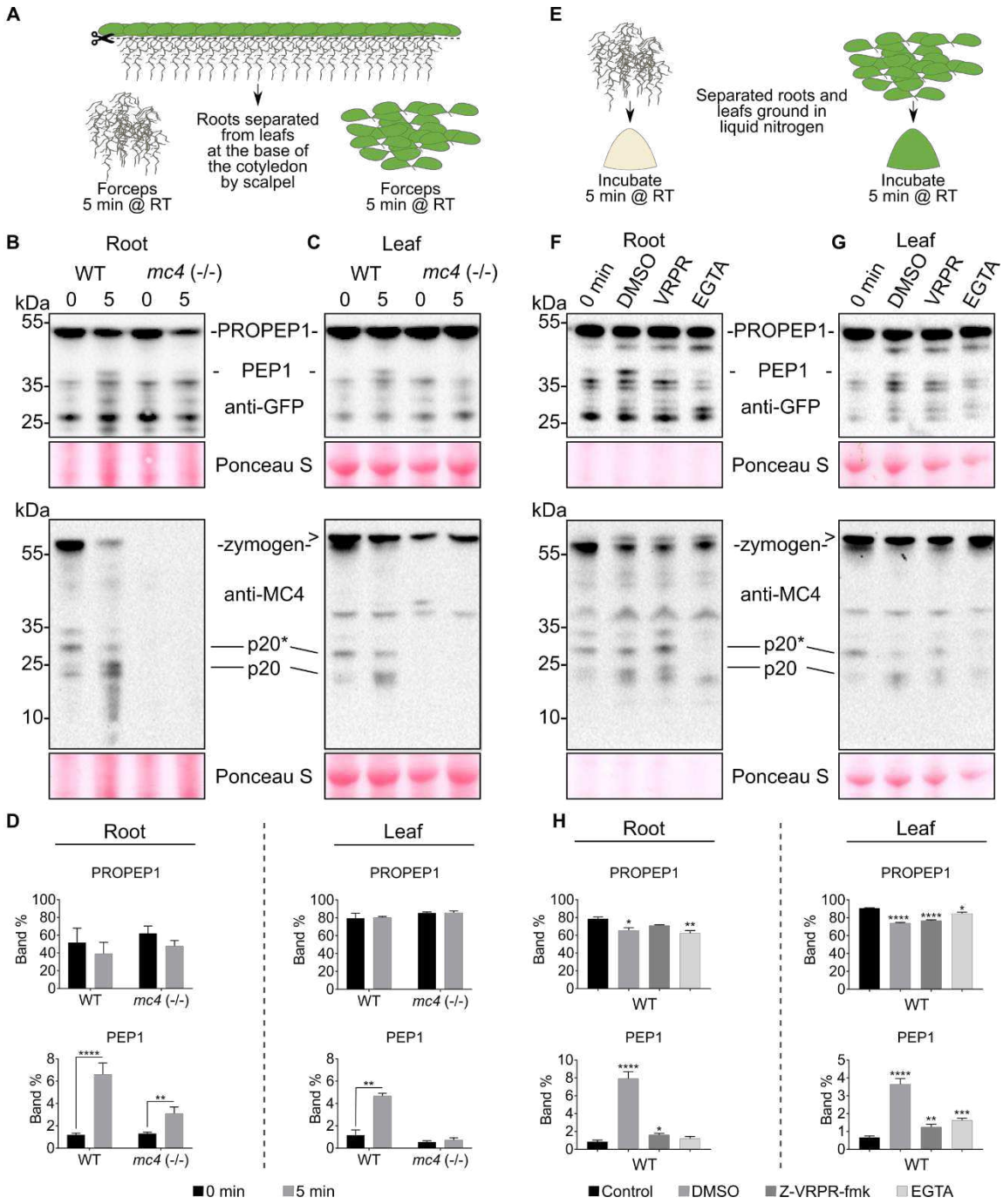


Fig. S3. MC4 cleaves PROPEP1 in root and leaf tissue *in vivo*. (A) Visual representation of the experiment performed in (B-D). (B-C) Immunoblot of PROPEP1-YFP (top) and MC4 (bottom) in WT and mc4(-/-) before and 5 minutes after damaging root (B) and leaf tissue (C) with forceps. Note that the anti-MC4 antibody had less background (band indicated with > in leaf tissue samples) in root tissue. Ponceau S stain indicates protein load. (D) Quantification of PROPEP1 and PEP1 band intensity in (B) and (C). Data are the mean of four experiments + SEM. (E) Visual representation of the experiment in (F-H). (F-G) Immunoblot of PROPEP1-YFP (top) and MC4 (bottom) in root and leaf tissue in combination with the following treatments: untreated

before thawing (0 min) and incubation at room temperature for 5 minutes in deionized water containing 0.5 % DMSO, 50 µM Z-VRPR-fmk or 5 mM EGTA. (H) Quantification of PROPEP1 and PEP1 band intensity in (F) and (G). Data are the mean of four experiments + SEM.

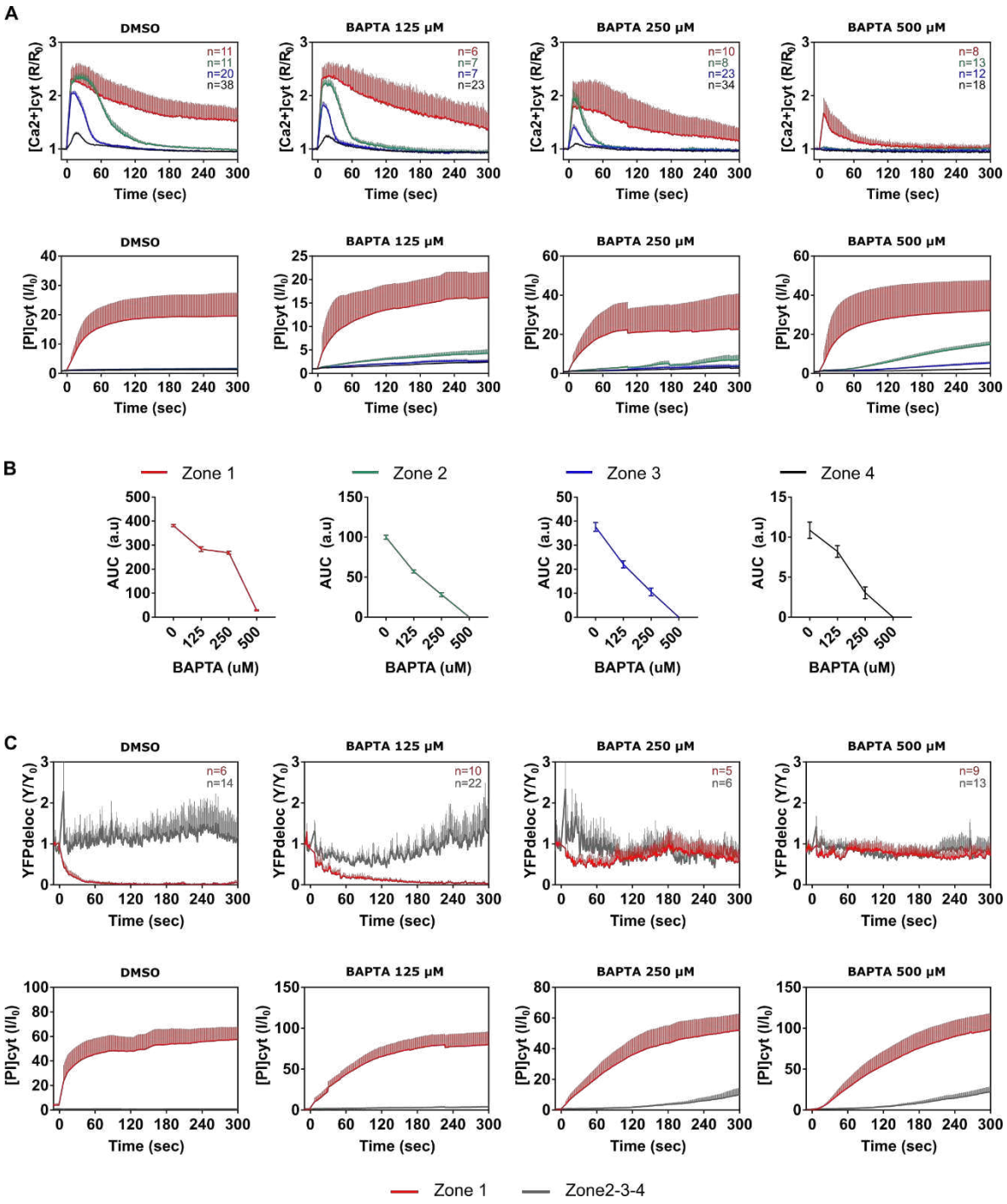


Fig. S4. Inhibition of laser-damage induced Ca^{2+} flux and PROPEP1-YFP delocalization by increasing concentrations of BAPTA. (A) $[Ca^{2+}]$cyt measured by

YC3.60-NES Ca^{2+} probe (top) or PI entry in epidermal cells of the root transition zone. Increasing concentrations of the calcium chelator BAPTA was added 5 minutes before imaging. Four percent of dimethyl sulfoxide (DMSO), equivalent to the amount in the 500 µM BAPTA samples, was used as a control. Curves are colored according the zones as shown in (B). (B) Area under the curve (AUC in arbitrary units) for the zones in (A) at a given concentration of BAPTA (0 = DMSO control). (C) Quantification of PROPEP1-YFP delocalization of PROPEP1-YFP fusion protein during laser damage (top) and corresponding PI entry (bottom) for zone 1 and the combination of zones 2,3 and 4. Curves are the mean of N cells + SD in (A) and +/- SEM in (B) and (C).

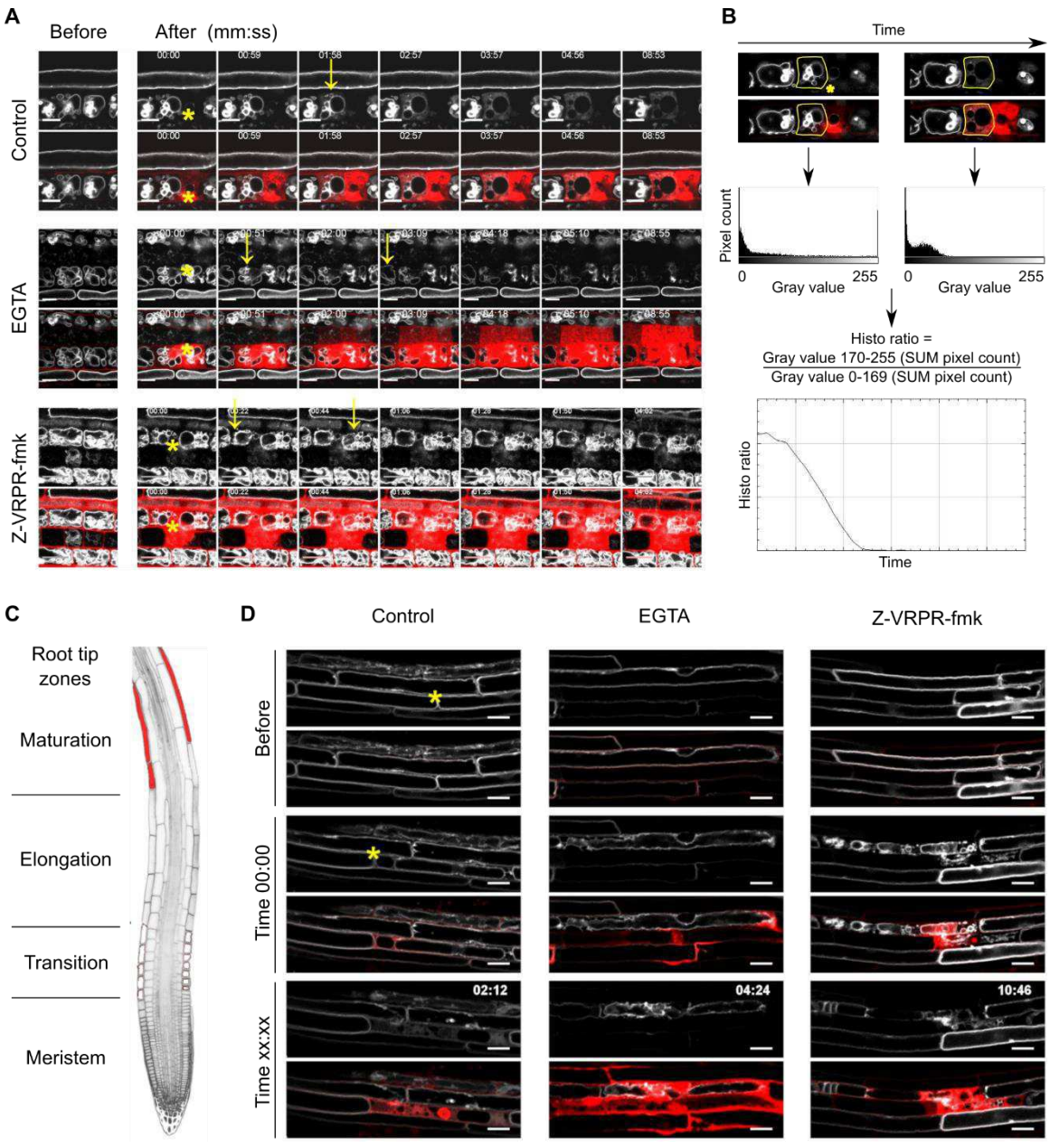


Fig. S5. EGTA and Z-VRPR-fmk inhibit PROPEP1-YFP delocalization from the tonoplast to the cytosol after laser wounding of epidermal cells of the root transition and maturation zone. (A) Laser wounding of root transition zone epidermal cells in: control condition (10 mM MES buffer pH 5.7), 1 mM EGTA (in 10 mM MES buffer pH 5.7) and 50μM Z-VRPR-fmk (in 10 mM MES buffer pH 5.7). In control samples, the sharp PROPEP1-YFP signal delocalized from tonoplast to the diffuse signal in the cytoplasm (indicated with an arrow) in the cell adjacent to the wounded cell (indicated with an asterix). Delocalization in the wounded cell and the adjacent cell did not take place for EGTA and Z-VRPR-fmk, indicated with arrows in timeframe 00:51 and 00:22, respectively. The surrounding cells started accumulating PI slightly later (indicated with arrowheads at the later time points). Also in these cells the transition was absent. (B) Visual explanation of YFP delocalization quantification as explained in the Material and methods part. The images in (B) were derived from the control experiment in (A) and only serve as example. (C) A scheme indicating in red the localization of the epidermal cells of the maturation zone. (D) Laser wounding of root maturation zone epidermal cells in the same conditions as (A). The targeted cell is indicated (*) on the image before ablation (time point 0, T0). In the subsequent control images at time points 00:00 min and xx:xx min, the center of imaging was moved to the right for better focus on the targeted cell. At each time point, the upper image shows the PROPEP1-YFP signal and the lower shows the overlap with PI. PROPEP1-YFP delocalization was inhibited in the presence of EGTA and Z-VRPR-fmk, indicated by an incomplete overlap of YFP and PI signals. Experiment was repeated 13 times for control, 7 times for EGTA and 9 times for Z-VRPR-fmk, all in two different PROPEP1-YFP plant lines with similar outcome. Scale bar is 20 μm.

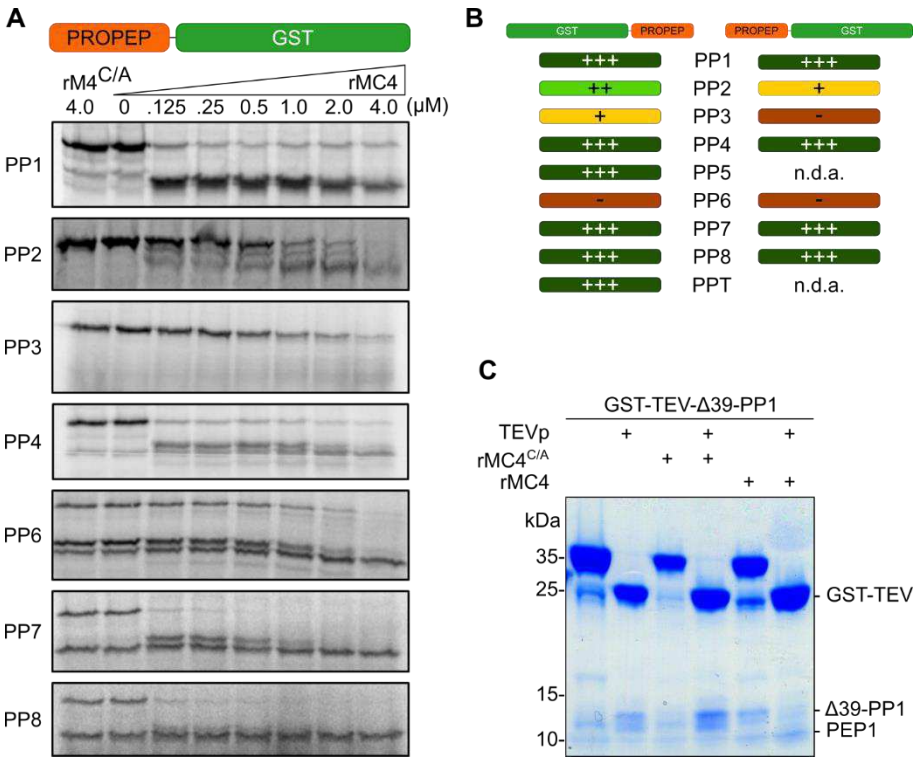


Fig. S6. MC4 cleaves certain Arabidopsis and tomato PROPEPs *in vitro* (A) TNT®-protease assays of PROPEP-GST fusion proteins of the Arabidopsis PROPEPs (PP1-PP8) and the tomato orthologue of PROPEP1 (PPT). Because size-jump of the processed bands is rather small for the GST-PROPEP fusion proteins (Fig. 4D), all assays were repeated with PROPEP-GST, except for PP5 and PPT, for which no data could be obtained (n.d.a.). All assays were repeated at least twice and were run together with PP1 as a positive control. (B) Gradations of processing were observed for the various constructs; cleaving well (+++) to bad (-). (C) Coomassie stained SDS-PAGE gel of purified GST-TEV-Δ39-PP1 fusion protein treated as indicated with Tobacco Etch Virus protease (TEVp; 1/100 ratio), rMC4 or rMC4C/A (1/50 ratio). GST-TEV-Δ39-PP1 cleavage fragments are indicated on the right. Δ39-PP1 cleavage fragments were running in the front of the SDS-PAGE gel, so PEP1 (2.5 kDa) cannot be discerned from the rest of Δ39-PP1.

Table S3.

Primer list

Primer name	Sequence (5'to 3')	Description
MC4_fw	ATGACGAAAAAGGCGGTGCTTA	Forward primer for AtMC4 (At1g79340)
MC4_rv	TCAACAGATGAAAGGAGCGTTGGC	Reverse primer for AtMC4 (At1g79340)
Ds3-1	ACCCGACCGGATCGTATCGGT	T-DNA primer for CSHL_GT7237 (*mc4*)
MC9_fw	ATGGATCAACAAGGGATGGT	Forward primer for AtMC9 (At5g04200)
MC9_rv	TCAAGGTTGAGAAAGGAACG	Reverse primer for AtMC9 (At5g04200)
GABI_LB	CCCATTTGGACGTGAATGTAGACAC	T-DNA Left Border primer for GABI_540H06 (*mc9*)
attB1-TEVrs-S40PROPEP1-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCGAAAACCTGTATTTTCAGTCCTCTCCCGGAACTTC	Cloning of GST-TEV-Δ39-PP1
attb2-PROPEP1_s-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTCTAATTATGTTGGCCAGGAC	Cloning of GST-TEV-Δ39-PP1

Movie S1

Spatiotemporal [Ca^{2+}]cyt dynamics in and surrounding laser-damaged root transition zone epidermal cells. A time series of confocal microscopy images, starting 10 seconds before laser wounding and taken every second up to 5 minutes after laser wounding. The top panel is YC3.60-NES signal (overlap of CFP and YFP FRET signal; blue means low calcium concentration, yellow means high calcium concentration). The middle panel is propidium iodide (PI) stain. Damaged cells accumulate PI. The bottom panel is the merged image of the above panels. The root was imaged in pure (deionized) water.

Movie S2

EGTA inhibits the spatiotemporal [Ca^{2+}]cyt dynamics in and surrounding laser-damaged root transition zone epidermal cells. Similar experiment as in Movie S1, but the root was imaged in pure (deionized) water containing 1 mM EGTA.

Movie S3

PROPEP1-YFP signal delocalization in laser-damaged root transition zone epidermal cells. A time series of confocal microscopy images, starting 10 seconds before laser wounding and taken every second up to 5 minutes after laser wounding. The top panel is PROPEP1-YFP signal. The middle panel is propidium iodide (PI) stain. Damaged cells accumulate PI. The bottom panel is the merged image of the above panels. The root was imaged in pure (deionized) water.

Movie S4

EGTA inhibits PROPEP1-YFP signal delocalization in laser-damaged root transition zone epidermal cells. Similar experiment as in Movie S3, but the root was imaged in pure (deionized) water containing 1 mM EGTA.

Movie S5

Z-VRPR-fmk inhibits PROPEP1-YFP signal delocalization in laser-damaged root transition zone epidermal cells. Similar experiment as in Movie S3, but the root was imaged in pure (deionized) water containing 50 μM Z-VRPR-fmk.

Movie S6

Extended time series of a laser wounding experiment for PROPEP1-YFP delocalization. The wound site was continuously imaged for up to 45 min after laser wounding. Laser target is indicated with ‘wound’. The YFP signal (white; left channel) delocalized immediately from the tonoplast (PROPEP1-YFP) to the cytosol (PEP1-YFP) at time 00:00 in the three closest cells surrounding the wound that accumulated PI (red signal overlaid with YFP in the right channel). In the next time points (01:18 to 08:14 min), YFP signal in the cell on the right of the wounded cells delocalized; however, this exceptionally occurred before PI entry. In the later time points (31:12 to 35:58 min), YFP signal in the cell on the left (indicated with ‘response’) delocalized, concomitant with PI entry.

Movie S7

Only the RFP moiety of a double labelled fluorophore construct mCitrine-PROPEP1-RFP is released to the cytosol in laser-damaged root transition zone epidermal cells. A time series of confocal microscopy images, starting 10 seconds before laser wounding and taken every two seconds for up to 5 minutes after laser wounding. The top panel is mCitrine. The middle panel is RFP and the bottom panel is the merged image of the two. The laser target site is indicated with an asterisk (*). The three adjacent cells were damaged. In surrounding undamaged cells, both mCitrine and RFP of the mCitrine-PROPEP1-RFP construct localized to the tonoplast. In the damaged cells only RFP localized to the cytosol (signal becomes more diffuse; indicated in one of the three damaged cells with a > sign), whereas mCitrine remained attached to the tonoplast.

REFERENCES

Acosta-Maspons, A., Sepulveda-Garcia, E., Sanchez-Baldoquin, L., Marrero-Gutierrez, J., Pons, T., Rocha-Sosa, M., and Gonzalez, L. (2014). Two aspartate residues at the putative p10 subunit of a type II metacaspase from Nicotiana tabacum L. may contribute to the substrate-binding pocket. Planta **239:** 147-160.

Bartels, S., and Boller, T. (2015). Quo vadis, Pep? Plant elicitor peptides at the crossroads of immunity, stress, and development. J Exp Bot **66:** 5183-5193.

Bartels, S., Lori, M., Mbengue, M., van Verk, M., Klauser, D., Hander, T., Boni, R., Robatzek, S., and Boller, T. (2013). The family of Peps and their precursors in Arabidopsis: differential expression and localization but similar induction of pattern-triggered immune responses. J Exp Bot **64:** 5309-5321.

Behera, S., Zhaolong, X., Luoni, L., Bonza, M.C., Doccula, F.G., De Michelis, M.I., Morris, R.J., Schwarzlander, M., and Costa, A. (2018). Cellular Ca(2+) Signals Generate Defined pH Signatures in Plants. Plant Cell **30:** 2704-2719.

Beloshistov, R.E., et al. (2018). Phytaspase-mediated precursor processing and maturation of the wound hormone systemin. New Phytol **218:** 1167-1178.

Coll, N.S., Vercammen, D., Smidler, A., Clover, C., Van Breusegem, F., Dangl, J.L., and Epple, P. (2010). Arabidopsis type I metacaspases control cell death. Science **330:** 1393-1397.

Costa, A., Luoni, L., Marrano, C.A., Hashimoto, K., Koster, P., Giacometti, S., De Michelis, M.I., Kudla, J., and Bonza, M.C. (2017). Ca2+-dependent phosphoregulation of the plasma membrane Ca2+-ATPase ACA8 modulates stimulus-induced calcium signatures. J Exp Bot **68:** 3215-3230.

Deutsch, C.A., Tewksbury, J.J., Tigchelaar, M., Battisti, D.S., Merrill, S.C., Huey, R.B., and Naylor, R.L. (2018). Increase in crop losses to insect pests in a warming climate. Science **361:** 916-919.

Fortin, J., and Lam, E. (2018). Domain swap between two type-II metacaspases defines key elements for their biochemical properties. Plant J **96:** 921-936.

Ghorbani, S., Hoogewijs, K., Pecenkova, T., Fernandez, A., Inze, A., Eeckhout, D., Kawa, D., De Jaeger, G., Beeckman, T., Madder, A., Van Breusegem, F., and Hilson, P. (2016). The SBT6.1 subtilase processes the GOLVEN1 peptide controlling cell elongation. J Exp Bot **67:** 4877-4887.

Heil, M., and Land, W.G. (2014). Danger signals - damaged-self recognition across the tree of life. Front Plant Sci **5:** 578.

Himschoot, E., Krebs, M., Costa, A., Beeckman, T., and Vanneste, S. (2018). Calcium Ion Dynamics in Roots: Imaging and Analysis. Methods Mol Biol **1761:** 115-130.

Huffaker, A., Pearce, G., and Ryan, C.A. (2006). An endogenous peptide signal in Arabidopsis activates components of the innate immune response. Proc Natl Acad Sci U S A **103:** 10098-10103.

Huffaker, A., Pearce, G., Veyrat, N., Erb, M., Turlings, T.C., Sartor, R., Shen, Z., Briggs, S.P., Vaughan, M.M., Alborn, H.T., Teal, P.E., and Schmelz, E.A. (2013). Plant elicitor peptides are conserved signals regulating direct and indirect antiherbivore defense. Proc Natl Acad Sci U S A **110:** 5707-5712.

Keinath, N.F., Waadt, R., Brugman, R., Schroeder, J.I., Grossmann, G., Schumacher, K., and Krebs, M. (2015). Live Cell Imaging with R-GECO1 Sheds Light on flg22- and Chitin-Induced Transient [Ca(2+)]cyt Patterns in Arabidopsis. Mol Plant **8:** 1188-1200.

Klauser, D., Desurmont, G.A., Glauser, G., Vallat, A., Flury, P., Boller, T., Turlings, T.C., and Bartels, S. (2015). The Arabidopsis Pep-PEPR system is induced by herbivore feeding and contributes to JA-mediated plant defence against herbivory. J Exp Bot **66:** 5327-5336.

Kohler, D., Montandon, C., Hause, G., Majovsky, P., Kessler, F., Baginsky, S., and Agne, B. (2015). Characterization of chloroplast protein import without Tic56, a component of the 1-megadalton translocon at the inner envelope membrane of chloroplasts. Plant Physiol **167:** 972-990.

Krogman, W., Sparks, J.A., and Blancaflor, E.B. (2020). Cell Type-Specific Imaging of Calcium Signaling in Arabidopsis thaliana Seedling Roots Using GCaMP3. Int J Mol Sci **21**.

Krol, E., Mentzel, T., Chinchilla, D., Boller, T., Felix, G., Kemmerling, B., Postel, S., Arents, M., Jeworutzki, E., Al-Rasheid, K.A., Becker, D., and Hedrich, R. (2010). Perception of the Arabidopsis danger signal peptide 1 involves the pattern recognition receptor AtPEPR1 and its close homologue AtPEPR2. J Biol Chem **285:** 13471-13479.

Lori, M., van Verk, M.C., Hander, T., Schatowitz, H., Klauser, D., Flury, P., Gehring, C.A., Boller, T., and Bartels, S. (2015). Evolutionary divergence of the plant elicitor peptides (Peps) and their receptors: interfamily incompatibility of perception but compatibility of downstream signalling. J Exp Bot **66:** 5315-5325.

McGurl, B., Pearce, G., Orozco-Cardenas, M., and Ryan, C.A. (1992). Structure, expression, and antisense inhibition of the systemin precursor gene. Science **255:** 1570-1573.

McLuskey, K., Rudolf, J., Proto, W.R., Isaacs, N.W., Coombs, G.H., Moss, C.X., and Mottram, J.C. (2012). Crystal structure of a Trypanosoma brucei metacaspase. Proc Natl Acad Sci U S A **109:** 7469-7474.

Minina, E.A., Stael, S., Van Breusegem, F., and Bozhkov, P.V. (2014). Plant metacaspase activation and activity. Methods Mol Biol **1133:** 237-253.

Pearce, G., Yamaguchi, Y., Munske, G., and Ryan, C.A. (2008). Structure-activity studies of AtPep1, a plant peptide signal involved in the innate immune response. Peptides **29:** 2083-2089.

Poncini, L., Wyrsch, I., Denervaud Tendon, V., Vorley, T., Boller, T., Geldner, N., Metraux, J.P., and Lehmann, S. (2017). In roots of Arabidopsis thaliana, the damage-associated molecular pattern AtPep1 is a stronger elicitor of immune signalling than flg22 or the chitin heptamer. PLoS One **12:** e0185808.

Razzell, W., Evans, I.R., Martin, P., and Wood, W. (2013). Calcium flashes orchestrate the wound inflammatory response through DUOX activation and hydrogen peroxide release. Curr Biol **23:** 424-429.

Ross, A., Yamada, K., Hiruma, K., Yamashita-Yamada, M., Lu, X., Takano, Y., Tsuda, K., and Saijo, Y. (2014). The Arabidopsis PEPR pathway couples local and systemic plant immunity. EMBO J **33:** 62-75.

Schardon, K., Hohl, M., Graff, L., Pfannstiel, J., Schulze, W., Stintzi, A., and Schaller, A. (2016). Precursor processing for plant peptide hormone maturation by subtilisin-like serine proteinases. Science **354:** 1594-1597.

Schindelin, J., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods **9:** 676-682.

Shannon, E.K., Stevens, A., Edrington, W., Zhao, Y., Jayasinghe, A.K., Page-McCaw, A., and Hutson, M.S. (2017). Multiple Mechanisms Drive Calcium Signal Dynamics around Laser-Induced Epithelial Wounds. Biophys J **113:** 1623-1635.

Stael, S., Wurzinger, B., Mair, A., Mehlmer, N., Vothknecht, U.C., and Teige, M. (2012). Plant organellar calcium signalling: an emerging field. J Exp Bot **63:** 1525-1542.

Stes, E., Laga, M., Walton, A., Samyn, N., Timmerman, E., De Smet, I., Goormachtig, S., and Gevaert, K. (2014). A COFRADIC protocol to study protein ubiquitination. J Proteome Res **13:** 3107-3113.

Tang, J., Han, Z., Sun, Y., Zhang, H., Gong, X., and Chai, J. (2015). Structural basis for recognition of an endogenous peptide by the plant receptor kinase PEPR1. Cell Res **25:** 110-120.

Tintor, N., Ross, A., Kanehara, K., Yamada, K., Fan, L., Kemmerling, B., Nurnberger, T., Tsuda, K., and Saijo, Y. (2013). Layered pattern receptor signaling via ethylene and endogenous elicitor peptides during Arabidopsis immunity to bacterial infection. Proc Natl Acad Sci U S A **110:** 6211-6216.

Toyota, M., Spencer, D., Sawai-Toyota, S., Jiaqi, W., Zhang, T., Koo, A.J., Howe, G.A., and Gilroy, S. (2018). Glutamate triggers long-distance, calcium-based plant defense signaling. Science **361:** 1112-1115.

Tsiatsiani, L., Van Breusegem, F., Gallois, P., Zavialov, A., Lam, E., and Bozhkov, P.V. (2011). Metacaspases. Cell Death Differ **18:** 1279-1288.

Tsiatsiani, L., Timmerman, E., De Bock, P.J., Vercammen, D., Stael, S., van de Cotte, B., Staes, A., Goethals, M., Beunens, T., Van Damme, P., Gevaert, K., and Van Breusegem, F. (2013). The Arabidopsis metacaspase9 degradome. Plant Cell **25:** 2831-2847.

Uren, A.G., O'Rourke, K., Aravind, L.A., Pisabarro, M.T., Seshagiri, S., Koonin, E.V., and Dixit, V.M. (2000). Identification of paracaspases and metacaspases: two ancient families of caspase-like proteins, one of which plays a key role in MALT lymphoma. Mol Cell **6:** 961-967.

Venne, A.S., Solari, F.A., Faden, F., Paretti, T., Dissmeyer, N., and Zahedi, R.P. (2015). An improved workflow for quantitative N-terminal charge-based fractional diagonal chromatography (ChaFRADIC) to study proteolytic events in Arabidopsis thaliana. Proteomics **15:** 2458-2469.

Vercammen, D., van de Cotte, B., De Jaeger, G., Eeckhout, D., Casteels, P., Vandepoele, K., Vandenberghe, I., Van Beeumen, J., Inze, D., and Van Breusegem, F. (2004). Type II metacaspases Atmc4 and Atmc9 of Arabidopsis thaliana cleave substrates after arginine and lysine. J Biol Chem **279:** 45329-45336.

Vercammen, D., Belenghi, B., van de Cotte, B., Beunens, T., Gavigan, J.A., De Rycke, R., Brackenier, A., Inze, D., Harris, J.L., and Van Breusegem, F. (2006). Serpin1 of Arabidopsis thaliana is a suicide inhibitor for metacaspase 9. J Mol Biol **364:** 625-636.

Watanabe, N., and Lam, E. (2011a). Calcium-dependent activation and autolysis of Arabidopsis metacaspase 2d. J Biol Chem **286:** 10027-10040.

Watanabe, N., and Lam, E. (2011b). Arabidopsis metacaspase 2d is a positive mediator of cell death induced during biotic and abiotic stresses. Plant J **66:** 969-982.

Wen, S., Ma, Q.M., Zhang, Y.L., Yang, J.P., Zhao, G.H., Fu, D.Q., Luo, Y.B., and Qu, G.Q. (2013). Biochemical evidence of key residues for the activation and autoprocessing of tomato type II metacaspase. FEBS Lett **587:** 2517-2522.

Willems, P., Horne, A., Van Parys, T., Goormachtig, S., De Smet, I., Botzki, A., Van Breusegem, F., and Gevaert, K. (2019). The Plant PTM Viewer, a central resource for exploring plant protein modifications. Plant J **99:** 752-762.

Willems, P., Ndah, E., Jonckheere, V., Stael, S., Sticker, A., Martens, L., Van Breusegem, F., Gevaert, K., and Van Damme, P. (2017). N-terminal Proteomics Assisted Profiling of the Unexplored Translation Initiation Landscape in Arabidopsis thaliana. Mol Cell Proteomics **16:** 1064-1080.

Wong, A.H., Yan, C., and Shi, Y. (2012). Crystal structure of the yeast metacaspase Yca1. J Biol Chem **287:** 29251-29259.

Wrzaczek, M., et al. (2015). GRIM REAPER peptide binds to receptor kinase PRK5 to trigger cell death in Arabidopsis. EMBO J **34:** 55-66.

Yamada, K., Yamashita-Yamada, M., Hirase, T., Fujiwara, T., Tsuda, K., Hiruma, K., and Saijo, Y. (2016). Danger peptide receptor signaling in plants ensures basal immunity upon pathogen-induced depletion of BAK1. EMBO J **35:** 46-61.

Yamaguchi, Y., Huffaker, A., Bryan, A.C., Tax, F.E., and Ryan, C.A. (2010). PEPR2 is a second receptor for the Pep1 and Pep2 peptides and contributes to defense responses in Arabidopsis. Plant Cell **22:** 508-522.

Zhang, H., Gannon, L., Hassall, K.L., Deery, M.J., Gibbs, D.J., Holdsworth, M.J., van der Hoorn, R.A.L., Lilley, K.S., and Theodoulou, F.L. (2018). N-terminomics reveals control of Arabidopsis seed storage proteins and proteases by the Arg/N-end rule pathway. New Phytol **218:** 1106-1126.

Chapter 5

"Identification of molecular partners and substrates of *AtMC4* during wounding stress".

Álvaro Daniel Fernández-Fernández[1,2], Patrick Willems[1,2], An Staes[3,4], Kris Gevaert[3,4], Simon Stael[1,2,3,4] and Frank Van Breusegem[1,2]

[1] Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium

[2] Center for Plant Systems Biology, VIB, 9052 Ghent, Belgium

[3] VIB-UGent Center for Medical Biotechnology, Ghent University, Technologiepark Zwijnaarde 75, Ghent 9052, Belgium.

[4] Department of Biomolecular Medicine, Ghent University, Technologiepark Zwijnaarde 75, Ghent 9052, Belgium.

Author contributions:

ADFF and SS conceived and designed the experiments, performed the research, analysed the data. ADFF wrote the manuscript with the help of SS and FVB. SS, FVB and KG conceived and supervised the project. PW and AS performed the MS/MS analysis and contributed to the computational analysis.

Aim and context

Metacaspases are a family of conserved cysteine proteases in non-metazoans, including plants. Two classes, type-I and type-II, are present in most plant clades. However, our current knowledge on the functions and downstream signalling caused by plant metacaspase proteolytic activities are still scarce. Type-I metacaspases are arduous to study, given some recalcitrant properties to be recombinantly produced and purified in non-host systems. On the other hand, six type-II metacaspase genes are expressed in Arabidopsis, with possible redundant activities. Among them, AtMC4 is the most abundantly expressed metacaspase in seedlings (Böllhoner et at, 2014; Hander et al, 2019) and it is the major protease responsible for the processing of the elicitor peptide AtPROPEP1 into the bioactive AtPep1 during wounding stress (Hander and Fernández-Fernández et al, 2019). Several interesting open questions arise here, like what is the identity and functionalities of other protein substrates of AtMC4; how is AtMC4 activity regulated prior and after self-processing and what is the fate of AtMC4 after exerting its proteolytic activity?

In order to answer some of these questions and thereby enhance our knowledge of the function of AtMC4 in wounding and other stresses we have set up a series of proteomics experiments that allow us to chart the substrate landscape, interactors and potential regulators.

Abstract:

Metacaspases are a family of cysteine proteases with functions related to stress and developmental responses. Here, we focus on the functions of AtMC4, a calcium dependent metacaspase responsible for the maturation of the immunomodulatory peptide AtPep1. Through a combined terminomics and shotgun proteomics approaches, we compared wild-type Arabidopsis with mutants lacking AtMC4 activity during wounding stress. We have validated AtMC4 substrate candidates by assessing their cleavage *in vitro* or *in vivo*. We identified Cystatin-1 a protease inhibitor as a new substrate of AtMC4 with possible implication in the control of the activity of AtMC4. Moreover, we studied changes in the proteome dynamics rapidly after wounding stress and between genotypes. This is the first N-terminomics study of wounding stress in plant roots, providing with an ample set of protease substrates and putative metacaspases targets. The data obtained will be a useful resource for assessing early wounding events and identification of metacaspases and other protease substrates in wounding.

Keywords: metacaspases, proteomics, *Arabidopsis thaliana*, N-terminomics, proteases, protease inhibitors, degradomics, wounding.

INTRODUCTION

Plant proteases are important enzymes for the correct functioning of plants under normal and stress conditions (van der Hoorn, 2008). Through proteolysis, many signalling pathways can be controlled, guarded or initiated. Some of these processes are studied in depth due to their implications in resistance to stresses and pathogens (Hou et al., 2021). Nevertheless, most proteolytic processes in plants and their mechanism are only partially scrutinized. In fact, the current knowledge on the mode of action of proteolytic enzymes and the identity of their targets are largely missing, existing only for a few plant proteases, which has been elaborated in chapter 2. Among these hydrolytic enzymes, metacaspases are likely the most studied family. Metacaspases are cysteine proteases involved in the response to multiple stresses (Tsiatsiani et al., 2011; Minina et al., 2017), at times in an antagonistic manner (Coll et al., 2010). For a better overview of metacaspase classification and their functions in plants to date, I refer to the chapter 1 of this thesis. Most type-II metacaspases depended on an increase in the levels of available calcium to self-process and cleave their substrates (Vercammen et al., 2004; Watanabe and Lam, 2011a; Hander et al., 2019; Shen et al., 2019),with the exception of a subgroup represented by AtMC9 and alike, which in contrast are activated at low pH values (Vercammen et al., 2004).

Due to the irreversibility of the action of proteases, the processed substrates display a new N- and C-termini with free amine and carboxyl group. This property has been exploited to chemically target these groups and identify the proteases substrates. In the last decades, different proteomic techniques involving mass spectrometry have focused in the identification of protease substrates (Staes et al., 2008; Kleifeld et al., 2010; Staes et al., 2017; Solis and Overall, 2018). In plants, the implementation of N-terminomics has also enabled the identification of plant protease substrates. First, the N-terminome of AtMC9 pioneered the field in Arabidopsis plants at early time points after germination using CoFraDiC (Tsiatsiani et al., 2013). Other N-terminome studies in plants have been successfully completed, such as ChaFraDiC or HUNTER methods (Venne et al., 2015; Weng et al., 2019) and reviewed elsewhere (Demir et al., 2018; Perrar et al., 2019). Positional proteomics can also provide information about other modifications than proteolysis, like N-terminal acetylation (Bienvenut et al., 2012; Kohler et al., 2015; Linster et al., 2015) or identification of degradation substrates in the N-End Rule Pathway (Zhang et al., 2015; Zhang et al., 2018). Although positional

proteomics studies can result very helpful in the identification of protease/proteasomal targets, the number of studies in this area are not very abundant, especially the plant field. When properly designed, these type of approaches can reveal valuable information not accessible with “traditional” methods. It is worth to mention that a good design of these experiments includes contrasting samples with different proteolytic activity. For instance, by comparison of an overexpression or wild type line versus a mutant lacking proteolytic activity *in vivo*. In specific cases, reduction of the enzymatic activity can be achieved by addition of inhibitors in the proteome sample, *e.g.* the application of proteasomal inhibitors aid in the stabilization of peptides and protein lifetime of the targets of E3 ligase PROTEOLYSIS 6 (Zhang et al., 2015). Furthermore, studies *in vitro* by co-incubation of a proteome with a recombinant active or inactive protein can pinpoint to direct substrates and avoid misidentification of differentially regulated proteins by cause of the mutant background. Ideally, both *in vivo* and *in vitro* pipelines can be done in parallel (Tsiatsiani et al., 2013). N-terminomics studies can be useful in the identification of relevant substrates with importance in biological processes, which were previously unknown. For instance, the modulation of PEPCK1 activity by AtMC9 in Tsiatsiani et al.,(2013) has been later corroborated by independent studies showing that AtMC9 processing tunes PEPCK1 activity in the plant transition from heterotrophic to autotrophic growth (Rojas et al., 2020).

In this chapter, we assessed the N-terminome of AtMC4 under wounding conditions. We used a modification of the CoFraDiC protocol including enrichment of N-terminal peptides by strong cation exchange (SCX) in combination with shotgun proteomics allowing global protein identification. Based on our previous results and fast activation of AtMC4 according to PROPEP1 dynamics after wounding (See chapter 4), a timepoint of five minutes was set as sampling for wounded conditions. As such, we used wild type plants (*Ler*) and METACASPASE 4 deficient plants (*mc4 KO*) under control conditions (unwounded) or five minutes after wounding. Each one of the conditions was performed in four biological repeats for the N-terminome and shotgun proteomics pipeline. Combination of both omics approaches allows 1) the identification of N-terminal peptides in normal and wound conditions revealing possible AtMC4 substrates; and 2) characterization of relative protein abundance affected by the genotype and/or the wounding stress.

RESULTS

AtMC4 is a nucleo-cytoplasmic, abundantly expressed protease in unstressed conditions

Among the metacaspases in Arabidopsis, *AtMC4* (AT1G79340) is expressed both in aerial and underground tissues (Fig. 1A). AtMC4 was previously reported to be expressed in *Arabidopsis thaliana* roots (Watanabe and Lam, 2011b). Here we also show expression in specialized structures like trichomes, in root trichoblast cell lineage that develop into root hairs and in mature root hairs, in addition to the expression in cells of true leaves with an *AtMC4* transcriptional reporter line (*pAtMC4::GFP-GUS; Fig. 1A*). *AtMC4* has a higher level of expression when compared to other metacaspases (Bollhoner et al., 2013). The localization of AtMC4 fused to fluorescent proteins both at the C- and N-terminus of AtMC4 showed a predominant localization in the cytosol, with lower levels in the nucleus of *Nicotiana benthamiana* epidermal cells, when transiently expressed under the constitutive promoter 35S (CaMV 35S; Fig. 1B, upper panel). This localization pattern remained unchanged when the AtMC4 active site was mutated (Fig. 1B, lower panel).

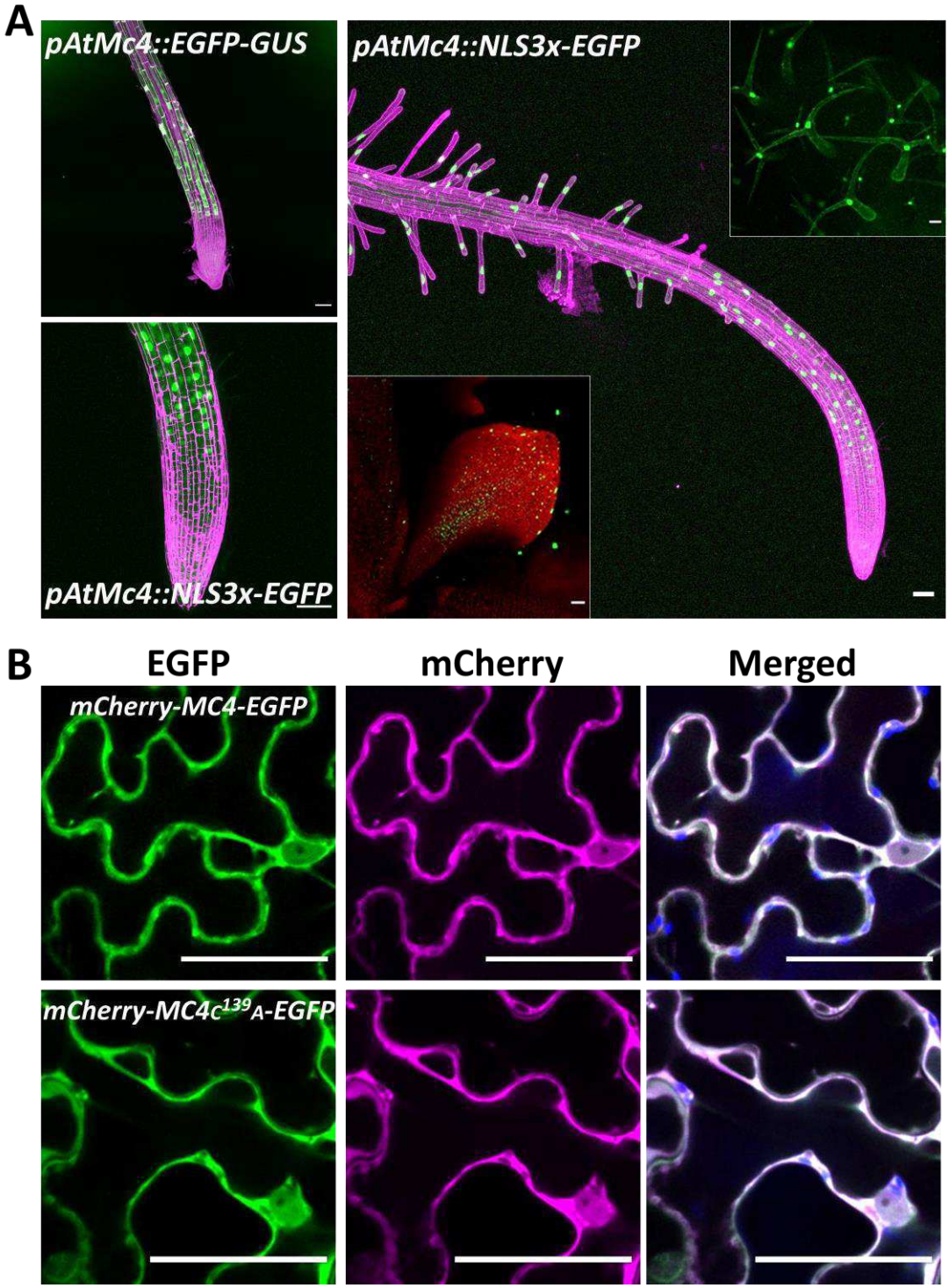


Figure 1. AtMC4 expression. A) *AtMC4* transcriptional line *pAtMC4::GFP-GUS* showing expression in trichoblasts above the root differentiation zone in *Arabidopsis thaliana*. Nucleus labelled in root hairs and trichoblasts (*pAtMC4::NLS-3xGFP*). Expression of AtMC4 by labelled nuclei in green leaves (Magenta: Propidium iodide; Green: GFP; Red: chlorophyll derived autofluorescence). **B)** Transient expression in *N. benthamiana* of mCherry-AtMC4-EGFP under the control of 35S promoter with a wild type version of AtMC4or a catalytically inactive AtMC4 C^{139}A.(Green: GFP, Magenta: mCherry, Blue: chlorophyll autofluorescence).

Wounding in AtMC4 depleted lines results in perturbed expression of stress genes

To find a response downstream of MC4 in wounding, we attempted to find wound-inducible MC4-dependent genes. Also, given that PROPEP maturation depends on metacaspases activity (Hander et al., 2019; Shen et al., 2019), and exogenous application of Peps leads to well-defined transcriptional reprogramming (Dressano et al., 2020; Rich-Griffin et al., 2020) we used wild type plants (Ler) and T-DNA insertion mutant in the Ler background (*mc4KO*) to measure the differential transcriptional changes at early wounding stages (Control, 5 and 15 minutes). We performed RT-qPCR for several genes known to be induced by exogenous Pep1 treatment (Personal communication with Claude Becker), herbivory (Scholz et al., 2014) and JA controlled genes (obtained by searches in Genevestigator). We found early induction of some stress-related genes DREB/ATERF109 (AT1G22810), ARM (AT3G02840), DRF (AT2G34930), NAC (AT3G12910) and ATMKK9 (AT1G73500). However, transcript expression for these known Pep1-induced genes did not show any statistically significant differences between wild type plants (Ler) and *mc4KO* mutants (Fig 2). In fact, DREB seemed to be induced in *mc4KO* compared to wild type at 15 minutes after wounding.

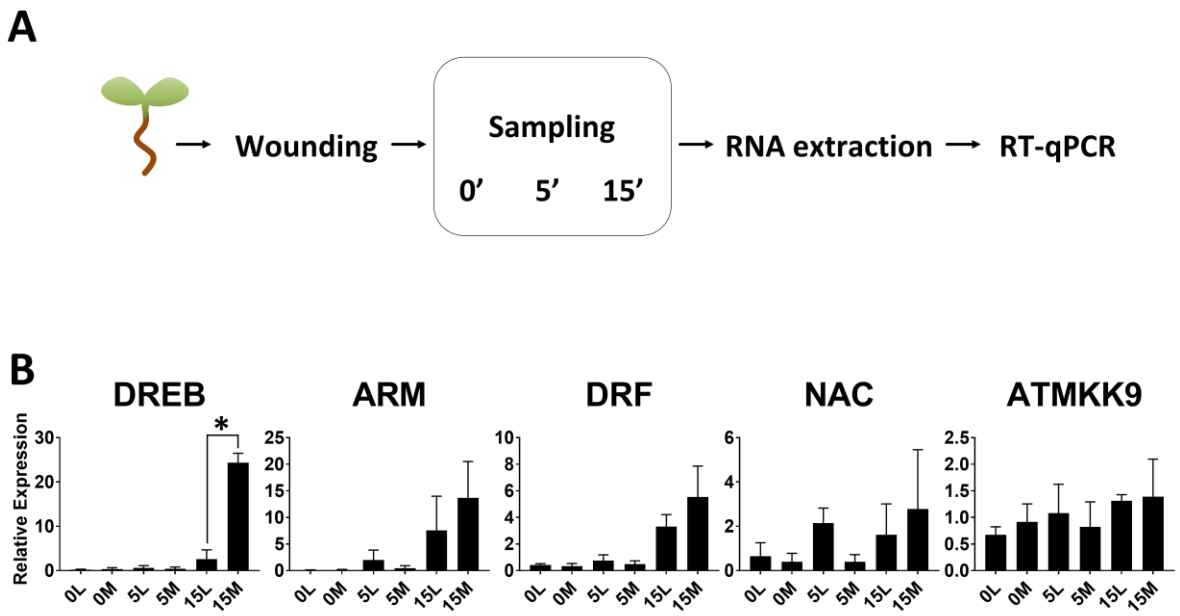


Figure 2. Transcriptional response of early wound-induced genes. A) Protocol followed for the RNA sampling of wild type and *mc4KO lines*. **B)** Bar graphs showing the relative expression normalized to ARP and UBP of Ler (L) and mc4KO (M) at control conditions (0L/0M), 5 and 15 minutes after wounding (5L/5M and 15L/15M, respectively).

N-terminome from wounded roots reveals AtMC4 substrate candidates

Although mutant plants for AtMC4 did not show any growth or developmental phenotypes under standardised lab conditions, we aimed to find substrates in order to gain knowledge of AtMC4 functions. We used wounded Arabidopsis root tissues as known triggering conditions for AtMC4 (Hander et al., 2019) (Fig 3A), for the detection of protease substrates and overall protein levels by N-terminomics and quantitative proteomics, respectively (Fig. 3B). For the N-terminome analysis, we used an adaptation of the combined fractional diagonal chromatography protocol (COFRADIC) based on the labelling of N-terminal amino acids with a heavy isotope of a NHS-ester of trideutero-acetate, and enrichment using strong cation exchange (SCX) (Staes et al., 2008). The roots of wild type and mutant plants were excised with a razor and directly frozen in liquid nitrogen as control plants, or wounded and incubated for five minutes at room temperature following transfer to liquid nitrogen. Four biological repeats were obtained for each biological condition (control wild-type, control mutant, wounded wild-type and wounded mutant) of each proteomic approach (SCX and quantitative proteomics). The study allows to obtain information about the protein abundance by summarizing the total protein coverage of peptide to protein mapping, including the calculated intensities that reveal the relative abundance. Additionally, it allows the overlap of the position in which the N-terminal peptides have been detected, with the information on the calculated values. Thus, by combination of both data we can estimate the abundance of protein and positional information of the cleavage sites (Fig. 3C). These data are calculated for each protein in the sampling conditions and all plotted in a 4 way graph (Fig. 3D).

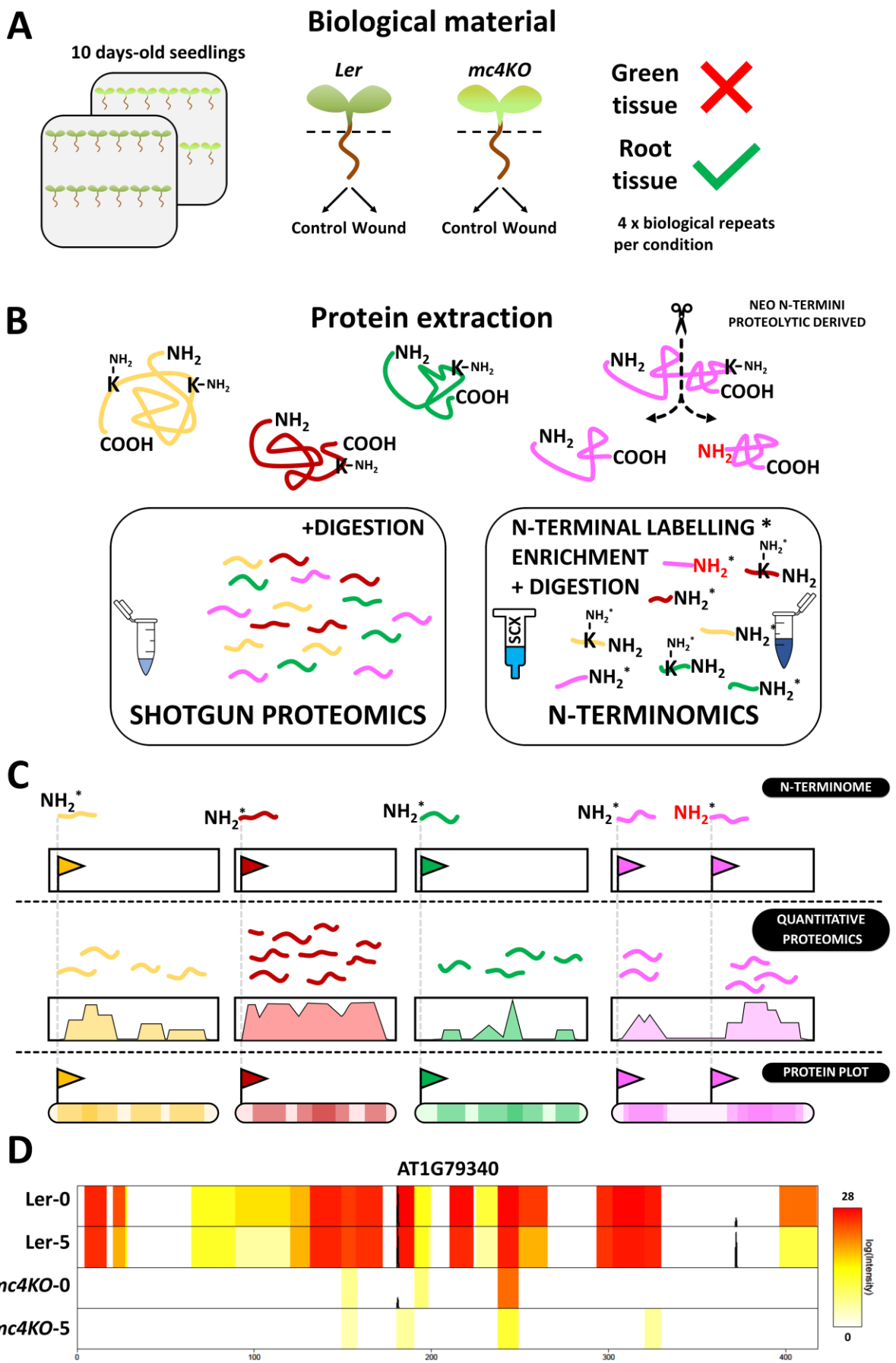

A
Biological material
10 days-old seedlings
Ler
mc4KO
Green tissue
Root tissue
Control Wound
Control Wound
4 x biological repeats per condition
B
Protein extraction
NEO N-TERMINI PROTEOLYTIC DERIVED
+DIGESTION
SHOTGUN PROTEOMICS
N-TERMINAL LABELLING *
ENRICHMENT
+ DIGESTION
SCX
N-TERMINOMICS
C
N-TERMINOME
QUANTITATIVE PROTEOMICS
PROTEIN PLOT
D
AT1G79340
Ler-0
Ler-5
mc4KO-0
mc4KO-5
28
log(Intensity)
0
0
100
200
300
400

Figure 3. Pipeline for the multi-proteomics used . A) Biological sampling 10-days old roots wild type and *mc4KO* Arabidopsis lines grown in ½ MS media grown under normal conditions and untreated (control) or wounded in 4 biological replicates. **B)** The total proteome of each sample was used for parallel shotgun proteomics and N-terminomics analysis. N-terminomics was based on the labelling of alpha and epsilon amines and enrichment of N-terminal peptides using SCX. **C)** Exemplification of the N-terminome and peptide abundance assembly obtained from the hypothetical proteins shown in B. Note that the Neo N-termini generated by proteolysis is marked in red. The N-terminome analysis allows to identified N-terminal peptide position and intensities in each of the conditions studied. N-terminal peptides includes peptides labelled at the beginning of their proteins (Methionine in position 1 or after methionine removal) as well as identification of processed proteins (pink protein). Shotgun proteomics allow the detection of peptides indicating relative abundance and/or detectability of the protein levels among the studied conditions. The combination of the data obtained by N-terminomics and shotgun proteomics can be superposed in a protein plot indicating both information. **D)** Example of a protein plot combining protein intensity and N-termini identified for AtMC4 (AT1G79340) among the different conditions used in this study.

The quantitative proteomics identified a total of 7,194 peptides matching to 5,939 protein groups and the N-terminome analysis identified a total number of 1,910 peptides in 1,297 protein groups considering all the conditions (Fig. 4A) (Online Suppl. Table 1). From the N-terminal peptides, almost a third contained the predicted protein N-terminus, of which 129 contained a peptide including the methionine in position 1, and 371 peptides contained the starting peptide with removal of the methionine. Both type of peptides were mostly endogenously acetylated (Fig. 4B). The 1,459 peptides with a position matching the amino acid protein position 3 or bigger were mostly heavy acetylated or showed no modification (free).

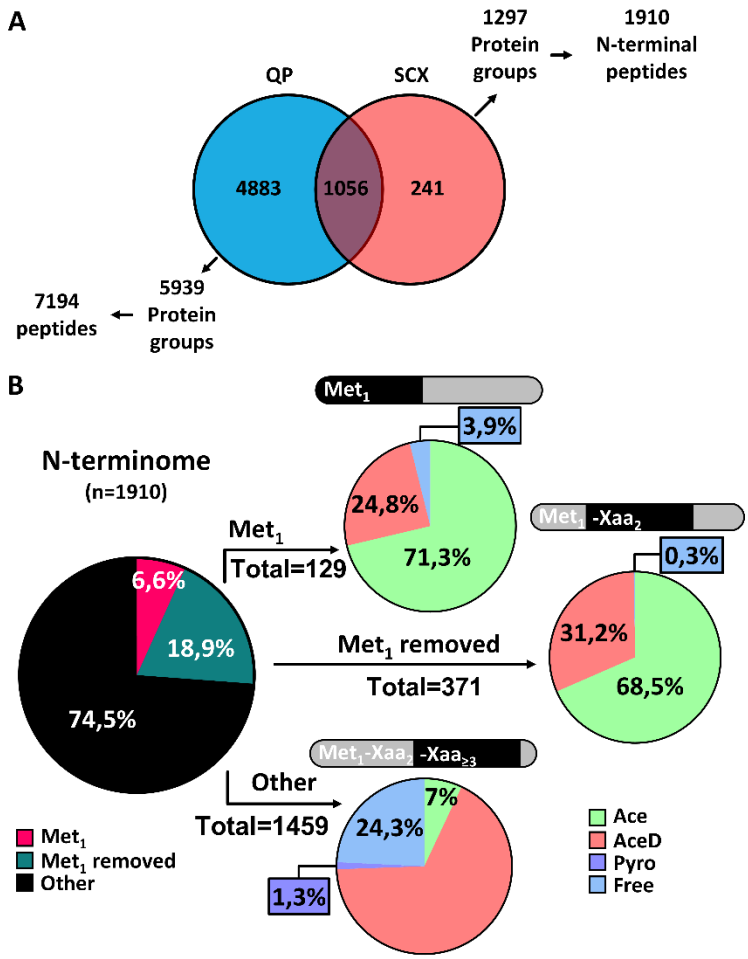


Figure 4. Overall number of identified proteins in the multi-proteomics approach A) Venn diagram showing overlap of the protein groups between QP and SCX identified proteins. The number of 5939 protein groups identified by QP corresponded to a number of 7194 peptides. The total number of 1297 protein groups contained a number of 1910 N-terminal peptides. **B)** Presence of N-terminal modifications identified in the N-terminome according to their position in the protein sequence (Fuchsia: peptides containing the first methionine; turquoise: peptides matching the protein second amino acid; black: peptides matching at a downstream position ≥ 3). Distribution of the modifications according to the peptide position in relation to the protein N-termini (Green: endogenous acetylated peptides; red: heavy acetylated peptides; purple: pyro-modified peptides; blue: unlabelled N-terminal peptides).

When considering the total pool of N-termini, independently of their position, 1,132 were heavily acetylated, 451 showed endogenous acetylation, while 19 presented a pyroglutamic modification, 360 peptides appeared to not have any additional modifications (Fig. 5A and 5B). Analysis of the amino acid preference before (position P1) and after cleavage (position P1') showed preference for arginine and methionine at position P1, and for alanine , serine and methionine at position P1' (Fig. 5C and D). This is evident in the heatmap and IceLogo generated for the complete N-terminome. Moreover, we found preferences for specific amino acids at other positions like lysine at P5 and valine at P2 and high preference for alanine at positions from P1' to P5' and lysine from P2' to P5' (Fig. 5D and E). The presence of methionine is likely due to the identification of multiple peptides mapping to predicted protein start sites that undergo methionine removal after translation. Two thirds of the total proteome can undergo methionine removal in some species (Martinez et al., 2008).

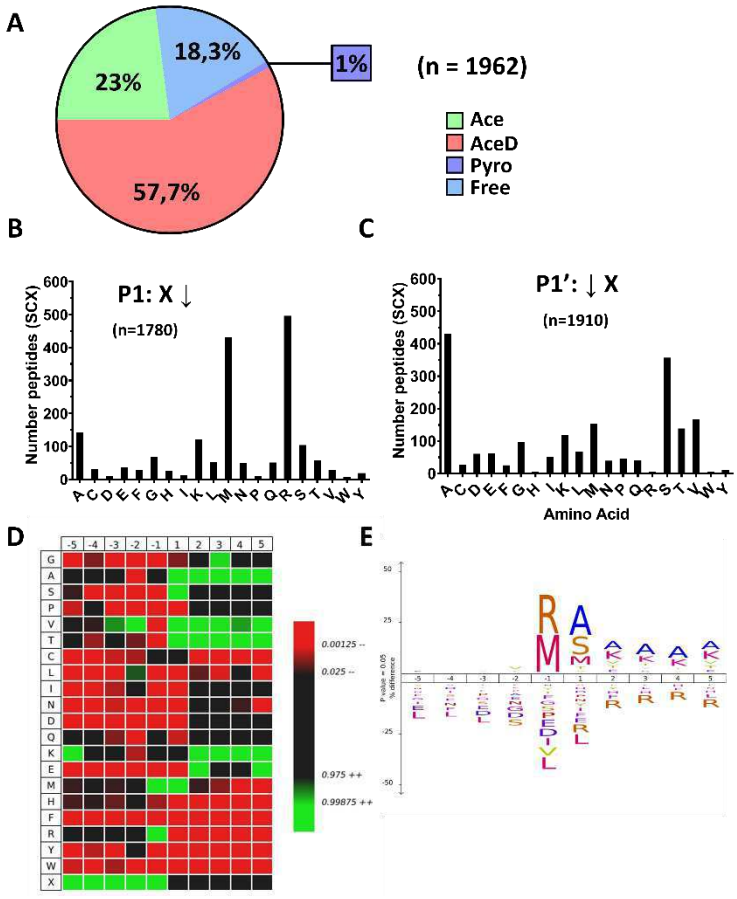


Figure 5. Root N-terminome of Arabidopsis in wounding A) Sector graph for N-terminal modification identified. (Green: endogenous acetylated; red: heavy acetylated; purple: pyro-modified; blue: free peptides). Number of heavy acetylated peptides (1132), endogenous acetylated (451), free peptides (360) and pyro-modified peptides (19) Note that some identical peptides where identified in two groups (n=1962). **B)** Prevalence of the amino acid found at P1 (P1 ↓) of the total number of the N-termini identified peptides. **C)** Prevalence of the amino acid found at P1' (↓ P1') of the total number of the N-termini identified peptides. **D)** Heatmap illustrating the amino acid frequencies of the P5-P5' spamming regions of the total heavily labelled peptides (n=1,910). **E)** IceLogo of all the N-terminal peptides identified in the N-terminome experiment (n=1,910).

In order to detect processed proteins resulting from wounding we retained only the heavy labelled N-termini (Fig. 6A). From the 1,132 selected peptides, 32 matched protein position 1, 116 matched protein position 2. Additionally 11 peptides were labelled at a position corresponding to an internal methionine in the protein sequence and 52 peptides labelled an amino acid preceded by internal methionine. We removed peptides labelled at a methionine or after a methionine, considering that those could be alternative translation initiation sites. The IceLogo showed high similitude to the total N-terminome, possibly due to the dominance of these peptides in the total N-terminome (Fig 6B). However, when observing the heatmap and Icelogo we could detect some differences such as an increase of leucine in position P2, increase of lysine and arginine at position P3 and a less noticeable increase for alanine and proline at P4 (Fig. 6C).

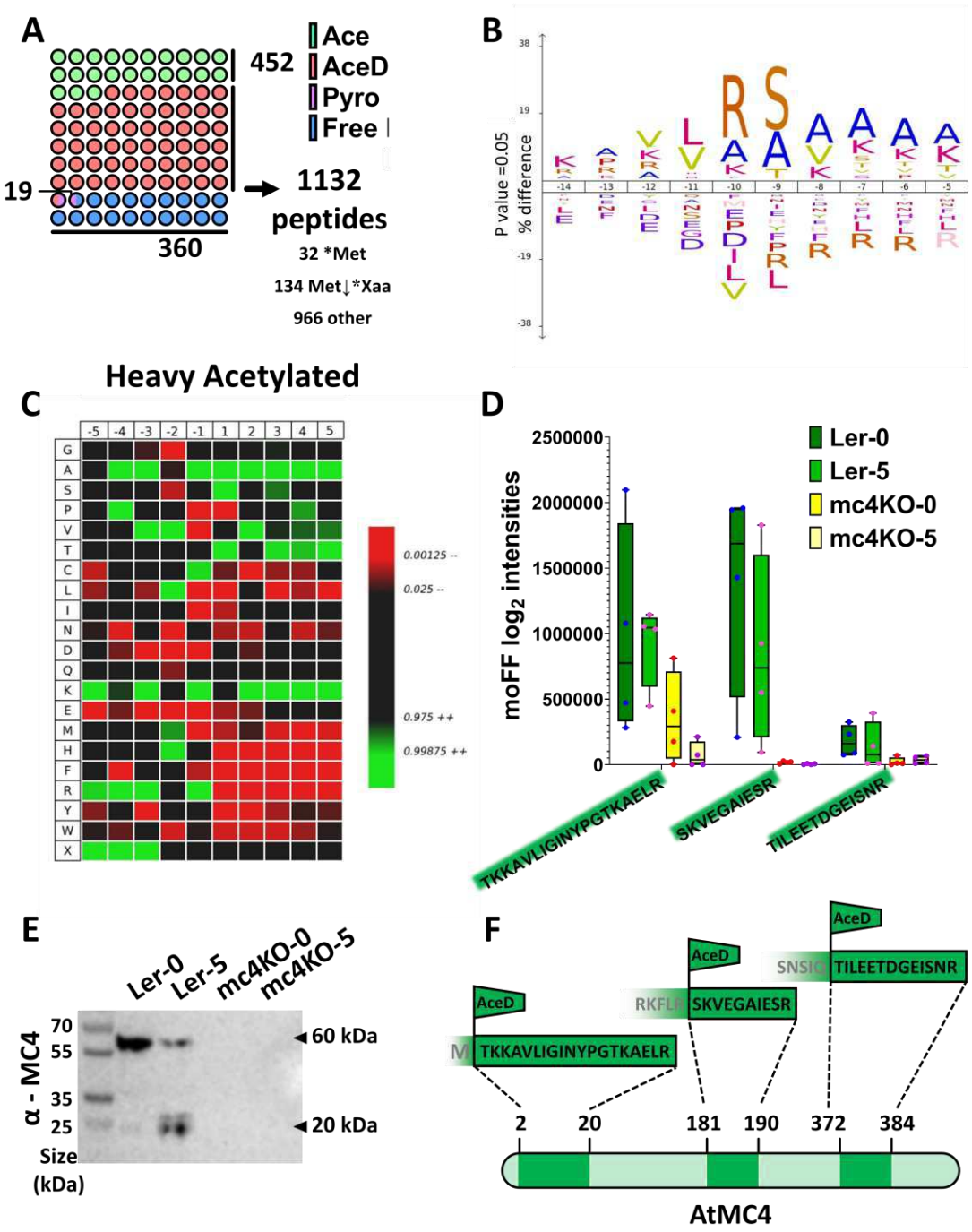


Figure 6. Heavy acetylated peptides characterize the fingerprint of Arabidopsis root terminome. **A)** Dot matrix of the total N-terminome. (n=1910). Of this a total number of 1132 were heavily acetylated. 32 peptides were heavily labelled in a methionine and 134 in an amino acid following a methionine. 966 peptides did not have methionine before or after the labelled position. **B)** IceLogo of the heavily labelled N-terminome **C)** Heatmap of the heavily labelled N-terminome (n=1132). **D)** N-terminal peptides found for Arabidopsis MC4 in the different conditions used in this study. **E)** Western-Blot detecting endogenous MC4 of protein loaded in a SDS-PAGE gel used for a biological repeat in this experiment. Bands around 60 and 20 kDa represents full size and self-processed AtMC4 respectively. **F)** Relative position of the heavy labelled peptides in the AtMC4 protein sequence.

We analysed the transformed peptide intensity for three heavy acetylated peptides matching AtMC4 and we observed a higher peptide abundance in wild type compared to the mutant background, which matches to the immunoblot for the detection of AtMC4 using specific antibody (Fig.6 D-F). The detection of peptides corresponding to the predicted N-terminus of AtMC4 (M ↓ TKKAVLIGINYPGTKAELR) in the mutant genotype could be explained by residual expression of the N-terminal part followed by protein truncation due to the T-DNA insertion.

To focus on possible substrate peptides of AtMC4 we selected heavy peptides with an arginine or a lysine in P1. Cleavage after arginine is dominant to lysine as seen before in the total N-terminome (Fig. 7A and B). To assess if arginylase activity displays different patterns to lysinase activity we splitter the results in two. For the peptides cleaved after arginine there was a predilection for serine and threonine at position P1', and valine, isoleucine and lysine at position P2', while alanine is preferred at positions P3'-5'. Additionally there was a preference for a histidine at position P2 (Fig 7C and D). The consensus sequence (P2-P5') for peptides cleaved after arginine would result in the following pattern: HR ↓ [S/T],[I/K/V]A[A/V]A (↓: denotes cleavage site). For peptides cleaved after lysine, it seems that another pattern appears. For the position P5-P3 lysines are preferred and leucine and valine at position P2. We found high promiscuity in the alanine and lysine positions after cleavage (P1'-P4' and P3'-P5' respectively), presence of valine in P1' and P2' and of serine in position P1' (Fig 7E and F). The consensus cleavage sequence would remain as the following: [A/K] [P/K] K [V/L] K ↓ [A/S/V] [A/V] [A/K] [A/K] K. Interestingly, while K at P1 would allow K at positions after cleavage, there is a high repression of arginine presence at any of the positions after cleavage.

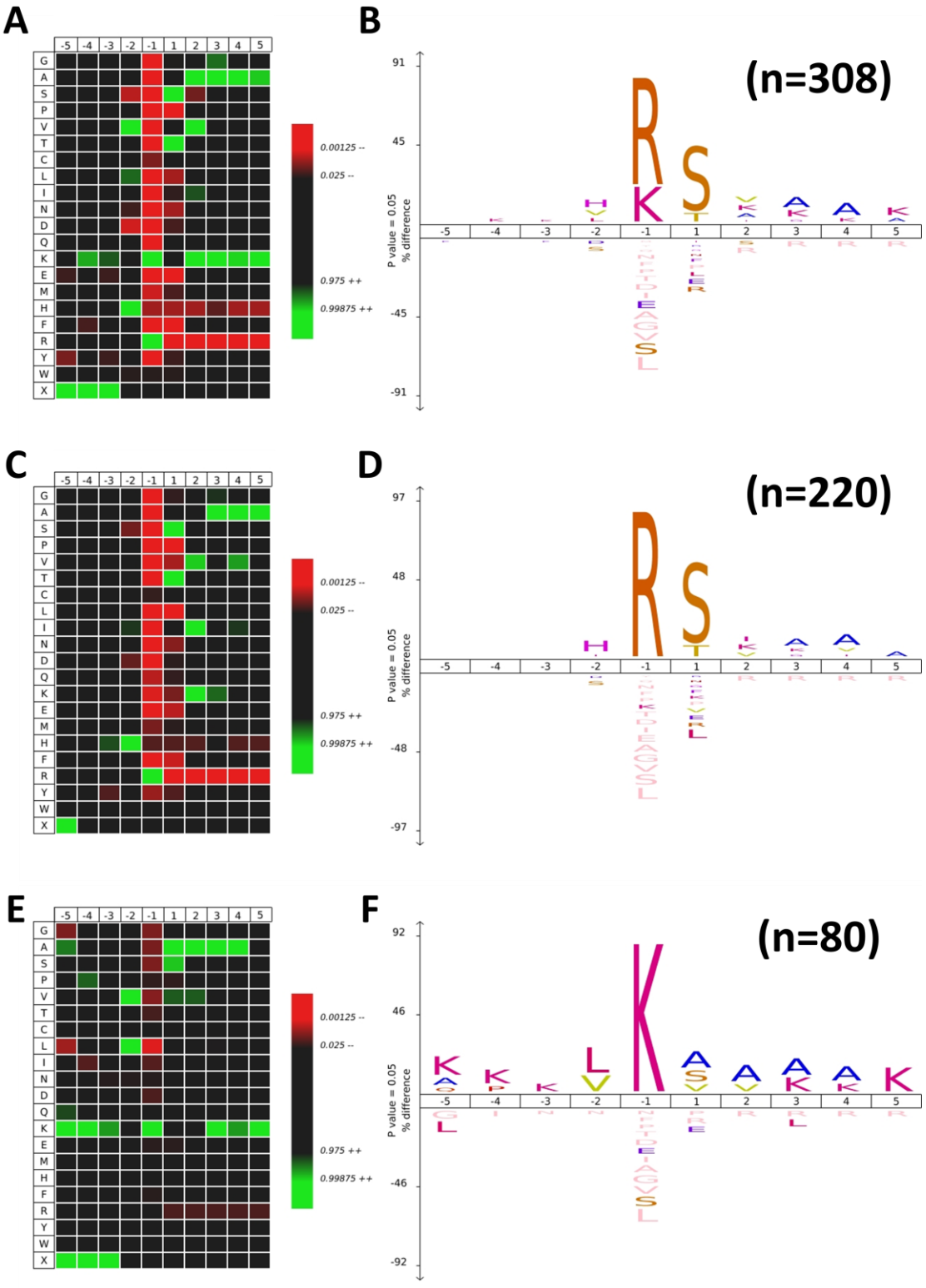


Figure 7. Patterns of cleavage for the metacaspases-like activity in Arabidopsis roots shown by the heatmaps and IceLogo of assembled amino acid sequences of the P5-P5' spamming regions of the heavily labelled peptides. A) and **B).** Representation for the argynylase and lysinase ([R/K] ↓ X) resulting peptides (n=308); **C)** and **D)** of the arginylase (R ↓ X)activity (n= 220) and **E)** and **F)** exclusive lysinase (K ↓ X)activity (n= 88).

Cleavage detection of AtMC4 candidate substrates

We used a Student's t-test for statistical analysis for the abundance of the heavy acetylated peptides for each condition, and a two-way ANOVA to compare genotype and treatment. 2-way ANOVA showed a list of 35 N-termini scoring as significant ($p<0.05$), 16 of which are cleaved after R/K (Table 1, Suppl. Fig 1). To validate AtMC4 substrates, selected candidates were first assessed for cleavage by an *in vitro* transcription and translation (TNT) assay, based of cell-free production of radioactively labelled substrate (^{S35}Met) and incubation with active and/or inactive AtMC4 (Minina et al., 2014). A protein with more abundant levels observed in the 2 way ANOVA and wounding conditions is NAI2 (AT3G15950). Four peptides were identified for NAI2 which is an ER-body associated protein with functions in wounding (Geem et al., 2019). We found 4 heavy labelled peptides for NAI2 (Fig. 8A). First we validated its cleavage using the TNT *in vitro* assays (Fig. 8C) and later its processing *in vivo* was assessed using an antibody against NAI2 (Fig. 8D). As observed in PYK10 blots, NAI2 cleavage in roots was more pronounced than in leaves (Fig. 8E).

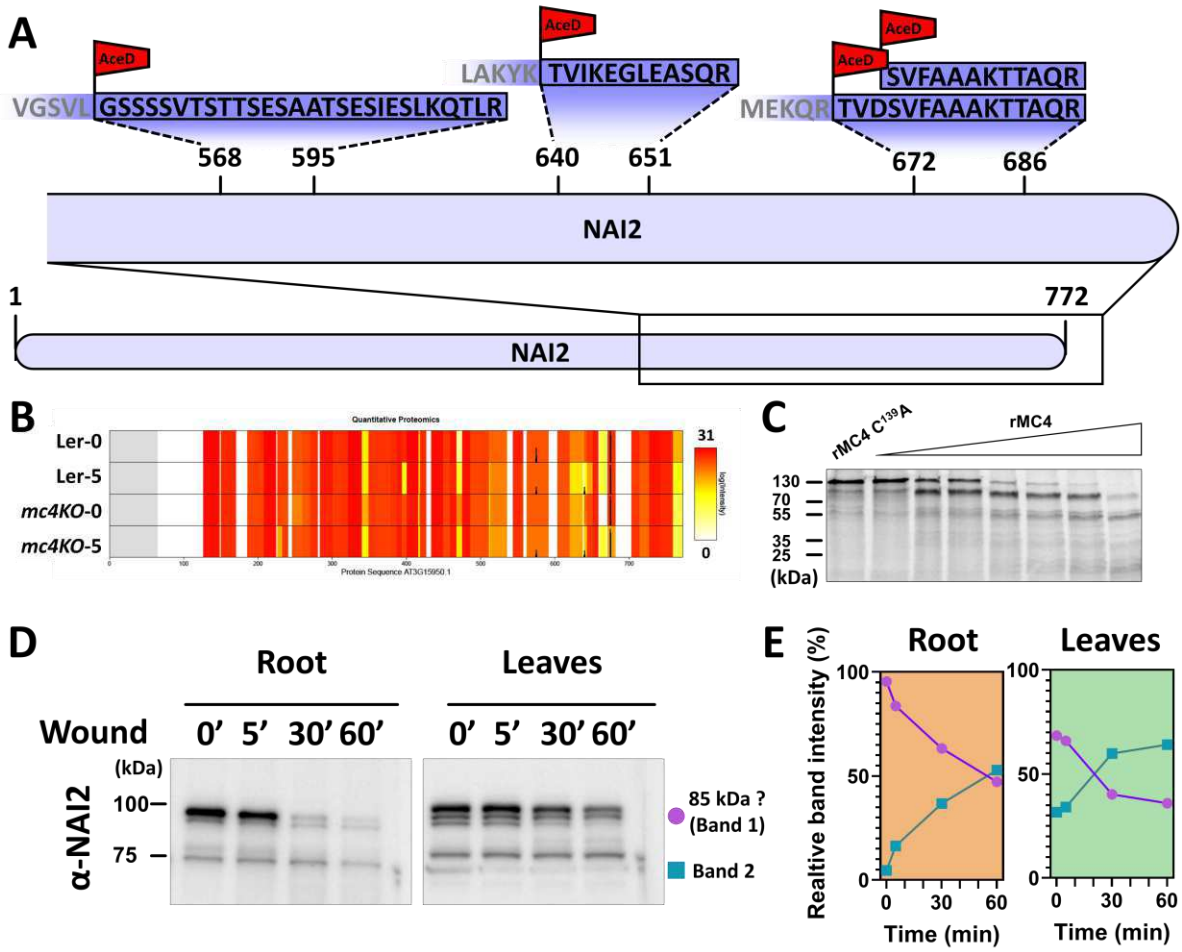


Figure 8. NAI2 is AtMC4 candidate substrate processed in response to wounding in roots and leaves. A) Heavy peptides matching NAI2 protein sequence. **B)** Overlap of the peptides from the quantitative proteomics and N-terminome detected positions in the NAI2 sequence. **C)** TNT assay for NAI2 incubated with catalytic inactive AtMC4 or active AtMC4 at different concentrations. **D)** Western Blot of root and leaf Arabidopsis protein samples subjected to wounding during 5, 30, 60 minutes or control. **E)** Measurement of proteolytic degradation as the relative signal detected in a Western Blot (Based on 1 repeat).

Moreover, we found multiple heavy and free peptides for PYK10 (AT3G09260), that were significant in wounding (Fig. 9A; Online suppl. table 1). PYK10 is a β-glucosidase expressed in Arabidopsis roots that locates to the ER (Nagano et al., 2005). Glucosidases hydrolase glucosinolates, which have functions in the defence against pathogens and pests in Brassicaceae pests (Rask et al., 2000). PYK10 was efficiently cleaved by AtMC4 in the TNT assay (Fig. 9C). Immunoblots of Arabidopsis extracted protein with endogenous antibody against PYK10 showed degradation dynamics after wounding in roots while bands in leaves remained stable (Fig. 9D). Analysis of cleavage showed a cleavage of a 50% in roots while it did not change in leaves (Fig.9E).

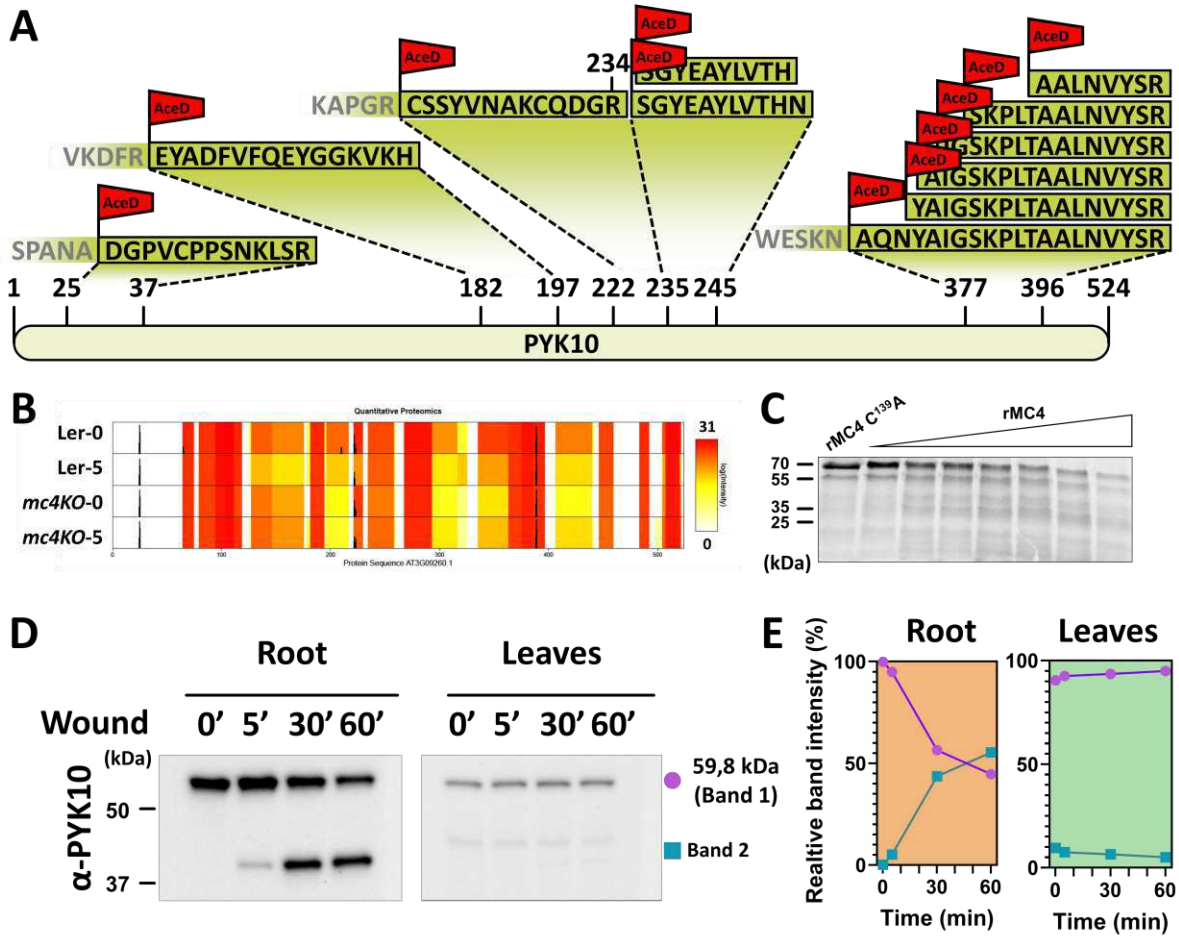


Figure 9. PYK10 is AtMC4 candidate substrate that is processed in response to wounding in Arabidopsis roots. A) Heavy peptides matching PYK10 sequences. **B)** Overlap of peptide coverage from the quantitative proteomics and N-terminome detected positions in the PYK10 sequence. **C)** TNT assay for PYK10 protein. **D)** Western Blot of root and leave Arabidopsis protein samples subjected to wounding during 5, 30, 60 minutes or control. **E)** Measurement of proteolytic degradation as the relative signal detected in a Western Blot (Based on 1 repeat).

We found 6 peptides matching the GLYCINE RICH PROTEIN 7 (GRP7, AT2G21660) in the terminome list (Fig. 10A). GRP7 is an RNA binding protein induced in cold conditions that regulates the plant responses to salt and dehydration stress (Yang et al., 2014; Meyer et al., 2017). Although the values recorded in our study did not reach significant values the 2 way ANOVA, GRP7 peptides at position 28 were significant when comparing the all the wild type samples versus the *mc4KO* conditions (Online Suppl. Table 2). Additionally, GRP7 was identified in the AtMC9 terminome study. The N-terminome results identified three peptides at the beginning of the protein at the amino acid starting positions of the peptides 2, 3 and 10. The other peptides matched

at positions 28, 33 and 78. The peptide for the last position was heavily labelled and cleaved after an arginine (DLDGR↓**SITVNEAQSR**; identified peptide in bold), being a possible metacaspase processing site. Protein levels in the coverage plot did not show differences among conditions (Fig. 10B). The *in vitro* cleavage test showed increase of cleavage when AtMC4 is present at higher concentration of AtMC4 (Fig. 10C).

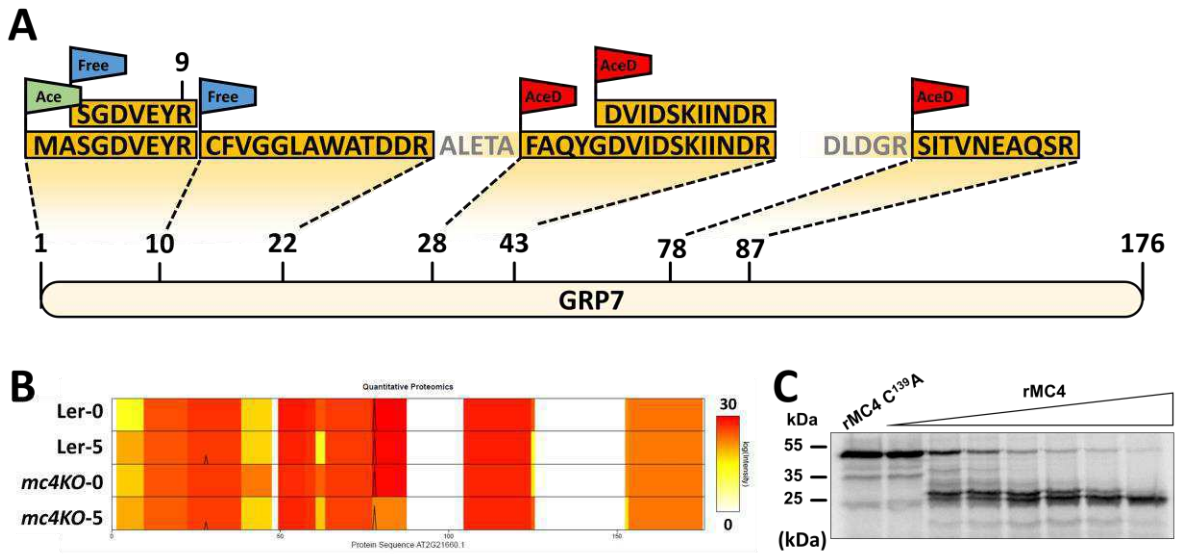


Figure 10. Active AtMC4 can process GRP7 when both are incubated in the *in vitro* assays. A) Peptides detected in the N-terminome approach matching GRP7 sequence. **B)** Overlap of the peptides from the quantitative proteomics and N-terminome positions. **C)** TNT assay showing the radiography detecting of ^{35}S-methionine corresponding to cell-free produced GRP7 and incubated with catalytic inactive AtMC4 or active AtMC4 at different concentrations gradient (increasing gradient left to right).

Moreover, some peptides identified were identical to other N-terminomics studies, like the mitochondrial MALATE DEHYDROGENASE (AT1G53240), identified in a ChaFraDiC experiment (Venne et al., 2015) and two Ribosomal L29 family proteins (AT2G39390, and AT5G02610) were also identified in a COFRADIC experiment comparing overexpression lines of AtMC9 and *mc9* mutants (Tsiatsiani et al., 2013). In the last case, the peptides were not identical, but were found to be adjacent: VVRK↓SIAQVLT↑VISQ (↓: cleavage identified in this study; ↑: cleavage identified in p35S:MC9 vs *mc9*). This can indicate that the region in the folded structure of the protein could be accessible by proteases and subjected to proteolytic attack by more than one protease. In addition, it could mean that a first cleavage can be followed by additional proteolysis by the newly exposed amino acid chain after first cleavage. We also found proteins that stabilize in AtMC4 active conditions like BPA1, a protein that interacts with accelerated cell death11 (ACD11) and that is degraded by the 26S

proteasome (Li et al., 2019). Whether AtMC4 absence might be compensated with induction of other proteolytic activities, including the proteasome is a hypothesis that requires more exploration.

We also could detect cleavage of other proteins that were not relevant in the statistical analysis like HOMOLOG OF ANTIOXIDANT 1 (ATX1, AT1G66240) and a Ribosomal L38e protein (RL38eI, AT3G59540). Three peptides located at ATX1 proximal positions 32, 39 and 41 (Suppl. Fig. 2A). Peptides 39 and 41 were heavy acetylated and the one starting to position 32 was endogenously acetylated. The Plant PTM viewer, a public resource that pools information on post-translational modifications of plant proteins including N-terminomics experiments (Willems et al., 2019), revealed that peptides initiating at sites 32 and 41 were already found before (Kohler et al., 2015; Zhang et al., 2018). Peptide abundance for the quantitative proteomics showed almost identical values between the four conditions (Fig Suppl. 2B). Moreover, *in vitro* assay showed cleavage (Fig Suppl. 2C). A peptide matching RL38eI was heavy acetylated after an arginine in position P1 (Fig Suppl. 2D). Moreover this sequence showed a dibasic motif before cleavage, which was previously observed in GRI processing as a hallmark of metacaspase cleavage (Wrzaczek et al., 2015), although this is not strictly adhered to, for example in the case of PROPEP1 (Hander et al., 2019). Protein levels were shown to be stable independently of the genotype and wounding conditions (Fig. Suppl. 8E). GST-RL38eI fusion protein was efficiently cleaved in the TNT assay at very low AtMC4 concentration (Fig. 8F).

Additionally, several peptides matched members of the Glutathione S-Transferase phi family (Fig. 11). N-terminal heavy labelled peptides were identified for GSTF8, GST9and GSTF10 and multiple positions (Fig, 11A, B and C). Surprisingly we found a great number of free peptides matching the C-termini of the three members, all of them preceded by a similar pattern [K/D]IS[S/A]R. When we examined the transformed intensity of the N-terminal peptides in the different conditions it turn out that they seem to be more abundant in wild type and / or wounding conditions in the wild type background (Fig 11D-K). The observations are clearer when comparing the intensity levels of each peptide individually as depicted for GSTF8 (Fig. 11G-K).

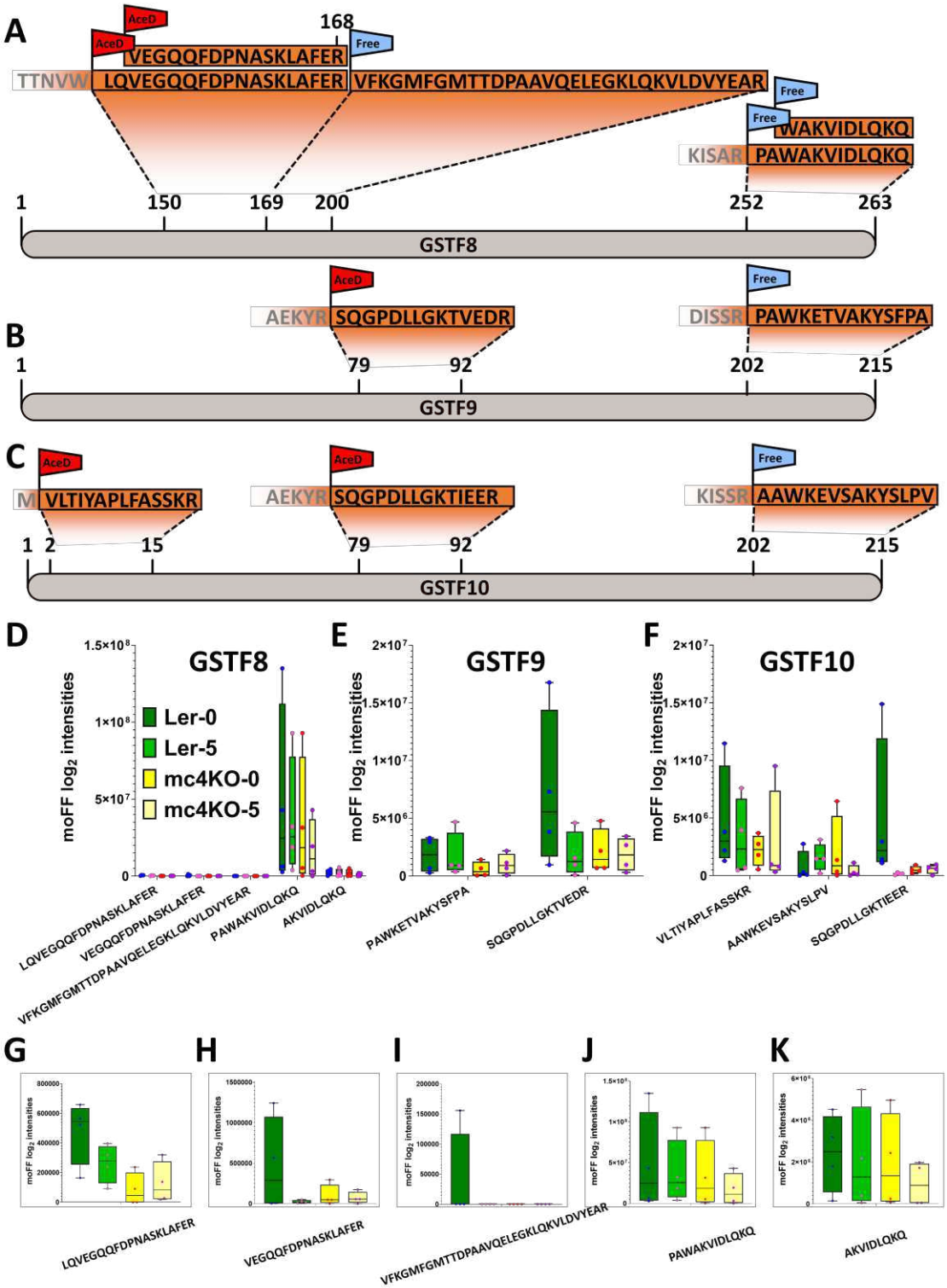


Figure 11. GSTF proteins are processed in a wound dependent manner. Peptides identified for GSTF8 **(A)**, GSTF9 **(B)** and GSTF10 **(C)**. Representation of the abundance of the N-terminal peptide identified for GSTF8 **(D)**, GSTF9 **(E)** and GSTF10 **(F)** (Ler control: dark green; Ler wounded: light green; *mc4KO* control: dark yellow; *mc4KO* wounded: light yellow). **G-H)** Individual representation of the N-terminal identified peptides for GSTF8in each condition.

In conclusion, we could validate AtMC4 substrates *in vitro* using TNT assays and incubation with active or inactive AtMC4. Three of these substrates , PYK10, NAI2 and GRP7, were statistically more abundant in wounding or in conditions with presence of AtMC4 in the genome. As such, three heavily labelled peptides corresponding to a PYK10 position (SKPLTAALNVYSR, IGSKPLTAALNVYSR and YAIGSKPLTAALNVYSR) showed higher abundance scores when comparing Ler-5 vs Ler-0 and mc4-5 vs mc4-0. This suggests that PYK10 is processed under wounding events, even in conditions when AtMC4 is not present. For NAI2, 4 peptides were heavily labelled in total. At least a peptide was significant , up or down, in each pairwise comparison and in the 2 way ANOVA (combination of genotype and treatment). Last, for GRP7 a peptide showed higher values in Ler conditions by comparison of the 8 Ler conditions (wounded and unwounded) vs the 8 *mc4KO* conditions (wounded and unwounded). Moreover, we could detect *in vivo* proteolysis after wounding using specific antibodies for PYK10 and NAI2, but the detection in lines with depleted metacaspase activity has yet to be confirmed. Also, our results indicate that GSTF peptide abundance depend on the presence of AtMC4 in the plant background, according to the peptide intensity levels.

The protease inhibitor CYS1 is likely an AtMC4 substrate *in vivo*

Another interesting protein identified is the proteinase inhibitor Cystatin 1 (CYS1; AT5G12140) (Belenghi et al., 2003). The N-terminal labelled peptides for CYS1 were present in the wild type genotype both in control and wounding conditions, but not in *mc4* mutant (Fig. 12A). One of the peptides (ADQQAGTIVGGVR) started at position 2 that was endogenously acetylated. The second peptide was detected at position 15, which was the site where the first peptide finished (Fig. 12B and Suppl. Fig.3). This second peptide was heavily labelled and preceded by an arginine (VGGVR↓**DIDANANDLQVESLAR**, in bold the peptide sequence detected) and present in Ler-0 conditions. It is possible that with the conditions used in the protein extraction protocol, AtMC4 could have been able to cleave substrates in the control conditions. Additionally a the identified location matches with the location of AtMC9 degradome. Thus, we further investigated CYS1 as a candidate interactor of AtMC4. For this, we transiently co-expressed in *N. benthamiana* C-terminal fused GFP-CYS1 with wild-type and catalytically inactive mutant versions of N- and C-terminally RFP tagged AtMC4 (AtMC4 C^{139}A-RFP, AtMC4-RFP, RFP-AtMC4 C^{139}A and RFP-AtMC4).

GFP-CYS1 derived signal was nucleo-cytoplasmic and overlapping with RFP signal of both wild-type and mutant AtMC4 constructs (Fig. 12C). Next, we performed a co-immunoprecipitation experiment to test if a N-terminal GFP tagged version of CYS1 (CYS1-GFP) could bind and/or be a substrate of AtMC4 (Fig 12D). Immunoblots showed detection of weak bands of wild-type versions of AtMC4-RFP and RFP-AtMC4 (lane 3 and 4 of the top panel in Fig. 12D). These bands would correspond to unprocessed size of AtMC4 for the lane 3. More interestingly, in the input blots, an extra lower band is present for GFP detection in those samples where wild type AtMC4 is expressed (second panel in Fig. 12D, indicated with an asterisk). Cleavage of GFP-CYS at the sites detected by N-terminomics (Fig. 12B) would generate a truncated GFP-CYS1 version with a deletion of 10 kDa from the C-terminus, similar in size to what it is seen on the blot (asterisk in Fig. 12D). This could mean that CYS1 is cleaved by AtMC4 upon co-expression in *N. benthamiana*. Further research is needed if CYS1 is just a substrate or if it also inhibits AtMC4 activity. Furthermore, cleavage should be confirmed with a mutated version of CYS1 at arginine in position 14 and in a *mc4* mutant background.

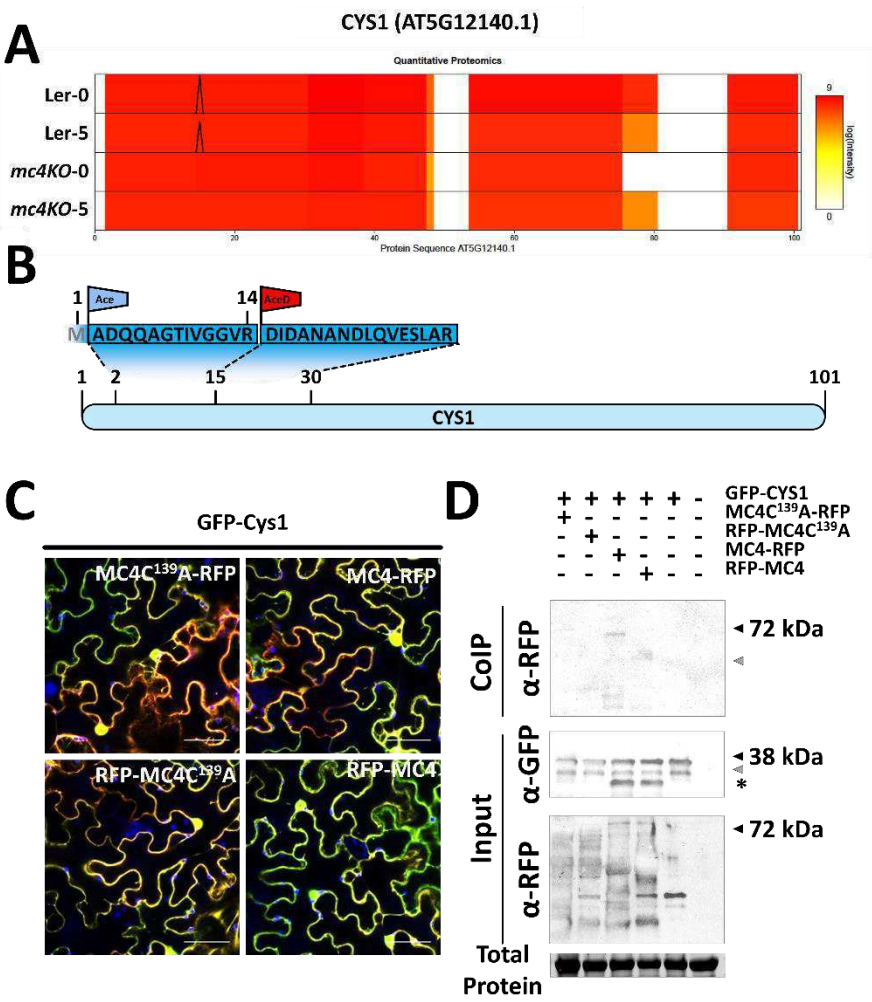


Figure 12. CYS1 is a potential substrate of AtMC4. A) Plot showing overlap of the N-termini scores for N-terminal identified peptides (black peaks) and log_2 fold change of the peptides abundance according to shotgun proteomics (yellow to red range). The plot shows 4 conditions for wild type control (*WT 0'*), wild type wounding (*WT 5'*), mutant control (*mc4 0'*) and mutant wounding (*mc4 5'*). **B)** Relative position of the peptides identified in the terminomics approach in the protein sequence of CYS1. (Blue flag endogenous acetylated peptide (Ace); red flag heavy acetylated peptide (AceD)).**C)** Confocal images of merged channels for the co-expression of GFP-CYS1 and AtMC4 RFP tagged versions for transient infiltration in *N. benthamiana*. **D)** Western-Blots showing the detection of GFP-CYS1 and RFP tagged versions of AtMC4 used for Co-IP performed in *N.benthamiana* as well as the blots showing input sampling. Weak bands were detected in the Co-IP blot corresponding to high molecular size of AtMC4 (Higher blot, back arrow). In the input blot we detected a band corresponding to GFP-CYS1 (38 kDa) and a second band of similar size (grey filled arrow). A third band only when active AtMC4 was expressed (third and fourth lanes from the left) which may correspond to the size of GFP-CYS with the release of its C-terminal part (indicated by an asterisk).

AtMC4 p20 interacts with SERPIN1

In parallel to CYS1,we tested AtMC4 and SERPIN1 interaction. Although it was not found in our proteomic study, SERPIN1 is a broad range peptidase inhibitor expressed in Arabidopsis roots (Suppl. Fig. 4A), able to inhibit AtMC9 and AtMC1 activities (Vercammen et al., 2006; Lema Asqui et al., 2018). We performed Co-IP using GFP of the SERPIN1 as bait at both its C- and N-termini when co-expressed with active and inactive versions of AtMC4 RFP-tagged at both termini (Suppl. Fig. 4B). AtMC4 is able to immuno-precipitate with SERPIN, independently of the labelling side of GFP to the inhibitor, but only the N-terminal side of active AtMC4 was eluted. The band detected had also undergone proteolysis indicating that possibly AtMC4 needs self-processing prior to SERPIN1 binding to the p20 subdomain. When the reactive centre loop of SERPIN1 was mutated to metacaspase cleavage motif tetrapeptide VRPR, interaction between SERPIN1 and AtMC4 remained stable (Suppl. Fig. 4D and E). We screened for expression patterns of SERPIN1 and identified its transcriptional induction in cold stresses according to available RNA-Seq datasets (Suppl. Fig. 4F). Temperature shifts are external stimuli that plants are able to sense and respond to. In fact, temperature shifts from room temperature to cold can induce calcium signalization (Knight et al., 1996). To test if the levels of calcium induction were similar to wounding or elicitors providing an environment for activation of AtMC4, we used the genetically encoded calcium reporter GCaMP3. We could detect immediate transient changes on calcium fluxes when the temperature was modified from 21 °C to 12 °C, and these levels were similar to those obtained with the elicitor peptide Pep1 treatment (Suppl. Fig. 4G and H). Further confirmation of the activation of AtMC4, or other metacaspases, in stresses involving calcium remain to be verified with additional experiments. These results indicate to possible novel activation of metacaspases in temperature shifts. Additionally, the same stress stimulate the transcription of SERPIN1 at later but subsequent time points. Our observations hint towards a pathway of rapid activation of metacaspases in cold stress, which is important for development of brassica species. Calcium release may allow for activity of metacaspases in a short time, and expression of its inhibitor in the frame of hours that may work as dampening mechanism. In any case, our indications are preliminary and they will require study of cold stress in combination with Arabidopsis mutant lines.

Shotgun proteomics indicate wound regulatory mechanisms involving plant proteostasis

Primarily, we aimed to use the shotgun quantitative proteomics as a comparative measure to identify potential stability, presence, or absence of protease substrate fragments (visualized by the protein coverage plots). However, we exploited this data further and investigated the proteome changes between wild type and *mc4KO* five minutes after damage, when AtMC4 is expected to be most active (Hander et al., 2019). Wounding can induce rapid events, such as hormone biosynthesis, which has been predicted to rapidly alter the plant proteome and protein stability (Larrieu et al., 2015). We detected 7,194 peptides organized in 5,939 protein groups that allowed us to understand some of the differences between conditions. We compared pairwise the difference between groups based on their genotype and treatment (Online Suppl. Table 3).

1) Ler-5 vs Ler-0.

We identified higher abundance of the protein PAMP-INDUCED COILED COIL (PICC; AT2G32240) and Jasmonate Resistant 1 (JAR1, AT2G46370) in wild type wounded plants (Suppl. Fig. 5 and Fig.13A, respectively). PICC is a protein localized at the ER membrane and contributes to callose deposition *(Venkatakrishnan et al., 2013; Wang et al., 2019)*. JAR1 is involved in the JA to JA-Ile transformation (Staswick and Tiryaki, 2004). Additionally, calmodulin CAM5 (AT2G27030) and microtubule-associated protein 18 MAP18 (AT5G44610), related to calcium signalling in root hair growth, were more abundant (Kang et al., 2017) (Kato et al., 2013). Lastly we also detected higher levels of AtMC4 in the wounding samples.

As representative of less abundant proteins, we found the calmodulin like 42 CML42 (AT4G20780), which is involved in herbivory (Vadassery et al., 2012) and ATECS1 (AT4G23100), that balances SA and phytoalexin production in response to bacterial infection (Ferrari et al., 2003). We found downregulation of the phenylpropanoid pathway enzymes *CINNAMATE-4-HYDROXYLASE C4H* (AT2G30490) and PHENYLALANINE AMMONIA-LYASE 2 PAL2 (AT3G53260).

2) Ler-5 vs mc4KO-5

Peptides for AtMC4 show low presence in *mc4KO* plants. (Fig. 11B). PICC is again abundant as well as FATTY ACID DESATURASE 7 (FAD7; AT3G11170), an enzyme

that catalyses the biosynthesis of a linoleic acid precursor of JA (Browse, 2009). The homolog of PAL2, PAL1 (AT2G37040) is more abundant in Ler-5 samples than in *mc4KO*-5 samples. CYSTATIN B is also overabundant (CYSB; AT3G12490)

Interestingly, we found abundant scores for KRATOS (AT3G23450), a predicted transmembrane protein that was previously identified in a peptidomics approach in a tracheary element cell culture with depletion of AtMC9 expression (Escamez et al., 2019). At the same, the levels of a protein enclosing an peptide with antagonistic effects to KRATOS and identified in the same study named BIA (AT3G24100) were found reduced. The kinase MRI (AT2G41970) necessary for pollen and root hair growth was also found downregulated (Boisson-Dernier et al., 2015).

3) Wild type-0 vs mc4KO-0

AtMC4 showed the highest abundance values, but we identified the glutathione-S transferase 8 phi (GSTF8; AT2G47730) that can bind to cadmium ions (Hou et al., 2019) and identified 12-OXOPHYTODIENOATE REDUCTASE 1 (OPR1; AT1G76680), important for stress detoxification (Schaller et al., 2000), was abundantly found together with PAL1 and PAL2, and the calcium dependent kinase (CDPK6/CPK3; AT4G23650), and of a pectin methylesterase (PMEPCRA; AT1G11580) (Fig 13C). As most representative downregulated proteins in this group we found several photosystem components like the subunits QA and Q-2 of the PHOTOSYSTEM II (PSBQA; AT4G21280 and PSBQ2; AT4G05180 respectively).

4) mc4KO-5 vs mc4KO-0

In the wound conditions, AtMC9, JAR1 and NPR3 (AT5G45110) homolog of NPR1 were more abundant (Fig. 13D). The cellulose synthase 1 (CESA1; AT4G32410) and some proteases like the subtilase SBT5.2 (AT1G20160) and the serine carboxypeptidase like SCPL37 (AT3G52010) were found upregulated. Contrarily, CML42, and the photosystem II subunit O-2 (PSBO2; AT3G50820), the photosystem I subunit E-2 (PSAE-2; AT2G20260)are downregulated together with KRATOS and the UNUSUAL PROTEASE INHIBITOR (UPI; AT5G43580).

By estimation of the quantitative proteomics data we have been able to determine the mutant background of the *mc4KO* lines. Also, we have been able to detect different levels in proteins linked to wound response and other calcium signalling events as well

as changes in the abundance of enzymes involved in plant hormone biosynthesis and defence to damage stresses.

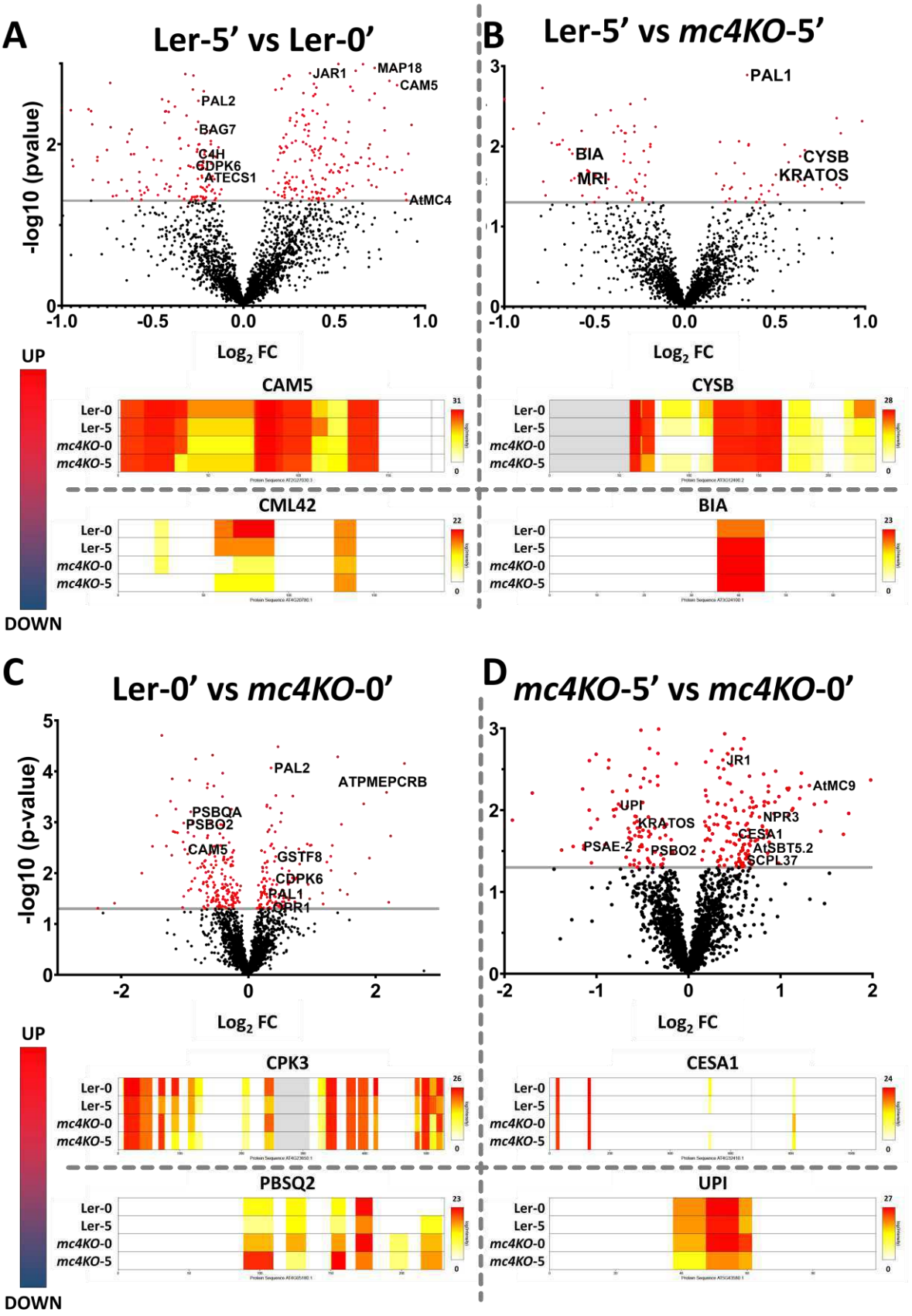


Figure 13. Protein abundance shown by volcano plots in the different conditions. Centred view of the total volcano plots view (Suppl. fig.2) showing protein levels by the values calculated for each protein were plotted considering their -log (p-value) and the log_2 FC. Significant values ($p<0.05$) above the significant threshold, displayed by the grey line, are coloured in red. Positive FDR values for each comparison means higher protein abundance while negative values represent lower abundance. Conditions plotted: **A)** Wounded wild type plants versus unwounded wild type plants (Ler-5 vs Ler-0) **B)** Unwounded wild type plants versus unwounded mutant plants (Ler-0 vs *mc4KO*-0) **C)** Wounded wild type plants versus wounded mutant plants (Ler-5 vs *mc4KO*-5) **D)** Wounded mutant plants versus unwounded mutant plants (*mc4KO*-5 vs *mc4KO*-0).

DISCUSSION

Proteases represent around 2% of the total genome in Arabidopsis (van der Hoorn, 2008; Lallemand et al., 2015), but they remain as an understudied group of enzymes when we compared them to other enzymes like kinases. Major hurdles in the study of proteases are the identification of the cellular conditions that activate them in an *in vivo* context and the detection of substrates. Here we tried to contribute to the current knowledge about a member of the metacaspase family in *Arabidopsis thaliana,* AtMC4 by expanding our knowledge of the candidate substrates in a condition known to activate AtMC4.

AtMC4 is highly expressed in root tissue as previously reported (Watanabe and Lam, 2011b) and as evidenced by available datasets showing expression levels at single cell level (Wendrich et al., 2020). Our promoter fusion showed strikingly high expression in root hairs, structures that require both calcium and ROS at their tip for its elongation. If AtMC4 or any of the other calcium dependent metacaspases is involved in this process remains to be studied with the aid of loss of function mutants and activated versions, like overactive AtMC4 (Zhu et al., 2020). We observed AtMC4 localization in nuclear and cytosolic fractions by confocal imaging. However, it could be that fluorescent-tagged versions of AtMC4 undergo partial cleavage with the corresponding free movement of the fluorescent protein within the cell. Other controls using the inactive version of AtMC4 could serve as additional proof to validate localization. Alternatively, confirmation of localization using a specific AtMC4 antibody by the usage of electron microscopy in plants could work as a definite method to confirm AtMC4 localization.

In our hands, no obvious phenotypes for AtMC4 were detected when plants were grown under normal conditions. Possibly, *mc4* mutants show mild or no phenotypic differences under normal circumstances and these lines will only display differences when primed with calcium induced stimuli. In other studies, AtMC4 mutants showed reduced responses to fungal toxin fumonisin B1 and *Pseudomonas syringae* (avrRpt2) when compared to wild type plants (Watanabe and Lam, 2011b). The causes of absent phenotype might be due to the redundancy in the metacaspase family, illustrated by the enhanced phenotypes of a quadruplet metacaspase mutant challenged with *Botrytis cinerea* infection (Shen et al., 2019). Additionally, in some organisms with sequence redundancy between the family members, the presence of a KO line with

partial expression of its mRNA can induce the expression of similar genes (El-Brolosy et al., 2019). By this compensatory mechanism, other type-II could be expressed when AtMC4 transcription is reduced and the other metacaspases may take over the functions of AtMC4. In this case, a higher order mutant will be required to determine the real contribution of the metacaspase family to plants in development and stress.

Yet, another explanation is that the resulting effects of AtMC4 absence are not detected visually and other approaches like ion leakage, cell death assays or phosphorylation assays, would be needed to detect a molecular phenotype. Following this idea of finding molecular phenotypes, we have performed a series of experiments aiming to identify molecular accomplices hinting to AtMC4 functions.

The transcriptional responses expected for genes involved in wounding and Pep1 signalling did not work as expected and this might be explained by different reasons. One of them is that, even if AtMC4 is the most abundantly expressed metacaspase in Arabidopsis, its function could be overruled by the action by other metacaspases. As enzymatic components, the activity rate of these proteins may be limited by their substrate affinity rather than by the amount of available protein. As such, calcium levels would be a limiting factor in the detection of metacaspases activity *in vivo*. A second reason could be the selection of sampling timepoints. Part of the transcriptional response expected to be obtained from Pep1 sensing, is generally measured at a later interval (4 to 8 hours) and therefore expression of the marker genes should be better evaluated at these conditions. Yet, another explanation could be due to the utilization of Ler and AtMC4 mutants in the Ler background. While most studies use the Col-0 ecotype, we predicted our responses based on wound and Pep1 induced responses studied in Col-0 (Dressano et al., 2020; Rich-Griffin et al., 2020). However transcriptional and stress responses in the Ler background could be different to the responses in other ecotypes like Col-0 that is the most used in research. In fact, *Arabidopsis thaliana* Ler plants contain unique genes (Zapata et al., 2016), and indels are frequently found in this background accumulating in disease resistance genes (Ziolkowski et al., 2009). This observation point to the fact that although Ler is a very useful resource to use when T-DNA lines are not available in Col-0 or corroborate the function of a gene in different backgrounds, known transcriptional responses should be better tested.

In this chapter, we combined proteomics techniques that allow the identification of substrates proteome and protein levels in wounding events, as the triggering mechanism of AtMC4 enzymatic activity. Previously we demonstrated that wounding works efficiently in the activation of AtMC4, which in turn cleaves PROPEP1 shortly after damage (Hander et al., 2019). In order to identify specific AtMC4 substrates we performed this study in wild type plants and insertion T-DNA lines for the AtMC4 gene.

In the N-terminomics studies, we detected a total number of 1,910 peptides, of which a third matched to starting or second position in the protein sequence. From the 1,132 were heavily labelled we observed detection of peptides after arginine and after methionine. Peptides matching to processing after methionine in a position different to the first in the protein than could be explained by methionine removal of an alternative splicing starting site. The N-terminal pattern revealed highly preference for peptides generated after an arginine. Two explanations might be given for this, the first is that proteases with arginylase activity are the dominant proteases in Arabidopsis roots, and the other is that truncated proteins after arginine would be more stable to degradation. The second hypothesis is less likely because the stability of the peptide would be conditioned by the amino acids in the peptidyl chain, not by the amino acids preceding the processed sequence. Thus, proteases cleaving after arginine are likely the predominant proteases in terms of enzymatic activity in Arabidopsis roots. When we analysed the cleavage events after arginine and lysine, considering these as possible metacaspases targeting sites we could distinguish two patterns. Substrates processed after arginine did not show requirements upstream but a histidine appear frequently position P2. For peptides processed after lysine, we could see a double basic motif upstream and preference for alanine in general at multiple positions after the cleavage site. Absence of arginine at close positions after cleavage could be explained by the trypsin action and the generation of small peptides that could not be recognized with peptide search parameter. Although the patterns resulting from compiling all the N-termini are not real proteins they could be used as synthetic minimal recognition sites to assess metacaspases activation by their cleavage.

We tested proteolytic cleavage for three substrate candidates, PYK10, NAI2 and GRP7 using TNT assays and cleavage of proteins in Arabidopsis protein extracts. PYK10 and NAI2 are two proteins in ER bodies with activity in stress and herbivory response (Nakazaki et al., 2019). How AtMC4 could modulate their activity is not clear.

For instance, PYK10 bears a KDEL ER retrieving signal and by cleavage, it might allow PYK10 to change its localization. We tested processing of RL38eI and Atx1, using a cell free system and validated its cleavage by AtMC4.

Further, we checked interaction *in vivo* with the protease inhibitor CYS1. Our preliminary results show that AtMC4 is likely able to cleave CYS1 as detected in the anti-GFP blot of co-expressed samples in *N. benthamiana*. Co-IP experiments resulted in detection of RFP-tagged AtMC4 bands at low intensity. Multiple questions arise in this respect, like if CYS1 can affect AtMC4 enzymatic function or if it guides to AtMC4 for degradation. We aimed to compare this interaction with the action of the other protease inhibitor, SERPIN1. We noticed that inactive versions of AtMC4 could not bind to SERPIN1 and that we only detected processed fragments of AtMC4 N-terminal part by Co-IP. We predict that the inhibitor binds exclusively to the p20 domain after processing. Thus, the AtMC4 zymogen needs to be activated by calcium and self-process itself to perform its activity which might be regulated by SERPIN1. We found hints that point to a possible induction of metacaspases through calcium in other stresses as cold in a system including transcriptional induction of SERPIN1 by low temperatures.

It remains essential to study how processing may affect these proteins by study of loss- and gain-of-function mutants in normal conditions and under damage. Validation of the cleavage in metacaspase deficient lines and with the use of inhibitors and with the addition of the calcium chelators EGTA or EDTA, reducing the activity of metacaspases are the strategies to follow in order to corroborate our candidates as metacaspase substrates.

The quantitative proteomic data suggests that substantial changes are detectable in in the Arabidopsis proteome and that those are not exclusive from the genotype but as rapid consequences of the wounding effects. Despite the traditional concept that protein abundance would not primarily change substantially, wounding shows to be a harsh stress capable of rapidly modifying the plant proteome. These changes are possibly related to other by fast events like calcium and ROS waves as rapid hormone conjugation that can affect protein stability like previously observed with JAZ modules.

These results could be pinpointing to AtMC4 as a protease directly or indirectly accounted for degradation of other proteins. For instance such regulation could take

place by AtMC4 cleaving other proteases which in response loses their capacity to process their downstream substrates diminishing their identification in quantitative proteomics approaches. Although this scenario is possible, our results would require stronger practical evidence and a bigger comprehension of proteolytic networks in plants. Moreover, we should take into consideration that omics techniques used to identify substrates, protein and peptide abundance was obtained from Arabidopsis root samples, while there is evidence that AtMC4 is also expressed in green tissues. Therefore we could not discard that part of the AtMC4 terminome, including other interesting substrates, are still undiscovered.

MATERIAL AND METHODS

Plant material

Arabidopsis thaliana plants Landsberg erecta (Ler) and *AtMC4* T-DNA mutant line (CSHL_GT7237) were grown in vertical plates containing solid 1/2 MS (no sugar was added to the medium). Arabidopsis seeds were sterilized using chlorine gas and placed for 3 days in the dark at 4ºC in sealed plates containing solid ½ Murashige-Skoog media. Plates were transferred to a growth room at 21ºC with 16/8 hours of light/dark cycle with a photo-intensity of 100 µmol · m^{-2} · s^{-1}. Roots of 10 days-old seedlings for Ler and *mc4 KO* lines were excised from the aerial tissue using a razor and directly frozen in liquid nitrogen (control) or subjected to physical damage (wounding) in four biological repeats rolling and applying pressure with a 50 ml tube to induce activation of AtMC4 (Fig. 3B). Damaged roots were incubated at room temperature for 5 minutes prior to being frozen in liquid nitrogen.

PCR and cloning

The individual DNA fragments were obtained by overlap extension PCR and the use of iProof polymerase (Bio-Rad) according to the manufacturer guide. Primers overlapping with internal sequences of the genes of interest and overhangs expanding to attB sites were designed. As template cDNA material extracted from Arabidopsis seedlings was used. The components for each reaction were done in a final volume of 50µl PCR as specified: 10µl iProof HF Master Mix, 1µl 10 mM dNTP, forward primer (0,5 µM final concentration), 1µl reverse primer (0,5 µM final concentration), 1µl DNA template (100-200 ng/µl), 1µl iProof High-Fidelity DNA Polymerase. The conditions for the PCR were: initial denaturation 98 ºC for 30 seconds, 30 cycles of denaturation 98 ºC for 10 seconds, annealing (54 ºC - 65 ºC) for 30 seconds, extension 72 ºC for 1-3 minutes (1kb / 30 seconds) and a final extension 72 ºC for 5 minutes. PCR products were kept in cold or directly run on a 1% agarose gel in 0,5xTBS and purified using GeneJET Gel Extraction Kit as indicated in the manufacturer guidelines.

For Gateway® BP cloning we used an adjusted volume for 50 ng of the linear PCR purified DNA product (2µl approximately), 2µl of the donor vector at a concentration of 50 ng/ml and 1µl of BP Gateway® recombinase. We adjusted with mQ H_2O for a final volume of 10µl and incubated overnight at 25 ºC. Next day, 1µl proteinase K was added and incubated at 37 ºC for 10 minutes to inactivate the clonase, and mixture

was incubated on ice before heat shock *E. coli* transformation. Bacterias were grown in their corresponding selection overnight and generally individual colonies were picked and assessed for correct size using colony PCR. Positive size colonies were plasmid isolated and an aliquot of the material was sent for Sanger sequencing using M13 primers. LR reaction was used using 2µl 50 ng/ml donor plasmid and 2µl 100 ng/ml destination vectors with 1µl LR clonase at 25 ºC overnight. Transformation was performed as described above and colony PCR was performed using destination specific primers. Promoter fusions of SERPIN1 pSER1:GFP lines (pSER1:GUS-GFP) were previously published (Ghorbani et al., 2016).

RT-Quantitative PCR Gene Expression Assays

Total RNA of 10 days-old seedlings was extracted with TRIzol reagent and further purified with an RNeasy Kit (Qiagen) according to the manufacturer's instructions. The amount of 1 µg of total RNA per sample was used for cDNA synthesis using iScript™ cDNA Synthesis Kit (Bio-Rad) as described by the manufacturer. RT-qPCR solution was pipetted with the help of a JANUS Liquid Handler Workstation (PerkinElmer) in a 384 well plate with 200 nM concentration of each primer, 10 times diluted cDNA and 50% (v/v) SYBRGreen mix (Invitrogen) in a final volume of 5 µl. RT-qPCR values were detected in a LightCycler® 480 Instrument II (Roche) and results were analyzed using qbase+ (Biogazelle).

Protocol for Multi-Proteomics (N-terminal, Quantitative proteomics and Peptidomics)

Total Protein Extraction From Arabidopsis Thaliana

Arabidosis tissue was collected, frozen in liquid nitrogen and ground into a fine powder with mortar and pestle. Approximately 0.5 g of frozen ground tissue was resuspended in 0.9 ml of pre-chilled (4 °C) proteome extraction buffer (50 mM HEPES pH7.5, 10% glycerol, 5 mM EDTA, 5 mM EGTA, 0.01%(w/v) CHAPS, 0.5% (w/v) deoxycholate, and fresh protease inhibitor tablet) and allowed to thaw on ice before transferring to a fresh tube. Samples were centrifuged for 10 minutes at 16,000×g at 4˚C, carefully aspirated the supernatant (0,8 ml) and transferred to a fresh tube. The previous step was repeated to remove any remaining debris. Total protein concentration was

measured using a DC BioRad kit. Solid guanidinium hydrochloride was added to the cleared lysate to a final concentration of 4 M (380 mg in 0.8 ml). Total sample volume is divided in equal parts (500µl and 500µl approximately) for proteomics (shotgun and SCX) and peptidomics analysis (trypsin and undigested peptidomics). Store the proteome extracts in -20 °C and preferably transferred on dry ice.

Reduction And Alkylation Of The Proteomes

A volume of 13µl of freshly prepared 570 mM TCEP HCl was added to each 0.5 ml of the proteomics (shotgun and SCX) samples to obtain a final concentration of 15 mM TCEP (pH 8). Following, 20µl of freshly prepared 750 mM iodoacetamide (pH 8) was added to each proteome sample to obtain a final concentration of 30 mM and incubated for 30 minutes at 37 °C in the dark while shaking at 400-600 rpm. Samples were spinned down for 2 minutes at 16,000×g at 4˚C. Desalting of the proteome mixtures was performed over NAP™-5 columns (GE Healthcare Life Sciences) as it follows: equilibration of the columns 3 times with with 3 ml of 2 M guanidinium hydrochloride in 100 mM sodium phosphate buffer (pH 8). Application of the 0.5 ml of each sample through the column and elution in 1 ml of 2 M guanidinium hydrochloride in 100 mM sodium phosphate buffer (pH 8). 200µl were placed aside for shotgun proteomics analysis, while 800µl used for the N-terminal labelling.

N-Terminal Labelling Of The Proteomes

Prepare fresh NHS-esters by dissolving 0.45 mg of NHS-(trideutero-)acetate per sample. To achieve this we added 10µl to each sample of stock NHS-ester solution (3.6 mg in 80µl 50% acetonitrile for 8 samples). The total amount of labelling reagents added suffices to label the equivalent of up to 1 mg of protein material. Samples were incubated for 1h at 30 °C while shaking at 400-600 rpm. Repeat addition of NHS ester and incubate 2 times. Add 10µl of 1 M of glycine to quench the non-reacted N-hydroxysuccinimide esters. This is twice the molar excess of glycine over the NHS ester and it was incubated for 10 minutes at 30 °C while shaking at 400-600 rpm. Add 1µl of hydroxylamine (16,95 µmol/µl). This is four times the molar excess of hydroxylamine over the NHS ester. Incubate for 10 minutes at 30 °C to revert possible O-acetylation of Ser, Thr, or Tyr residues. No longer, otherwise chemical cleavage of

proteins by hydroxylamine might occur. NAP™-10 columns (GE Healthcare Life Sciences) were pre-equilibrated by washing 3 times 5 ml of freshly prepared 10 mM ammonium bicarbonate. Desalt the protein mixtures over in 1,5 mL ammonium bicarbonate and collect in an eppendorf. Using 3-5 replicates per sample is advised for accurate concentration measurement. Use 50µl for 3 replicates, V_F=1,35 ml. This concentration serves us to estimate the amount of trypsin needed for digestion and the volume of trypsin (0,2 µg/µl) to be added in the proteome samples.

Trypsin Digestion

Protein mixture was incubated for 5 minutes at 95 °C and transferred immediately to ice for 5 minutes incubation. Add trypsin (Sequencing-grade modified trypsin from Promega) to an enzyme/substrate ratio of 1/50 (µg/µg total protein). Trypsin was dissolved in 10 mM ammonium bicarbonate prior to adding the required volume. No switching to other digestion enzymes has been done due the simple fact that trypsin exerts a strong activity and serves its purpose in this method. Incubation of the mixture was done overnight at 37 °C while shaking at 400-600 rpm. Peptide mixture was centrifuged for 10 minutes at 16000×g, collect the supernatant, and vacuum dried.
Pyroglutamic Acid Removal

25µl aliquot of purified pGAPase (625 mU) was activated by adding 1µl 800 mM NaCl, 1µl 50 mM EDTA, and 11µl of fresh prepared 50 mM cysteamine following incubation at 37 °C for 10 minutes while shaking 400-600 rpm. The proteome pellet was re-dissolved in a 212µl freshly made pyroglutamate buffer (16 mM NaCl, 0.5 mM EDTA, 3 mM cysteamine, 50 µM aprotinin). Trypsin inhibition by aprotinin prevents inactivation of pGAPase and glutamine cyclotransferase (Q-cyclase). Activated pGAPase (625 mU) was added to the sample with 25µl of Q-cyclase (1250 mU) and incubated for 1h at 37 °C while shaking 400-600 rpm.

Strong Cation Exchange (SCX) For Pre-Enrichment Of N-Terminal Peptides

Initially we prepared a buffer A (10mM ammonium dihydrogen phosphate, 17.5% ACN pH3) and buffer B (11.5 mM phosphoric acid, 50%ACN, pH3). The proteome samples were acidified to pH 3 by adding 1% (v/v) trifluoroacetic acid (100µl). Samples were centrifuged for 15 minutes at 16,000×g at room temperature. Pre-wetting of the SCX column (Si-C8-SCX Agilent Bond Elut Cat no. 1202145) was done with 1 ml 100%

acetonitrile, followed by 1 ml water and then pre-equilibrated the column with 2 ml buffer A, the flow-through was then discarded. Using a vacuum manifold speeds up the process. A maximum of 1 ml sample was loaded to avoid any pellet debris and the flow-through was collected immediately in a 2 ml tube by gravity. Avoiding pump usage at this point allows peptides more time to contact to bind to the column. Washing was performed with 1 ml buffer A and collection of wash together with the previous flow through sample. Elution was performed by additionof 1 ml buffer B for 6 times, collecting the flow-through in a labelled same falcon tube. The pH of each flow-through sample was adjusted 6 by adding approximately 12-15µl of 5M NaOH to avoid extreme pH values while drying. 1 ml of the elution (100 µg) was taken apart for the HPLC fractionation sample. The sample was vacuum-centrifuged until it was completely dried and stored at -20 °C until further use.

Sorting N-Terminal Repeats

The dried-pelleted sample was redissolved using 80µl of the HPLC buffer 10 mM ammonium acetate 2%ACN pH 5.5. 4µl of acetic acid was added in order to enhance the solubility of low soluble positive charged peptides. Centrifugation of the peptide mixture for 10 minutes at 16,000×g at room temperature and transfer the supernatant to an HPLC vial was followed by addition of 20µl of 3% (w/w) hydrogen peroxide in water and incubation of the sample for 30 minutes at 30°C to convert methione into its stable sulfoxide derivative

Pellets were Re-dissolved in 50µl 50mM sodium borate pH 9.5 and 10µl freshly made 15mM TNBS in 50mM sodium borate pH 9.5 was added to each fraction and incubated for 30 minutes at 37 °C in the dark. The addition of TNBS and incubation was repeated for 3 times. Each TNBS fraction was acidified with 4µl acetic acid and centrifuged for 5 minutes at 16,000×g. A secondary HPLC run was performed on every fraction. Fractions are pooled automatically by the HPLC fraction collector with a 12 minutes time interval.

Quantitative Proteomics

A calculated volume for a total protein amount of 50 µg was pipetted by using data from the sample protein concentration measurements. The volume was adjusted to 120µl by adding TEAB 50 mM. Buffer was exchanged to TEAB using Thermo Scientific

Protein Desalting Spin Columns according to manufacture protocol. Desalted samples were heated for 2 minutes at 95 ºC and placed immediately on ice. 5µl of trypsin 0.2 µg/µl was added to each sample and incubated overnight at 37ºC. The next day samples were acidified to pH 3 using 0.5 % TFA, spinned down for 5 minutes at full speed and the supernatant transferred to a mass spec vial and stored at -20 ºC.

Mass Spectrometry

MS samples were dissolved in a 20µl MS loading buffer (2% ACN, 0.1% TFA) and 2.5 µl injected in a nanoHPLC Dionex 3000 system, connected to the Orbitrap XL MS.

Peptide Identification And Quantification

Raw files were converted to MGF peak lists using RawConverter (He et al., 2015). Resulting MGF files were searched by the COMET engine built-in the Crux toolkit (McIlwain et al., 2014).For N-terminal labeled samples, static modifications were carbamidomethylation (+57.02146 Da) on cysteine and heavy acetylation (+47.03606 Da) on lysine residues. Variable modifications inclucded methionine oxidation and N-terminal acetylation (in vivo, +42.01056 Da), heavy acetylation (+47.03606 Da) or pyroglutamate formation on N-terminal glutamine (-17.02655 Da). In case of quantitative shotgun samples, no heavy acetylation on lysine residues or peptide N-termini were specified as modification. The default setting of a concatenated decoy database search was used (shuffled peptides) and target-decoy FDR scoring was applied by the ‘assign-confidence’ function. The output files were converted to the moFF input format using an in-house python script. Afterwards, moFF (Argentini et al., 2016) was run on default settings, performing matching-between-runs for all files and extracting MS1 APEX intensities. The moFF results file were loaded in MsqRob (Goeminne et al., 2016) using the importMSnSet function. For peptide-level analysis, the protein was defined as the protein combined with the respective cleavage site position – thus performing differential statistics for cleavage sites. For protein-level analysis, peptide intensities are used to fit a model to determine differential protein abundance. Preprocessing with the function preprocess_MSnSet included $\log_2$ transformation (logtransform=TRUE, base=2) and quantile normalization

(normalisation="quantiles"). Finally, the data was converted to a protdata object with the MSnSet2protdata function and a robust ridge regression model with empirical Bayes variance estimation was fitted for each protein with genotype as fixed factor, and peptide sequence, biological replicate and sample as random effects using the fit.model function. The contrast of interest was estimated with the test.contrast_adjust function that also corrects for multiple testing. The false discovery rate was controlled at the 5% level.

Transient expression in *N. benthamiana* and protein extraction

Agrobacterium cultures were incubated for 2-3 days at 28 ° C in liquid YEB with agitation with their corresponding resistance. The amount of culture corresponding to an OD_{600} of 1.5 was calculated for a final volume 10 ml of combined culture and centrifuged for 10 minutes at 4000 rpm. Pellets were resuspended in the calculated volume of the infiltration buffer (10 mM $MgCl_2$, 10 mM 2- (Nmorpholino) ethanesulfonic acid (MES, pH 5.6), 100 µM acetosyringone). After 2 h incubation with shaking at 28 °C plants were infiltrated. For co-infiltration, equal volumes of agrobacterium with equal OD_{600} were mixed and infiltrated in the same leaf area. Abaxial part of 4to 6 weeks old *N. benthamiana* leaves were infiltrated with a 1 ml syringe.

“Quick-protein extraction” protocol

Leaf tissue of *N. benthamiana* transiently transformed were ground using liquid nitrogen and the samples were stored at - 80ºC. An equal volume to the tissue sample of 2XLaemmli buffer with 5-10% BME was directly added in the containing eppendorf and heated for 10 minutes at 70ºC. *Arabidopsis thaliana* samples were ground in an eppendorf containing 2 metal beads (3 mm) in a homogenizer Retsch NM 400. Protein concentration was measured when necessary, using Bradford protein assay (Bradford, 1976). Samples were diluted with mQ H_2O and Laemmli buffer to obtain equal protein concentration.

Co-Immunoprecipitation protocol

Approximately 2 g fresh weight of transiently infiltrated *N. benthamiana* leaves 3 days after infiltration were used as input material. The tissue was deep frozen and stored in -70 °C or directly ground in liquid nitrogen using a mortar and a pestle. Fresh extraction buffer (150mM Tris-HCl pH 7.5; 150 mM NaCl; 10 % glycerol; 10 mM EDTA; 1mM sodium molybdate; 1 mM NaF; 10 mM DTT; 0.5 % (w/v) PVPP; 1 % (v/v) protease

inhibitor cocktail (P9599, Sigma (1 tablet/10ml EB); 1 % (v/v) NP-40) was added to the input material in a ratio of 2 ml/ g FW following incubation on ice for 15 minutes. The material was centrifuged at 4˚C 13000 rpm for 20 mins and supernatant was cleared through a 40µm cell filter or centrifuge at 4˚C 13000 rpm, 20 mins again. An aliquot was taken as input material. 20µl of GFP-Trap Magnetic beads (Chromotek) previously washed for 3 times with the extraction buffer were incubated with the protein extract for 2 hours at 4 °C. Beads were trapped with the help of a magnetic rack and wash and resuspended 3 times with fresh washing buffer (20 mM Tris-HCl pH 7.5; 150 mM NaCl; 0.5 % (v/v) NP-40). Elution from the beads was performed using 50µl of pre-warmed Laemmli buffer with 5% β-mercaptoethanol. The elution material was directly run on a SDS-PAGE.

Immunoblotting

Protein samples were loaded in a 12,5% SDS-PAGE gels in Tris-Glycine-SDS (TGS) buffer at a voltage of 200 V for 45 minutes approximately. Gels were transferred to a precast TransBlot Turbo membrane and transferred for 3 minutes at 2.5 amperes constant up to 25 V in a BioRad transfering system. Membranes were blocked in 5% skim milk in PBS-T (0,2% Tween) for 2 hours. The blotting of GFP fluorescent tagged proteins was done by 2 hours incubation in 1% skim milk with single step antibody (1:1,000) rabbit anti-GFP conjugated to HRP (Miltenyi) followed 3 washing steps for 10 minutes in PBS-T and detection. RFP tagged proteins were blotted by overnight incubation at 4 ºC with mouse anti-RFP 6G6 (1:2,000; Chromotek) in 1% PBS-T and anti mouse conjugated to HRP (1;10,000; Sigma) for at least 2 hours at room temperature. Detection was performed by exposure to a solution of developing and fixation agents (1:1) for 1 minute. In the cases membranes were stripped we incubated the membrane for 10 minutes with a fresh solution on 10 mM Glycine pH 2.7 and SDS 10% (1:1), following 6 times washings for 5 minutes and re-starting by blocking for 2 hours using 5% skim milk PBS-T solution. Primary antibodies for C-term of NAI2 and PYK10 were obtained from Agrisera and incubated as described for GFP immunoblots.

Cell-free synthesis of radioactively labelled substrates and proteolysis assay

Recombinant MC4 (rMC4) and mutated MC4 $C^{139}A$ fused to a His-tag in pDEST17 vector were expressed and purified from *E. coli* as previously described in chapter 3

and stored in 50% glycerol, 25 mM HEPES (pH 7,5). Candidate substrate proteins RBL38eI, ATX1 were fused to GST tags using GATEWAY cloning system (pDEST15 or pDEST24) and GRP7, NAI2 and PYK10 plasmids were obtained from RIKEN collection. TnT-protease assays were performed as previously described (Minina et al., 2014). CDSs were in vitro transcribed and translated using rabbit reticulocyte TNT® coupled transcription/translation system (Promega) in the presence of radiolabeled ^{35}S-methionine. Aliquots of this reaction were subsequently mixed and incubated with the indicated amounts of rMC4 or the inactive rMC4C^{139}A protease for 30 minutes at 30 °C in optimal MC4 reaction buffer (50 mM HEPES pH7,5; 150 mM NaCl; 10% glycerol (w/v); 50 mM $CaCl_2$ and 10 mM DTT). The enzymatic reaction was stopped by the addition of a Laemmli buffer (supplemented with 50 mM EGTA to avoid aberrant SDS-PAGE electrophoresis due to high levels of Ca^{2+} ions in the samples). Samples were separated on SDS-PAGE and visualized with storage phosphor screens.

Confocal fluorescence imaging

Arabidopsis seedlings grown for 10 to 15 days-old and *N. benthamiana* leaf disks were imaged in a Zeiss LSM-710 confocal microscope. GFP was excited at 488 nm; mCherry and RFP were excited at 561nm. For membrane staining, Arabidopsis seedling were incubated for 10 minutes prior to imaging in a concentration of 10 µM propidium iodide (PI) in ddH_2O and the samples were directly mounted in a cover slide and a cover slip avoiding damaging the plant tissues. Images showing analogous and comparable lines were imaged using the same settings between samples, in order to be able to compare results. Image analysis editing and addition of size bars was performed using ImageJ software.

Calcium imaging

To image calcium dynamics we used a Nikon Ti microscope with Perfect Focus System (PFSIII) for Z-drift compensation, equipped with an Ultraview spinning-disk system (PerkinElmer). 10 days-old Arabidopsis thaliana seedlings grown in vertical plates and expressing the calcium reporter GCaMP3 (Toyota et al., 2018) were used for imaging using a 488 nm light source excitation and an emission window between 500 and 530 nm in single-camera mode. For elicitor treatment, self-made metal chambers were used and dissolved elicitor at the corresponding concentration was added at the desired timepoints. For cold dynamics imaging a CherryTemp system as

described in (Wang et al., 2020)was used to induce cold changes from 21°C to 12°C in real time while imaging.

Acknowledgements

We thank Dr. Jonathan Michael Dragwidge for assistance in the calcium imaging with the CherryTemp system. We thank Iris de Jaeger for assistance in the TnT assays. We thank Prof. Simon Gilroy for sharing the *Arabidopsis thaliana* seedlings expressing GCaMP3.

Supplementary Data

Online supplementary material can be accessed at https://osf.io/6v2p5/

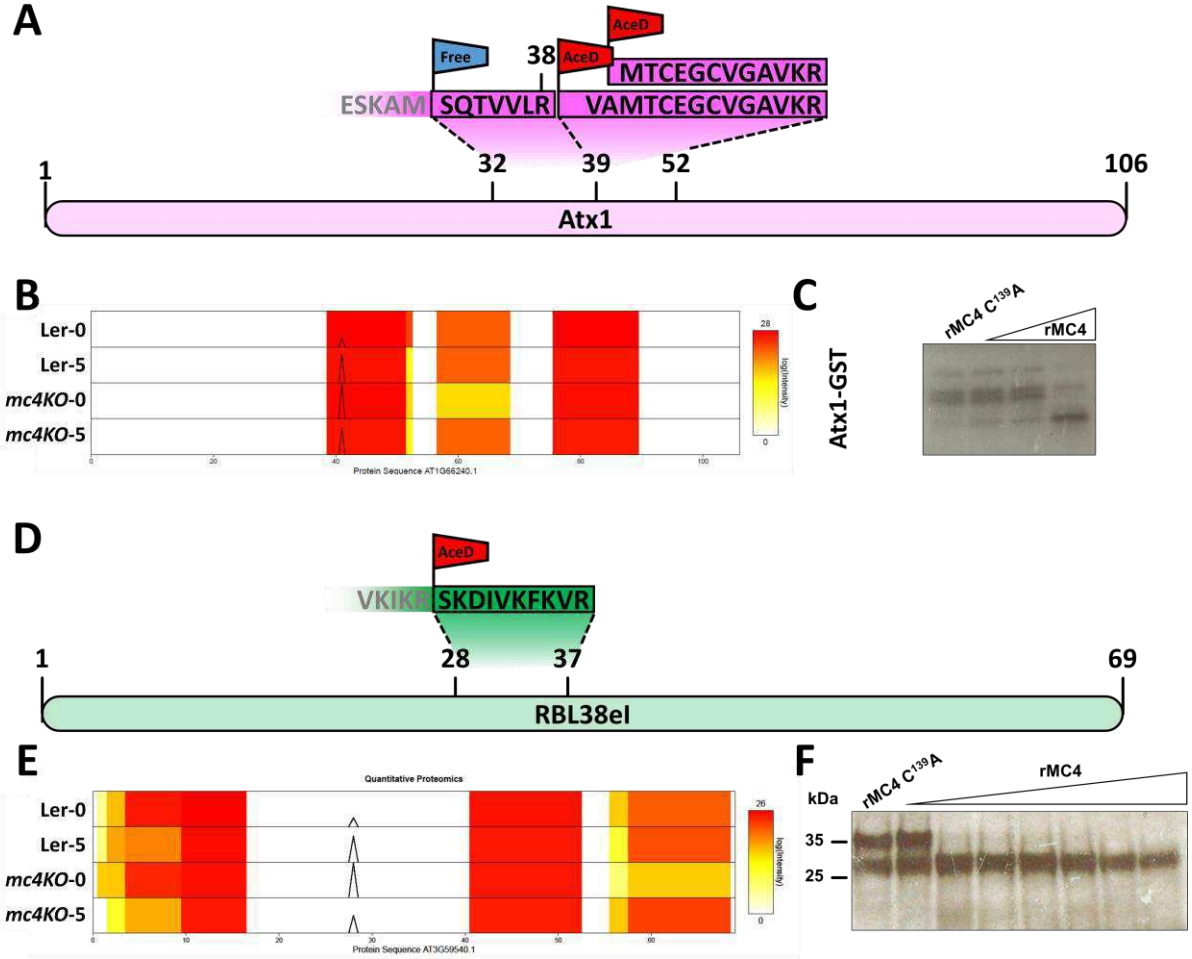
Peptide	AGI	Description	Pred. Localization	P5-P5'		PSMs	2-way ANOVA			
				p5	p5'		est	se	pval	qval
IKVVAPPER	AT1G49240.1(329) AT2G37620.1(329) AT2G42100.1(330) AT3G12110.1(329) AT3G18780.2(329) AT3G46520.1(329) AT3G53750.1(329) AT5G09810.1(329) AT5G59370.1(329)	actin 8 (ACT8)	C	PSSMK, PPSMK	IKVVA	1	-8.417	1.811	0.001	0.482
TVIKEGLEASQR	AT3G15950.1(640)	NAI2	ER	LAKYK	TVIKE	4	-2.471	0.557	0.001	0.524
SKVEIIKEKSNFIR	AT5G04590.1(75)	sulfite reductase (SIR)	P	TATKR	SKVEI	1	2.577	0.706	0.003	0.543
SPPIDEVIKAGVIPR	AT1G09270.1(112)	importin alpha isoform 4 (IMPA-4)	N	LSIER	SPPID	1	3.594	1.047	0.004	0.543
SQGPDLLGKTIEER	AT2G30870.1(79)	glutathione S-transferase PHI 10 (GSTF10)	C	AEKYR	SQGPD	9	-2.578	0.741	0.004	0.543
VSKVEVGTR	AT2G21160.1(222)	Translocon-associated protein (TRAP), alpha subunit	ER	KKTKK	VSKVE	2	2.527	0.679	0.005	0.543
SCPGIILALPILGITIGR	AT2G30490.1(446)	cinnamate-4-hydroxylase (C4H)	ER	GVGRR	SCPGI	1	3.866	0.864	0.005	0.543
LNSGKVIAVGPGSR	AT1G14980.1(36)	chaperonin 10 (CPN10)	M	EKSSK	LNSGK	1	4.062	0.921	0.005	0.557
GEVEVAKEAPAGVAR	AT3G10020.1(125)	unknown protein (HUP26)	N	VELVK	GEVEV	1	2.991	0.827	0.010	0.859
SEVVGYMGDDNLAKALE GADLVIIPAGVPR	AT1G53240.1(80)	Lactate/malate dehydrogenase protein (mMDH1)	M	HINTR	SEVVG	1	1.701	0.601	0.015	1.000
SIAQVLTVISQKQKSALR	AT2G39390.1(53) AT5G02610.2(76)	Ribosomal L29 family protein (RBL39)	C C	KVVRK	SIAQV	1	2.412	0.759	0.019	1.000
LAATCATAVESR	AT3G14990.1(135)	Class I glutamine amidotransferase-like superfamily protein (AtDJ1A)	G	VFMEK	LAATC	4	2.229	0.851	0.021	1.000
SFKAIELGSEETAR	AT1G54990.1(217)	AUXIN RESISTANT 4 (AXR4)	ER	NSKRR	SFKAI	1	-1.549	0.598	0.023	1.000
SQYIKEATELGKKAR	AT3G08930.1(250) AT5G01460.1(250)	LMBR1-like membrane protein	PM PM	AVITR	SQYIK	2	-2.231	0.899	0.029	1.000
VTANVVVAKDGTGKFKT VNEAVAAAPENSNTR	AT1G11580.2(264)	methylesterase PCR A (PMEPCRA)	Ex	PKTLK	VTANV	5	2.150	0.897	0.030	1.000
KTLKTDGIAGLYR	AT3G08580.1(238)	ADP/ATP carrier 1 (AAC1)	M	VDVYR	KTLKT	2	-1.552	0.627	0.032	1.000

Table 1. Significant heavy labelled N-terminal peptides identified and processed after arginine or lysine (p value < 0.05). Peptide sequence, AGI codes matching to each peptide identification cleavage site between brackets, protein description predicted localization (Prediction of localization were done using SUBA and using the consensus results https://suba.plantenergy.uwa.edu.au/. C: cytoplasm; ER: Endoplasmic Reticulum; Ex: Extracellular G: Golgi; M: Mitochondrion; N: Nucleus; P: plastid; PM: Plasma Membrane), P5-P5'motif, number of peptides identified, estimate and p-value, for each peptide is shown.

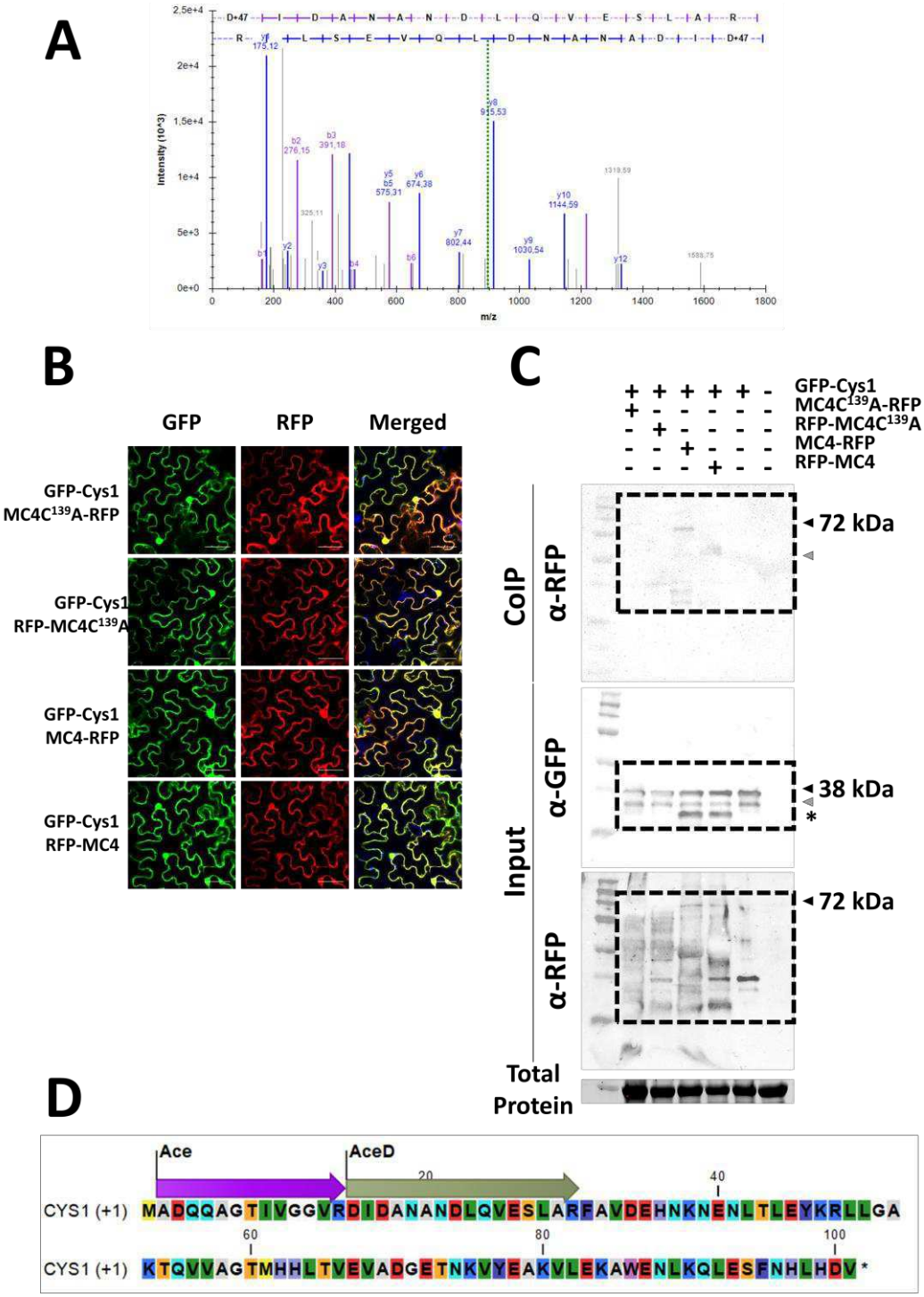
Oligo name	Sequence	Objective
attB1-MC4-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGACGAAAAAG GCGGTGCTT	Cloning
attB2-AtMc4_ns-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTGACAGATGAAAGGA GCGTTGGC	Cloning
attB2-AtMc4_s-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTTCAACAGATGAAAG GAGCGTTGGC	Cloning
GG_PAtMC4_A_FW	TTTTGGTCTCCACCTAGTTTCGACGATTTCAAAATTT	Cloning
GG_PAtMC4_B_RV	TTTTGGTCTCCTGTTGAGAAAAAATGATCTGATCGA	Cloning
qPCR-DREB-RV	ACGCTGGTACTGTGGACGATTC	RT-qPCR
qPCR-DREB-FW	ACGCTTTGATCGTTCCTCTTCTGG	RT-qPCR
qPCR-ARM-1-RV	ACGCTGGTACTGTGGACGATTC	RT-qPCR
qPCR-ARM-1-FW	ACGCTTTGATCGTTCCTCTTCTGG	RT-qPCR
qPCR-DRF-1-FW	AGAACTGAGTCGCTTGGAGACG	RT-qPCR
qPCR-DRF-1-RV	TGCTGCGAAGCTCTGAGGTATG	RT-qPCR
qPCR-NAC-1-FW	AAGCGCCGGGAAAGGAAACAAG	RT-qPCR
qPCR-NAC-1-RV	TCCGGCATAACGTCCATACTTCCG	RT-qPCR
qPCR-ATMKK9-1-FW	ACCTGTCGTCGGTGAAGTATCG	RT-qPCR
qPCR-ATMKK9-1-RV	TTACACCGCCGCGTAGTGATTC	RT-qPCR
attB1-GRP7-FW	GGGGACAACTTTGTATAATAAAGTTGGAATGGCGTCCGGTGAT GTTGAG	Cloning
attB2-GRP7ns-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTAACCCTTAATGGCAT CCTTCATTGA	Cloning
attB2-GRP7s-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTATTAACCCTTAATGG CATCCTTCATTGA	Cloning
RL38eI-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGCCTAAGCAA ATCCACGAG	Cloning
RL38eI_ns-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTGAAGGTCTTGTACA CTCAAACC	Cloning
RL38eI_s-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTTCAAAGGTCTTGTAC ACTCAAACC	Cloning
AtATX1-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGCTTAAAGAC TTGTTCCAA	Cloning
AtATX1_ns-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTGAGCCTTAGCAGTT TCACCT	Cloning
AtATX1_s-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTTTAAGCCTTAGCAGT TTCACCT	Cloning
attB1-CYS1-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCACCATGGCGGAT CAACAAGCAGGA	Cloning
attB2-CYS1ns-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTGAACATCGTGAAGG TGGTTGA	Cloning
attB2-CYS1s-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTTTAAACATCGTGAA GGTGGTTGA	Cloning
attB1-SER1-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCACCATGGACGTG CGTGAATC	Cloning
attB2-SER1ns-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTGATGCAACGGATCA ACAAC	Cloning
attB2-SER1s-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTTTAATGCAACGGAT CAACAAC	Cloning

Table 2. Primer list used in this study. Name of the primer, sequence of the primer (5' → 3'direction) and objective of the primer in this chapter.

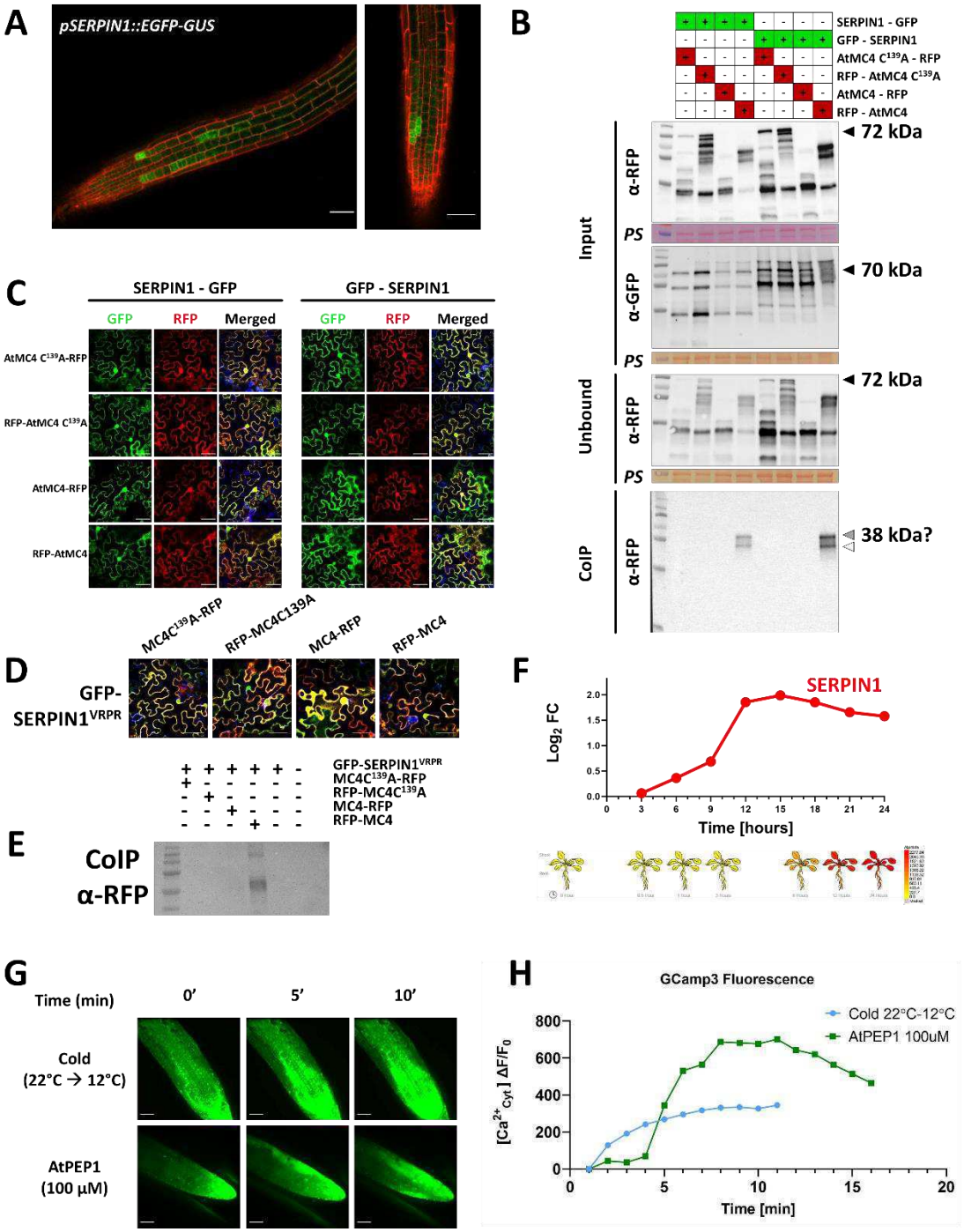
Suppl. figure 1. Plots for the significant (p values<0.05, 2-way ANOVA) of the heavy labelled peptides labelled after an arginine or a lysine show in Suppl. Table.1. Name abbreviations and AGI code are indicated for each gene. Plots showing overlap of the N-terminome detected peptides (black peaks) and peptide intensity identified from shotgun proteomics (yellow to red range). Selected proteins contain peptides that scored significantly (p<0.05) and that are processed after arginine of lysine. Each plot shows 4 conditions for wild type control (*WT 0'*), wild type wounding (*WT 5'*), mutant control (*mc4 0'*) and mutant wounding (*mc4 5'*).



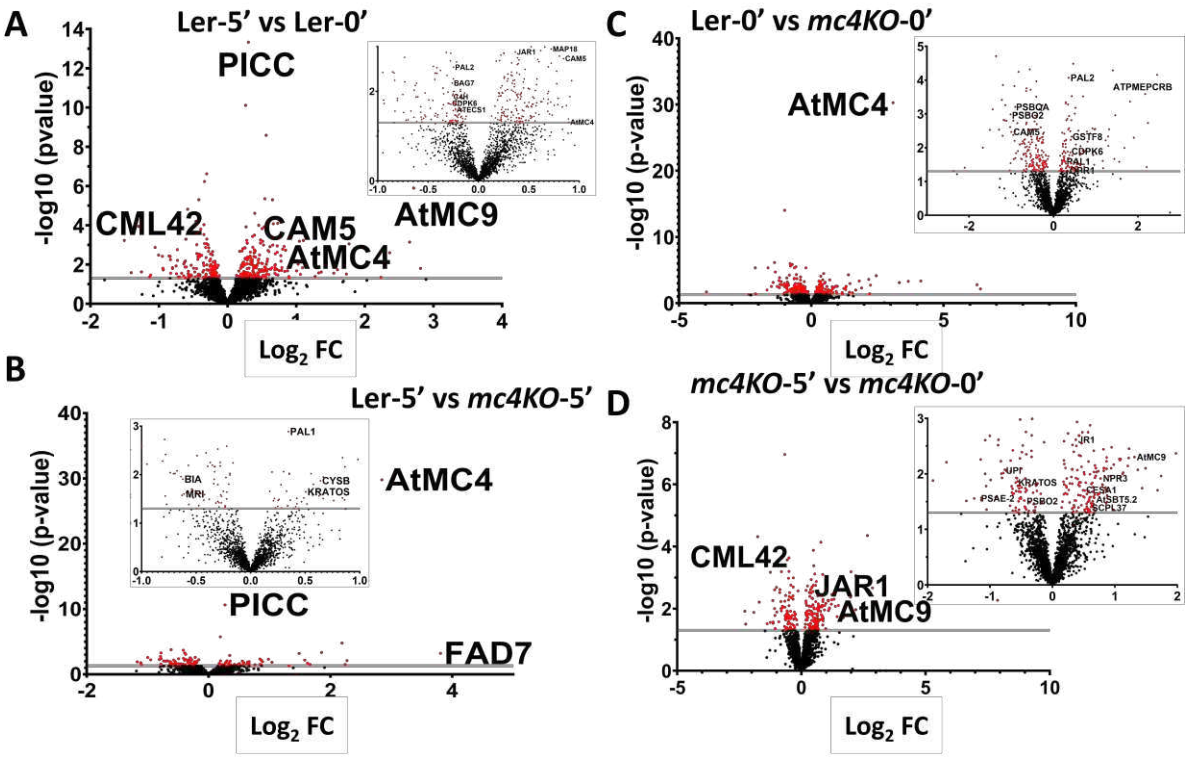
Suppl. figure 2. Identified putative substrates of AtMC4. Identified peptides in the degradome approach and their relative position for Atx1 **(A)** and RL38eI **(D)**. Plots showing intensities N terminal sites and the corresponding peptide abundances for Atx1 **(B)** and RL38eI **(E)**. TNT assays for in vitro test validating cleavage by AtMC4 for Atx1 **(C)** and RL38eI **(F)**.



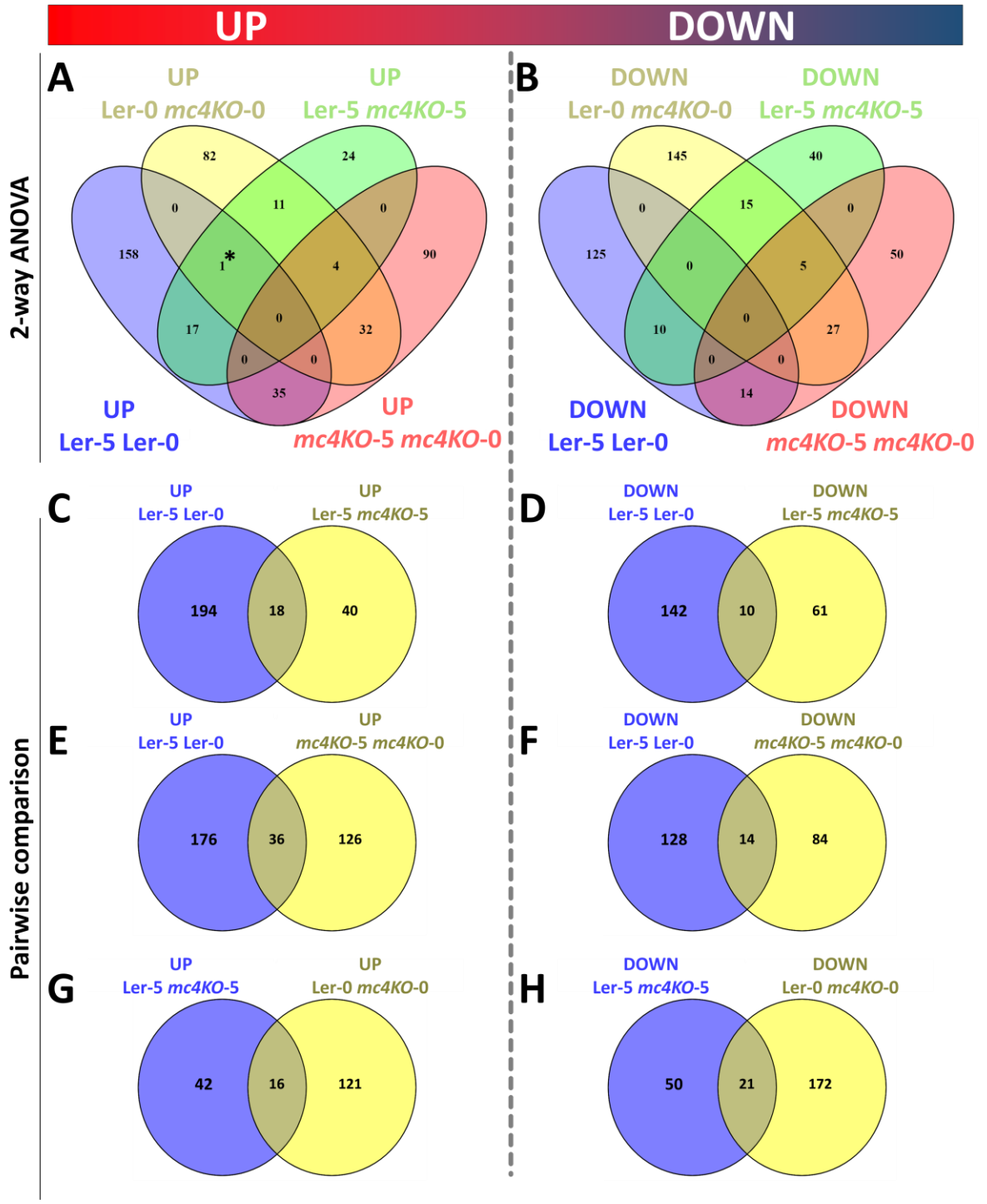
Suppl. figure 3. A) Annotated MS/ MS spectra for the heavy labelled peptide. **B)** Confocal images of co-expression of GFP-CYS1 and AtMC4 RFP tagged versions for transient infiltration in *N. benthamiana*. **C)** Western-Blots showing the detection of GFP-CYS1 and RFP tagged versions of AtMC4 used for Co-IP performed in *N. benthamiana* as well as the blots showing input sampling. Weak bands were detected in the Co-IP blot corresponding to high molecular size of AtMC4 (Higher blot, back arrow).In the input blot we detected a band corresponding to GFP-CYS1 (38 kDa) and a second band of similar size (grey filled arrow). A third band only when active AtMC4 was expressed (third and fourth lanes from the left) which may correspond to the size of GFP-CYS with the release of its C-terminal part (indicated by an asterisk). **D)** CYS1 sequence with the peptides annotation.



Suppl. figure 4. SERPIN1 interacts with the active form of AtMC4. A) Expression patterns of the line *pSERPIN1:GFP-GUS* in *A .thaliana* seedlings. **B)** Co-IP for *SERPIN1* fused to GFP C or N-terminally with the active and inactive versions of AtMC4. **C)** Confocal images of transiently expressed lines used for Co-IP. **D)** and **E)** Confocal images of Co-IP samples of the SERPIN1 mutated at its reactive centre loop VRPR (SERPIN1^{VRPR}) and Co-IP showing interaction of AtMC4 with the mutated SERPIN1^{VRPR}. **F)** Expression levels for *SERPIN1* in cold stresses (Data obtained from Genevestigator) and image representation (Obtained at www.thalemine.com). **G)** Images of calcium reporter GCaMP3 in *A. thaliana* seedlings undergoing elicitor or cold (21°C to 12°C) treatments and **H)** measurements of the calcium dynamics according to fluorescent ratios.



Suppl. figure 5. Complete view of the volcano plots showing protein abundance. Values calculated for each protein plotted considering their -log (p-value) and the log_2 FC. Significant values ($p<0.05$) above the significant threshold, displayed by the grey line, are coloured in red. Positive FDR values for each comparison means higher protein abundance while negative values represent lower abundance. All the points are displayed in the expanded version, while the values for the central region as shown in Fig. x are displayed in the squared boxes. Conditions plotted: **A)** Wounded wild type plants versus unwounded wild type plants (Ler-5 vs Ler-0). **B)** Unwounded wild type plants versus unwounded mutant plants (Ler-0 vs *mc4KO*-0) **C)** Wounded wild type plants versus wounded mutant plants (Ler-5 vs *mc4KO*-5) **D)** Wounded mutant plants versus unwounded mutant plants (*mc4KO*-5 vs *mc4KO*-0)



Suppl. figure 6. Venn diagram for common UP or common DOWN values obtained in the quantitative proteomics. We compared statistically significant values ($p<0.05$) between samples as indicated. Venn diagram showing the Comparison the four different conditions (Ler-5/Ler-0 vs Ler-0/*mc4KO*-0 vs Ler-5/*mc4KO*-5 vs *mc4KO*-5/*mc4KO*-0) for **A)** Upregulated values and **B)** Downregulated values. Venn diagrams showing the pairwise compares on the values for the Ler-5/Ler-0 vs Ler-5/*mc4KO*-5 for **C)** UP and **D)** DOWN values; Ler-5/Ler-0 vs *mc4KO*-5/*mc4KO*-0for **E)** UP and **F)** DOWN values; and Ler-5/*mc4KO*-5 vs Ler-0/*mc4KO*-0 for **G)** UP and **H)** DOWN values.

REFERENCES

Argentini, A., Goeminne, L.J., Verheggen, K., Hulstaert, N., Staes, A., Clement, L., and Martens, L. (2016). moFF: a robust and automated approach to extract peptide ion intensities. Nature methods **13:** 964-966.

Belenghi, B., Acconcia, F., Trovato, M., Perazzolli, M., Bocedi, A., Polticelli, F., Ascenzi, P., and Delledonne, M. (2003). AtCYS1, a cystatin from Arabidopsis thaliana, suppresses hypersensitive cell death. Eur J Biochem **270:** 2593-2604.

Bienvenut, W.V., Sumpton, D., Martinez, A., Lilla, S., Espagne, C., Meinnel, T., and Giglione, C. (2012). Comparative large scale characterization of plant versus mammal proteins reveals similar and idiosyncratic N-alpha-acetylation features. Mol Cell Proteomics **11:** M111 015131.

Boisson-Dernier, A., Franck, C.M., Lituiev, D.S., and Grossniklaus, U. (2015). Receptor-like cytoplasmic kinase MARIS functions downstream of CrRLK1L-dependent signaling during tip growth. Proc Natl Acad Sci U S A **112:** 12211-12216.

Bollhoner, B., Zhang, B., Stael, S., Denance, N., Overmyer, K., Goffner, D., Van Breusegem, F., and Tuominen, H. (2013). Post mortem function of AtMC9 in xylem vessel elements. New Phytol **200:** 498-510.

Browse, J. (2009). The power of mutants for investigating jasmonate biosynthesis and signaling. Phytochemistry **70:** 1539-1546.

Coll, N.S., Vercammen, D., Smidler, A., Clover, C., Van Breusegem, F., Dangl, J.L., and Epple, P. (2010). Arabidopsis type I metacaspases control cell death. Science **330:** 1393-1397.

Demir, F., Niedermaier, S., Villamor, J.G., and Huesgen, P.F. (2018). Quantitative proteomics in plant protease substrate identification. New Phytol **218:** 936-943.

Dressano, K., Weckwerth, P.R., Poretsky, E., Takahashi, Y., Villarreal, C., Shen, Z., Schroeder, J.I., Briggs, S.P., and Huffaker, A. (2020). Dynamic regulation of Pep-induced immunity through post-translational control of defence transcript splicing. Nat Plants **6:** 1008-1019.

El-Brolosy, M.A., et al. (2019). Genetic compensation triggered by mutant mRNA degradation. Nature **568:** 193-197.

Escamez, S., Stael, S., Vainonen, J.P., Willems, P., Jin, H., Kimura, S., Van Breusegem, F., Gevaert, K., Wrzaczek, M., and Tuominen, H. (2019). Extracellular peptide Kratos restricts cell death during vascular development and stress in Arabidopsis. J Exp Bot **70:** 2199-2210.

Ferrari, S., Plotnikova, J.M., De Lorenzo, G., and Ausubel, F.M. (2003). Arabidopsis local resistance to Botrytis cinerea involves salicylic acid and camalexin and requires EDS4 and PAD2, but not SID2, EDS5 or PAD4. Plant J **35:** 193-205.

Geem, K.R., Kim, D.H., Lee, D.W., Kwon, Y., Lee, J., Kim, J.H., and Hwang, I. (2019). Jasmonic acid-inducible TSA1 facilitates ER body formation. Plant J **97:** 267-280.

Ghorbani, S., Hoogewijs, K., Pecenkova, T., Fernandez, A., Inze, A., Eeckhout, D., Kawa, D., De Jaeger, G., Beeckman, T., Madder, A., Van Breusegem, F., and Hilson, P. (2016). The SBT6.1 subtilase processes the GOLVEN1 peptide controlling cell elongation. J Exp Bot **67:** 4877-4887.

Goeminne, L.J., Gevaert, K., and Clement, L. (2016). Peptide-level Robust Ridge Regression Improves Estimation, Sensitivity, and Specificity in Data-dependent Quantitative Label-free Shotgun Proteomics. Molecular & cellular proteomics : MCP **15:** 657-668.

Hander, T., et al. (2019). Damage on plants activates Ca(2+)-dependent metacaspases for release of immunomodulatory peptides. Science **363**.

He, L., Diedrich, J., Chu, Y.Y., and Yates, J.R., 3rd. (2015). Extracting Accurate Precursor Information for Tandem Mass Spectra by RawConverter. Analytical chemistry **87:** 11361-11367.

Hou, S., Zhang, J., and He, P. (2021). Stress-induced activation of receptor signaling by protease-mediated cleavage. Biochem J **478:** 1847-1852.

Hou, X., Tan, L., and Tang, S.F. (2019). Molecular mechanism study on the interactions of cadmium (II) ions with Arabidopsis thaliana glutathione transferase Phi8. Spectrochim Acta A Mol Biomol Spectrosc **216:** 411-417.

Kang, E., Zheng, M., Zhang, Y., Yuan, M., Yalovsky, S., Zhu, L., and Fu, Y. (2017). The Microtubule-Associated Protein MAP18 Affects ROP2 GTPase Activity during Root Hair Growth. Plant Physiol **174:** 202-222.

Kato, M., Aoyama, T., and Maeshima, M. (2013). The Ca(2+) -binding protein PCaP2 located on the plasma membrane is involved in root hair development as a possible signal transducer. Plant J **74:** 690-700.

Kleifeld, O., Doucet, A., auf dem Keller, U., Prudova, A., Schilling, O., Kainthan, R.K., Starr, A.E., Foster, L.J., Kizhakkedathu, J.N., and Overall, C.M. (2010). Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products. Nat Biotechnol **28:** 281-288.

Knight, H., Trewavas, A.J., and Knight, M.R. (1996). Cold calcium signaling in Arabidopsis involves two cellular pools and a change in calcium signature after acclimation. Plant Cell **8:** 489-503.

Kohler, D., Montandon, C., Hause, G., Majovsky, P., Kessler, F., Baginsky, S., and Agne, B. (2015). Characterization of chloroplast protein import without Tic56, a component of the 1-megadalton translocon at the inner envelope membrane of chloroplasts. Plant Physiol **167:** 972-990.

Lallemand, J., Bouche, F., Desiron, C., Stautemas, J., de Lemos Esteves, F., Perilleux, C., and Tocquin, P. (2015). Extracellular peptidase hunting for improvement of protein production in plant cells and roots. Front Plant Sci **6:** 37.

Larrieu, A., et al. (2015). A fluorescent hormone biosensor reveals the dynamics of jasmonate signalling in plants. Nat Commun **6:** 6043.

Lema Asqui, S., Vercammen, D., Serrano, I., Valls, M., Rivas, S., Van Breusegem, F., Conlon, F.L., Dangl, J.L., and Coll, N.S. (2018). AtSERPIN1 is an inhibitor of the metacaspase AtMC1-mediated cell death and autocatalytic processing in planta. New Phytol **218:** 1156-1166.

Li, Q., Ai, G., Shen, D., Zou, F., Wang, J., Bai, T., Chen, Y., Li, S., Zhang, M., Jing, M., and Dou, D. (2019). A Phytophthora capsici Effector Targets ACD11 Binding Partners that Regulate ROS-Mediated Defense Response in Arabidopsis. Mol Plant **12:** 565-581.

Linster, E., et al. (2015). Downregulation of N-terminal acetylation triggers ABA-mediated drought responses in Arabidopsis. Nat Commun **6:** 7640.

Martinez, A., Traverso, J.A., Valot, B., Ferro, M., Espagne, C., Ephritikhine, G., Zivy, M., Giglione, C., and Meinnel, T. (2008). Extent of N-terminal modifications in cytosolic proteins from eukaryotes. Proteomics **8:** 2809-2831.

McIlwain, S., Tamura, K., Kertesz-Farkas, A., Grant, C.E., Diament, B., Frewen, B., Howbert, J.J., Hoopmann, M.R., Kall, L., Eng, J.K., MacCoss, M.J., and Noble, W.S. (2014). Crux: rapid open source protein tandem mass spectrometry analysis. Journal of proteome research **13:** 4488-4491.

Meyer, K., Koster, T., Nolte, C., Weinholdt, C., Lewinski, M., Grosse, I., and Staiger, D. (2017). Adaptation of iCLIP to plants determines the binding landscape of the clock-regulated RNA-binding protein AtGRP7. Genome Biol **18:** 204.

Minina, E.A., Stael, S., Van Breusegem, F., and Bozhkov, P.V. (2014). Plant metacaspase activation and activity. Methods Mol Biol **1133:** 237-253.

Minina, E.A., Coll, N.S., Tuominen, H., and Bozhkov, P.V. (2017). Metacaspases versus caspases in development and cell fate regulation. Cell Death Differ **24:** 1314-1325.

Nagano, A.J., Matsushima, R., and Hara-Nishimura, I. (2005). Activation of an ER-body-localized beta-glucosidase via a cytosolic binding partner in damaged tissues of Arabidopsis thaliana. Plant Cell Physiol **46:** 1140-1148.

Nakazaki, A., Yamada, K., Kunieda, T., Sugiyama, R., Hirai, M.Y., Tamura, K., Hara-Nishimura, I., and Shimada, T. (2019). Leaf Endoplasmic Reticulum Bodies Identified in Arabidopsis Rosette Leaves Are Involved in Defense against Herbivory. Plant Physiol **179:** 1515-1524.

Perrar, A., Dissmeyer, N., and Huesgen, P.F. (2019). New beginnings and new ends: methods for large-scale characterization of protein termini and their use in plant biology. J Exp Bot **70:** 2021-2038.

Rask, L., Andreasson, E., Ekbom, B., Eriksson, S., Pontoppidan, B., and Meijer, J. (2000). Myrosinase: gene family evolution and herbivore defense in Brassicaceae. Plant Mol Biol **42:** 93-113.

Rich-Griffin, C., et al. (2020). Regulation of Cell Type-Specific Immunity Networks in Arabidopsis Roots. Plant Cell **32:** 2742-2762.

Rojas, B.E., Hartman, M.D., Figueroa, C.M., and Iglesias, A.A. (2020). Proteolytic cleavage of Arabidopsis thaliana phosphoenolpyruvate carboxykinase-1 modifies its allosteric regulation. J Exp Bot.

Schaller, F., Biesgen, C., Mussig, C., Altmann, T., and Weiler, E.W. (2000). 12-Oxophytodienoate reductase 3 (OPR3) is the isoenzyme involved in jasmonate biosynthesis. Planta **210:** 979-984.

Scholz, S.S., Vadassery, J., Heyer, M., Reichelt, M., Bender, K.W., Snedden, W.A., Boland, W., and Mithofer, A. (2014). Mutation of the Arabidopsis calmodulin-like protein CML37 deregulates the jasmonate pathway and enhances susceptibility to herbivory. Mol Plant **7:** 1712-1726.

Shen, W., Liu, J., and Li, J.F. (2019). Type-II Metacaspases Mediate the Processing of Plant Elicitor Peptides in Arabidopsis. Mol Plant **12:** 1524-1533.

Solis, N., and Overall, C.M. (2018). Identification of Protease Cleavage Sites and Substrates in Cancer by Carboxy-TAILS (C-TAILS). Methods Mol Biol **1731:** 15-28.

Staes, A., Van Damme, P., Helsens, K., Demol, H., Vandekerckhove, J., and Gevaert, K. (2008). Improved recovery of proteome-informative, protein N-terminal peptides by combined fractional diagonal chromatography (COFRADIC). Proteomics **8:** 1362-1370.

Staes, A., Van Damme, P., Timmerman, E., Ruttens, B., Stes, E., Gevaert, K., and Impens, F. (2017). Protease Substrate Profiling by N-Terminal COFRADIC. Methods Mol Biol **1574:** 51-76.

Staswick, P.E., and Tiryaki, I. (2004). The oxylipin signal jasmonic acid is activated by an enzyme that conjugates it to isoleucine in Arabidopsis. Plant Cell **16:** 2117-2127.

Toyota, M., Spencer, D., Sawai-Toyota, S., Jiaqi, W., Zhang, T., Koo, A.J., Howe, G.A., and Gilroy, S. (2018). Glutamate triggers long-distance, calcium-based plant defense signaling. Science **361:** 1112-1115.

Tsiatsiani, L., Van Breusegem, F., Gallois, P., Zavialov, A., Lam, E., and Bozhkov, P.V. (2011). Metacaspases. Cell Death Differ **18:** 1279-1288.

Tsiatsiani, L., Timmerman, E., De Bock, P.J., Vercammen, D., Stael, S., van de Cotte, B., Staes, A., Goethals, M., Beunens, T., Van Damme, P., Gevaert, K., and Van Breusegem, F. (2013). The Arabidopsis metacaspase9 degradome. Plant Cell **25:** 2831-2847.

Vadassery, J., Reichelt, M., Hause, B., Gershenzon, J., Boland, W., and Mithofer, A. (2012). CML42-mediated calcium signaling coordinates responses to Spodoptera herbivory and abiotic stresses in Arabidopsis. Plant Physiol **159:** 1159-1175.

van der Hoorn, R.A. (2008). Plant proteases: from phenotypes to molecular mechanisms. Annu Rev Plant Biol **59:** 191-223.

Venkatakrishnan, S., Mackey, D., and Meier, I. (2013). Functional investigation of the plant-specific long coiled-coil proteins PAMP-INDUCED COILED-COIL (PICC) and PICC-LIKE (PICL) in Arabidopsis thaliana. PLoS One **8:** e57283.

Venne, A.S., Solari, F.A., Faden, F., Paretti, T., Dissmeyer, N., and Zahedi, R.P. (2015). An improved workflow for quantitative N-terminal charge-based fractional diagonal chromatography (ChaFRADIC) to study proteolytic events in Arabidopsis thaliana. Proteomics **15:** 2458-2469.

Vercammen, D., van de Cotte, B., De Jaeger, G., Eeckhout, D., Casteels, P., Vandepoele, K., Vandenberghe, I., Van Beeumen, J., Inze, D., and Van Breusegem, F. (2004). Type II metacaspases Atmc4 and Atmc9 of Arabidopsis thaliana cleave substrates after arginine and lysine. J Biol Chem **279:** 45329-45336.

Vercammen, D., Belenghi, B., van de Cotte, B., Beunens, T., Gavigan, J.A., De Rycke, R., Brackenier, A., Inze, D., Harris, J.L., and Van Breusegem, F. (2006). Serpin1 of Arabidopsis thaliana is a suicide inhibitor for metacaspase 9. J Mol Biol **364:** 625-636.

Wang, J., Mylle, E., Johnson, A., Besbrugge, N., De Jaeger, G., Friml, J., Pleskot, R., and Van Damme, D. (2020). High Temporal Resolution Reveals Simultaneous Plasma Membrane Recruitment of TPLATE Complex Subunits. Plant Physiol **183:** 986-997.

Wang, Z., Li, X., Wang, X., Liu, N., Xu, B., Peng, Q., Guo, Z., Fan, B., Zhu, C., and Chen, Z. (2019). Arabidopsis Endoplasmic Reticulum-Localized UBAC2 Proteins Interact with PAMP-INDUCED COILED-COIL to Regulate Pathogen-Induced Callose Deposition and Plant Immunity. Plant Cell **31:** 153-171.

Watanabe, N., and Lam, E. (2011a). Calcium-dependent activation and autolysis of Arabidopsis metacaspase 2d. J Biol Chem **286:** 10027-10040.

Watanabe, N., and Lam, E. (2011b). Arabidopsis metacaspase 2d is a positive mediator of cell death induced during biotic and abiotic stresses. Plant J **66:** 969-982.

Wendrich, J.R., et al. (2020). Vascular transcription factors guide plant epidermal responses to limiting phosphate conditions. Science **370**.

Weng, S.S.H., Demir, F., Ergin, E.K., Dirnberger, S., Uzozie, A., Tuscher, D., Nierves, L., Tsui, J., Huesgen, P.F., and Lange, P.F. (2019). Sensitive Determination of Proteolytic Proteoforms in Limited Microscale Proteome Samples. Mol Cell Proteomics **18:** 2335-2347.

Willems, P., Horne, A., Van Parys, T., Goormachtig, S., De Smet, I., Botzki, A., Van Breusegem, F., and Gevaert, K. (2019). The Plant PTM Viewer, a central resource for exploring plant protein modifications. Plant J **99:** 752-762.

Wrzaczek, M., et al. (2015). GRIM REAPER peptide binds to receptor kinase PRK5 to trigger cell death in Arabidopsis. EMBO J **34:** 55-66.

Yang, D.H., Kwak, K.J., Kim, M.K., Park, S.J., Yang, K.Y., and Kang, H. (2014). Expression of Arabidopsis glycine-rich RNA-binding protein AtGRP2 or AtGRP7 improves grain yield of rice (Oryza sativa) under drought stress conditions. Plant Sci **214:** 106-112.

Zapata, L., Ding, J., Willing, E.M., Hartwig, B., Bezdan, D., Jiao, W.B., Patel, V., Velikkakam James, G., Koornneef, M., Ossowski, S., and Schneeberger, K. (2016). Chromosome-level assembly of Arabidopsis thaliana Ler reveals the extent of translocation and inversion polymorphisms. Proc Natl Acad Sci U S A **113:** E4052-4060.

Zhang, H., Deery, M.J., Gannon, L., Powers, S.J., Lilley, K.S., and Theodoulou, F.L. (2015). Quantitative proteomics analysis of the Arg/N-end rule pathway of targeted degradation in Arabidopsis roots. Proteomics **15:** 2447-2457.

Zhang, H., Gannon, L., Hassall, K.L., Deery, M.J., Gibbs, D.J., Holdsworth, M.J., van der Hoorn, R.A.L., Lilley, K.S., and Theodoulou, F.L. (2018). N-terminomics reveals control of Arabidopsis seed storage proteins and proteases by the Arg/N-end rule pathway. New Phytol **218:** 1106-1126.

Zhu, P., Yu, X.H., Wang, C., Zhang, Q., Liu, W., McSweeney, S., Shanklin, J., Lam, E., and Liu, Q. (2020). Structural basis for Ca(2+)-dependent activation of a plant metacaspase. Nat Commun **11:** 2249.

Ziolkowski, P.A., Koczyk, G., Galganski, L., and Sadowski, J. (2009). Genome sequence comparison of Col and Ler lines reveals the dynamic nature of Arabidopsis chromosomes. Nucleic Acids Res **37:** 3189-3201.

Chapter 6

"Genetically encoded fluorescent reporters for the detection of proteolysis in plants"

Álvaro Daniel Fernández-Fernández[1,2], Simon Stael[1,2,3,4] and Frank Van Breusegem[1,2]

[1] Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium

[2] Center for Plant Systems Biology, VIB, 9052 Ghent, Belgium

[3] VIB-UGent Center for Medical Biotechnology, Ghent 9052, Belgium.

[4] Department of Biomolecular Medicine, Ghent University, Ghent 9000, Belgium.

Author contributions:
ADFF, SS and FVB conceived and designed the experiments and project. ADFF performed the research, analyzed the data. ADFF wrote the manuscript with the help of SS and FVB.

Aim and context

Proteolysis is a widespread post-translational modification consisting of the hydrolysis of the peptidyl bonds within a protein. Every protein undergoes proteolysis during degradation, but specifically targeted proteolytic cleavages can also work to tune protein functions leading for example to changes in enzymatic activity or interaction partners. Despite the impact of proteases on cell signal transduction, many proteolysis processes, especially in plants, remain to be studied. The main bottlenecks in the protease field are the low number of substrates identified and the difficulty to understand and visualize the proteolytic activation mechanisms. While multiple tools have emerged in the last years to identify substrates using proteomics, imaging of the proteolytic activity *in vivo* is becoming a very attractive subfield to understand and characterize the protease studied. While an increasing number of biosensors have been used in plants to track redox conditions, ion fluxes, metabolites, hormones and enzymes like kinases, the use of proteolytic sensors is highly underexplored. We aimed to test some of the existing sensors, tested mainly in mammalian studies, for the detection of proteolysis in plants combining proof of concept and specific sensors for metacaspases.

Abstract:

Proteolytic enzymes are present in all types of organisms. Proteases can work as molecular switchers by irreversible processing of their substrates affecting their functions. However, proteolytic activity is not occurring permanently over time and its action is restricted to specific occasions and localized events. Taking into account proteases nature and their tightly regulation it is essential to understand the mode of action of proteases in their *in vivo* environment. We have implemented the use of three fluorescent genetically encoded proteolytic sensors in plants reporting the activity of specific proteases. Using known proteases, we have been able to set up good basis for the development of robust sensors in the plant field. We have monitored the activity of caspase-3, TEV protease and metacaspases with successful results. Moreover, we have been able to monitor the *in situ* activity of a viral protease during infection in the plant host system. Lastly, we set up the development of FRET sensors that will allow the detection of metacaspase activity in their natural existing locations like the root and by triggering conditions like wounding. We have also identified the strengths and the weak points of each reporter and proposed models for their improvement in future sensor designs.

Keywords: proteolysis, genetically encoded sensors, *Arabidopsis thaliana*, proteases, metacaspases, Nicotiana, degradomics, ZipGFP, FRET.

INTRODUCTION

Understanding the enzymatic activity of proteases is a crucial aspect to understand their biological roles in plants and animals (van der Hoorn, 2008; Salvesen et al., 2016). The use and development of biochemical probes and fluorescent reporters have resulted to be very effective to investigate in the enzymatic dynamics of proteolytic enzymes especially in the animal field with biomedical applications (Oliveira-Silva et al., 2020) and in plants to understand function of proteases in development and stress (Morimoto and van der Hoorn, 2016).

The use of genetically encoded biosensors have raised in the last years as extremely powerful tools that in parallel to other approaches, allow to study in detail biochemical processes in an in vivo cellular context. Since the discovery of the light emitting proteins their sequences have been altered, mutated, engineered and fused to other protein domains that can change their excitation and emission properties (Ai et al., 2014). In such fashion, a myriad of biosensors have been developed in order to image cellular events (Shaner et al., 2005; Kostyuk et al., 2019). Both fluorescent and luminescent biosensors are generally cheap and easy to produce by organisms where genetic transformation is feasible. On the contrary, chemical probes require specialists for synthesis and can be costly to produce. When comparing fluorescent and luminescent systems, luciferase-based sensors have the fastest maturation rates but they require the addition of chemical co-factors. On the other hand fluorescent sensors are excited by lasers that may induce phototoxicity.

In plants, the use of tools for the detection of proteolytic activity is mainly restricted to activity based protein profiling with protease family-type specificity (Morimoto and van der Hoorn, 2016). Live-cell imaging of genetically encoded biosensors is a field currently expanding in plants (Grossmann et al., 2018). The use of biosensors, initially developed in other fields of molecular biology, is being implemented rapidly in plants, i.e. ATP, calcium, glutamate redox balance and kinase activities (Monshausen et al., 2008; Ast et al., 2017; De Col et al., 2017; DeFalco et al., 2017; Exposito-Rodriguez et al., 2017; Toyota et al., 2018; Nietzel et al., 2019; Zaman et al., 2019). Specific to plants, phytohormone studies have engineered their own specific sensors, polishing their function over time (Jones et al., 2014; Waadt et al., 2014; Liao et al., 2015). With

the exceptions of BRET sensors for the detection of cysteine protease ATG4 activity in autophagy (Woo et al., 2014) and to measure the DEVDase activity in stresses (Zhang et al., 2009), the use of biosensors in the plant protease field remains largely unexplored.

In this chapter we have implemented the use of genetically encoded biosensors to detect proteolytic activity in plants. We have tested the use of ZipGFP (To et al., 2016), which allows the detection of apoptosis in animal cells, and we have newly engineered a re-localization probe, NLS^{TEVrs}NES-GFP, and a sensor based on Förster resonance energy transfer (FRET). We showed proof-of-concept for each sensor using known proteases to validate their performance and we have further explored the applications of the FRET biosensor to detect viral presence by its proteolytic activity and assess metacaspases activity in *Nicotiana benthamiana* and *Arabidopsis thaliana*.

RESULTS

Fluorescent reporter Switch-ON ZipGFP can be used in plants for the detection of proteolysis events

Initially, we tested ZipGFP, which is a ratiometric reporter that is characterized for its gain-of-fluorescence (Switch-ON) by GFP complementation after proteolytic processing (To et al., 2016). (Fig. 1A). Cleavage of the protease substrate sites releases two binding peptides that block complementation of two parts of the GFP protein. The system has a constitutive signal due to expression of mCherry after a 2A-like peptide that induces ribosomal skipping, making this probe ratiometric. We cloned ZipGFP containing a Tobacco Etch Virus (TEV) cleavage site into two independent cassettes using GreenGate cloning system (Lampropoulos et al., 2013) to obtain *pUBI10::ZipGFP1-10^{TEVrs}-2A-mCherry* and *pUBI10::ZipGFP11^{TEVrs}* (Fig. 1B). Additionally, we cloned the TEV protease sequence into a cassette with a N-terminal HA tag (*p35S::HA-TEVp*). All cassettes were transformed in *Agrobacterium tumefaciens* and the lines containing both biosensor parts were co-expressed with or without the line expressing TEVp in *Nicotiana benthamiana*. When imaged in the confocal microscope we could observe a gain of fluorescence in the green channel for an abundant number of *N. benthamiana* epidermal cells (Fig. 1C). Quantification of the fluorescence levels as the ratio of GFP/mCherry signal revealed significant differences and a four-fold ratio increase signal when TEVp was co-expressed but no differences observable when one or both cassettes encoding the reporter were expressed (Fig. 1D).

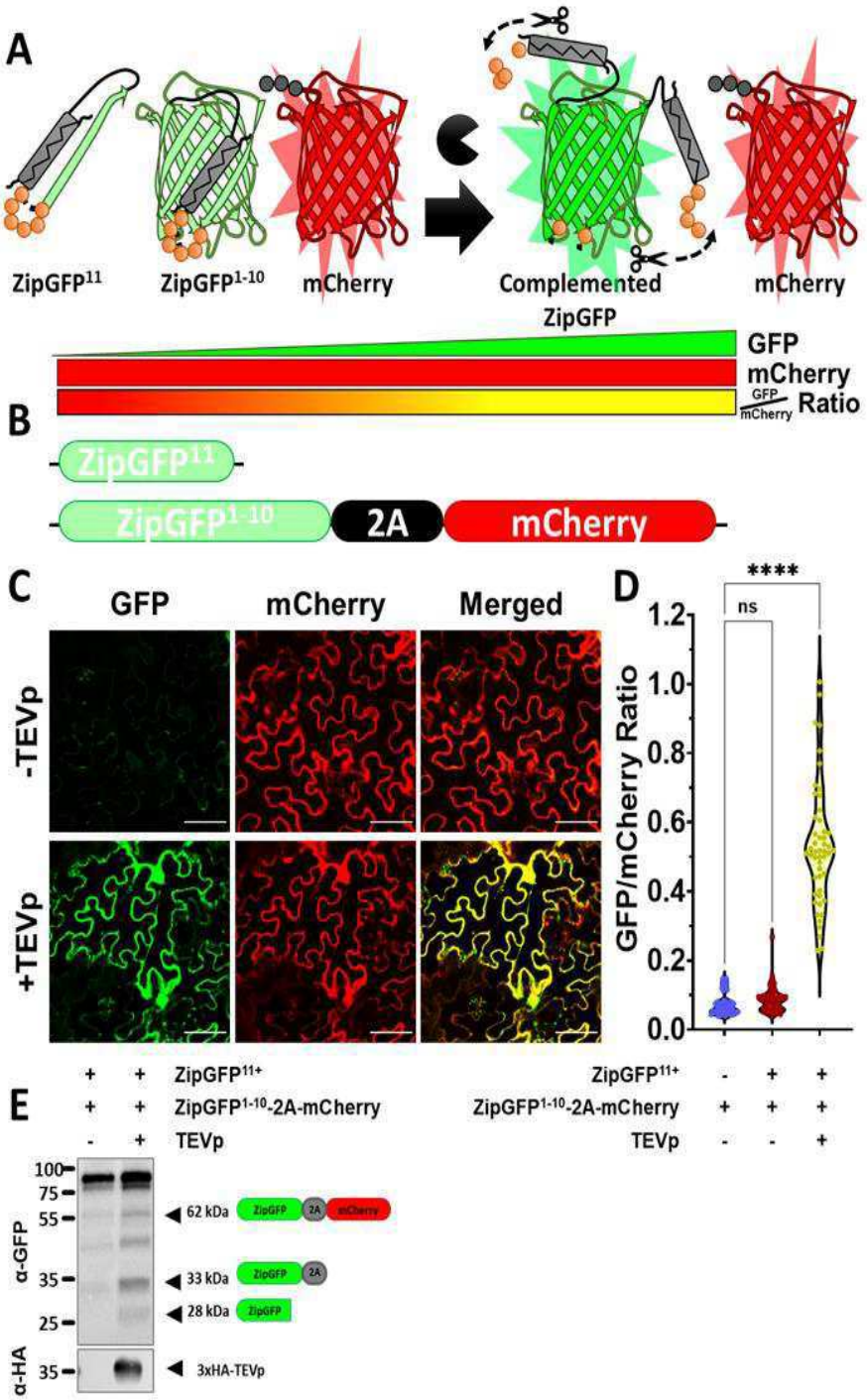


Figure 1. Use of the proteolytic reporter ZipGFP in plants. A) Rationale of the ZipGFP. Complementation mechanism of ZipGFP11 and ZipGFP1-10 occurs only after processing in two sites. Bars below show mCherry fluorescence remaining unaffected by the proteolysis status and gain of GFP signal in TEVp-mediated proteolytic conditions. Proteolysis can be quantified as the relative ratio of GFP signal divided by the mCherry signal. **B)** ZipGFP cassette used for detection of proteolysis in plants. One of the parts ZipGFP^{11} was expressed in a different cassette than ZipGFP$^{1-10}$-2A-mCherry. The second cassette bears a 2A-like peptide that induces ribosomal skipping and release of the mCherry protein. **C**) Confocal images of *N. benthamiana* epidermal cells 3 days after *A. tumefaciens* co-infiltration with ZipGFP^{11} and ZipGFP$^{1-10}$-2A-mCherry together or in absence of TEVp. **D)** Quantification of the GFP/mCherry ratio levels obtained from confocal image of transient co-infiltrated lines as indicated in *N. benthamiana* leaves (n=50 per condition; Student's t-test, p-value < 0.05)(Bars represent 50 µm). **E)** Western blots showing anti-GFP-HRP antibody detecting multiple bands from ZipGFP$^{1-10}$-2A-mCherry . Blots were stripped and re-probed for the HA tag fused to the TEVp.

Detection of proteolysis by re-localization of GFP

Further, we devised a simple GFP protease reporter that could report proteolysis based on localization changes rather than variation of the signal intensity. We used the *SIMIAN VIRUS 40* nuclear localization signal (NLS) sequence followed by a TEV cleavage that would process one single amino acid above a nuclear export signal (NES) from the Squash Leave Curl Virus (Ward and Lazarowitz, 1999) which is fused to EGFP, and we denominated it NLS^{TEVrs}NES-GFP (Fig. 2A). We forecasted that TEVp would remove the NLS from the reporter amino acid chain, thus exposing the NES that should induce exit of the reporter to the cytosol (Fig. 2B). Following the same procedure as for ZipGFP we infiltrated *N. benthamiana* leaves with or without *p35S::HA-TEVp*. The preliminary results showed differences in the localization pattern. Comparison of randomly chosen areas while imaging revealed changes in the localization from nuclear to cytosolic regions in sectors where TEVp was expressed (Fig. 2C, E and F). However, many cells still showed GFP fluorescence in the nuclei when focusing on the right plane. This observation could be explained by different reasons. The first would be the presence of an uncleaved subfraction of the biosensor in its original location. A second reason could be that after initial cleavage by TEVp and transport to the cytosol, the NES could be degraded and NES-free GFP may re-enter to the nuclei as it is reported for free GFP expressing plants. Also, immunoblots of the reporter showed dimerization of the reporter when TEVp was not expressed (Fig. 2D). While we have no explanation for the effect of TEVp presence on inhibiting the dimerization events of GFP, we concluded that NLS^{TEVrs}NES-GFP would benefit from further optimization steps.

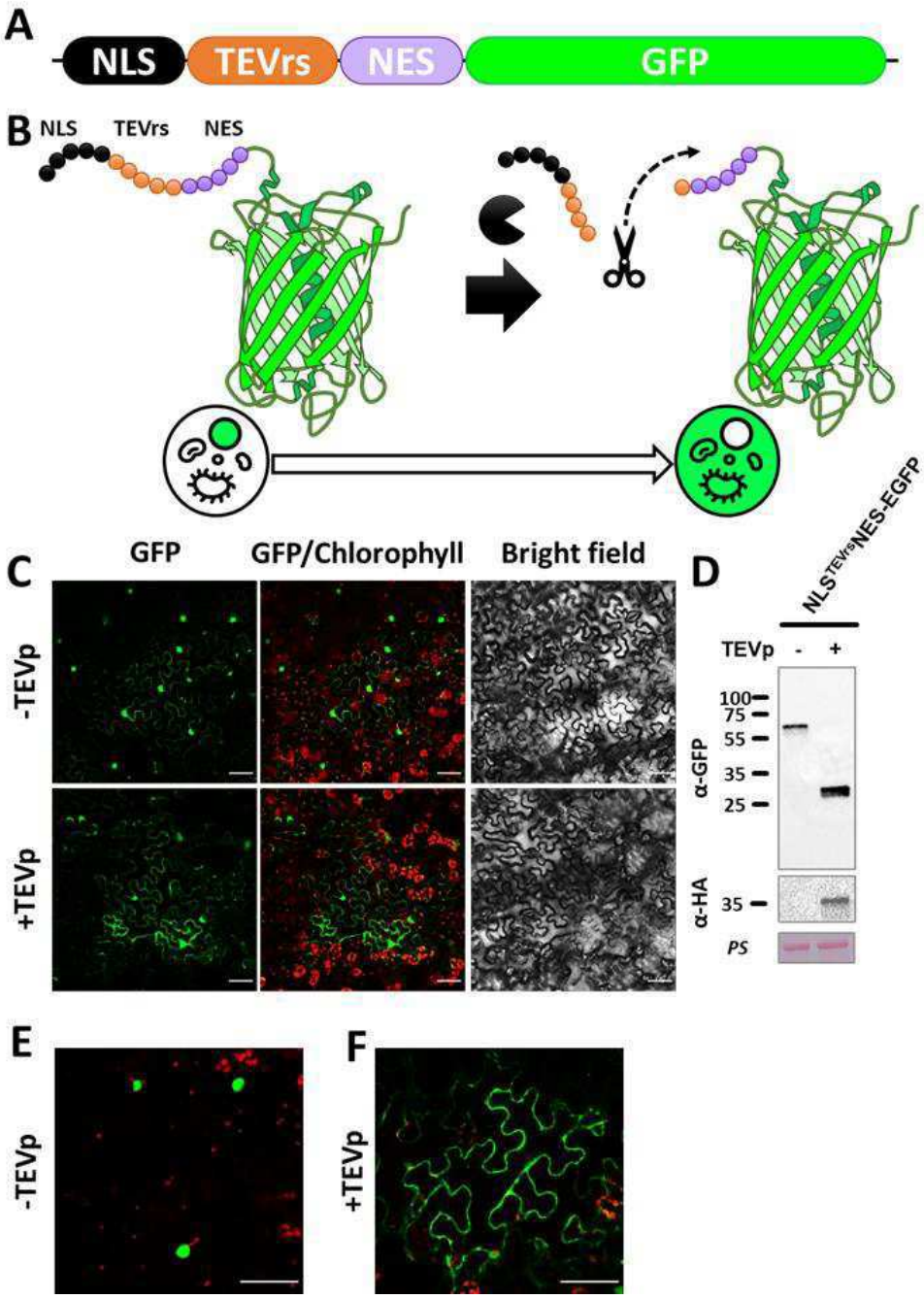


Figure 2. NLS^{TEVrs}NES-GFP as a proteolysis reporter by localization changes. A) Scheme of the reporter containing *SV40* NLS, a TEV recognition site (ENLYFQS) and NES preceding GFP. **B)** Rationale of NLS^{TEVrs}NES-GFP indicating cleavage (scissors) and its corresponding localization changes from nuclear to cytosolic. **C)** Confocal images were taken three days after infiltration of NLS^{TEVrs}NES-GFP in random areas of *N. benthamiana* leaves. **D)** Anti-GFP and anti-HA immunoblots for the NLS^{TEVrs}NES-GFP samples with or without co-expression of *p35S::HA-TEVp*. In absence of TEVp the reporter seems to dimerize around 65 kDa. When the protease is present the reporter size is 28kDa. Relative protein load is represented by Ponceau staining (PS) of the rubisco large subunit. **E)** Merged image of confocal images zooming on epidermal cells expressing the reporter showing mainly nuclear localization, or **(F)** the reporter and TEVp showing cytosolic difusion. (Bars represent 50 µm).

FRET reporters can efficiently report proteolysis in plants

Next, we pursued the use of biosensors with two fluorescent proteins. In order to increase our confidence in the protease reporters, we also added additional controls. For this, we sub-cloned the active domain of mammalian caspase-3 into a plant compatible cassette. The advantage of using caspase-3 is that there exists plenty of tools and extensive literature in the mammalian field to study and apply in a research question. To test caspase-3 cleavage in plants we used CR(DEVD)$_2$, a chemical reporter that turns fluorescent by proteolysis upon cleavage (Fig. 3A). Also, CR(DEVD)$_2$ has been previously used in plants to detect caspase-3 like activity as a cell death proxy, showing its permeability (Bosch and Franklin-Tong, 2007; Lin et al., 2020). When applied to *N. benthamiana* leaves we could detect high levels of fluorescence indicating that the transiently expressed caspase-3 is active in plants (Fig. 3B). These results allowed us to use caspase-3 as additional control in the development of our plant FRET sensors. FRET is based on the transmission of energy between two fluorescent proteins that are in a proper spatial orientation and proximity (Lam et al., 2012). During FRET, part of the energy of an excited donor fluorophore is transferred to an acceptor fluorescent protein (acceptor) in close proximity, which is thereby excited and emits light. Protease reporters relying on FRET are engineered with a linker resembling a protease substrate, which unlike FRET sensing of other stimuli like ATP or calcium, is irreversible when cleaved (Fig. 3C). We took advantage of the overlapping emission and excitation spectra of the green-yellow mNeonGreen and the orange-red mRuby3 fluorescent proteins (Fig. 3D) (Bajar et al., 2016). This fluorescent protein pair has been tested in plants for intermolecular FRET with good results (Denay et al., 2019) and similar pairs have been used in other model organisms for the detection of histidine kinase activity by FRET (Duvall and Childers, 2020). Additionally, we added two peptides WW, derived from the human YAP65 and Wp, an interactor with moderate affinity for WW (WW:PDDVPLP, Wp: GTPPPPYTVG). Both peptides were cloned at the FRET sensor N- and C-terminus ends respectively. These peptides induce a weak interaction bond and thereby enhance energy transfer between the FRET fluorophores (Grunberg et al., 2013). Between either fluorophores we cloned a protease recognition site (RS) or a four-times repeat of a glycine-glycine-serine linker, as negative control.

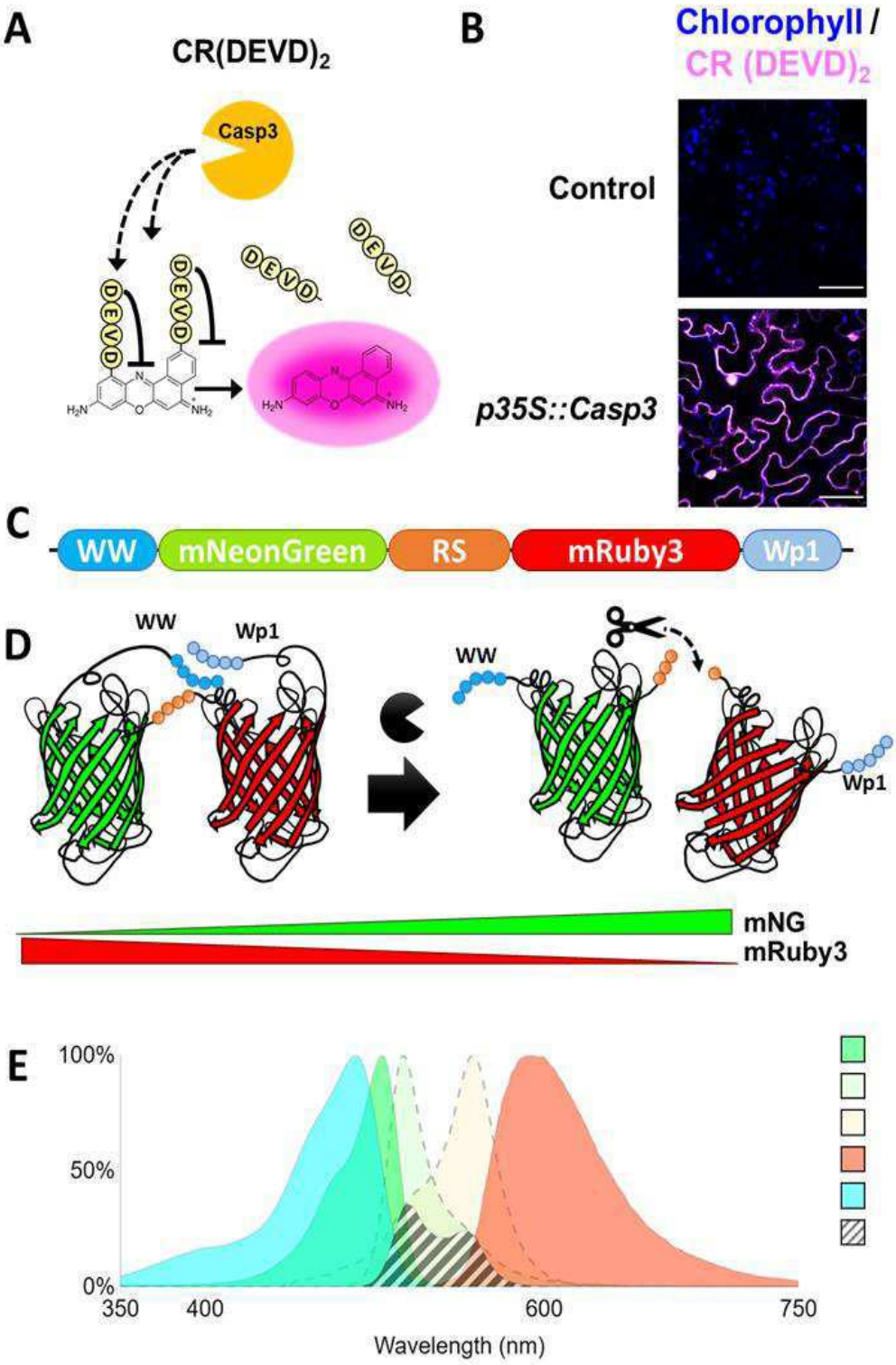


Figure3. A) CR(DEVD)$_2$ scheme depicting the mechanism to detect caspase-3. **B)** Confocal images show that caspase-3 works efficiently in *N.benthamiana* epidermal cells as detected by fluorescence of CR(DEVD)$_2$ (Bars represent 50 µm). **C and D)** Conceptual design of the FRET reporters for the detection of proteolytic activity. mNG: mNeonGreen; WW and WP1 are the weak FRET helpers. **E)** Excitation and emission properties of mNeonGreen and mRuby3. FRET overlapping Emission of mNeonGreen and mRuby absorbance is streaked in white and black.

Prior to assessing the effect of proteolysis on the sensor we used the negative control FRET^{6xGS} (Fig. 4A) and performed a photo-bleaching experiment. By acceptor bleaching, the energy to be transferred to the acceptor is almost fully used instead by the donor, in this case mNeonGreen, and consequently the ratio of the acceptor/donor signal is decreased. Bleaching of selected areas showed a strong localized decrease in the FRET ratio of approximately 60% (Fig. 4B). Fluorescence measurements show an immediate decrease in the FRET mRuby3 derived signal and a corresponding increase of mNeonGreen, indicating that the biosensor is capable to work as a FRET reporter. The changes the FRET ratio value (mRuby/mNeonGreen) decreased from an initial value normalized to the unit, to values under 0,5 (Fig.4C, black values). The signal decreased of 2-fold times after multiple photo-bleaching would indicate the minimal FRET values. Also our values stay in the range of the results for the ratio (mCitrine/mTurquoise) of the original FRET sensor using WW and Wp (Grunberg et al., 2013), indicating feasibility of mRuby/mNeonGreen as FRET pair.

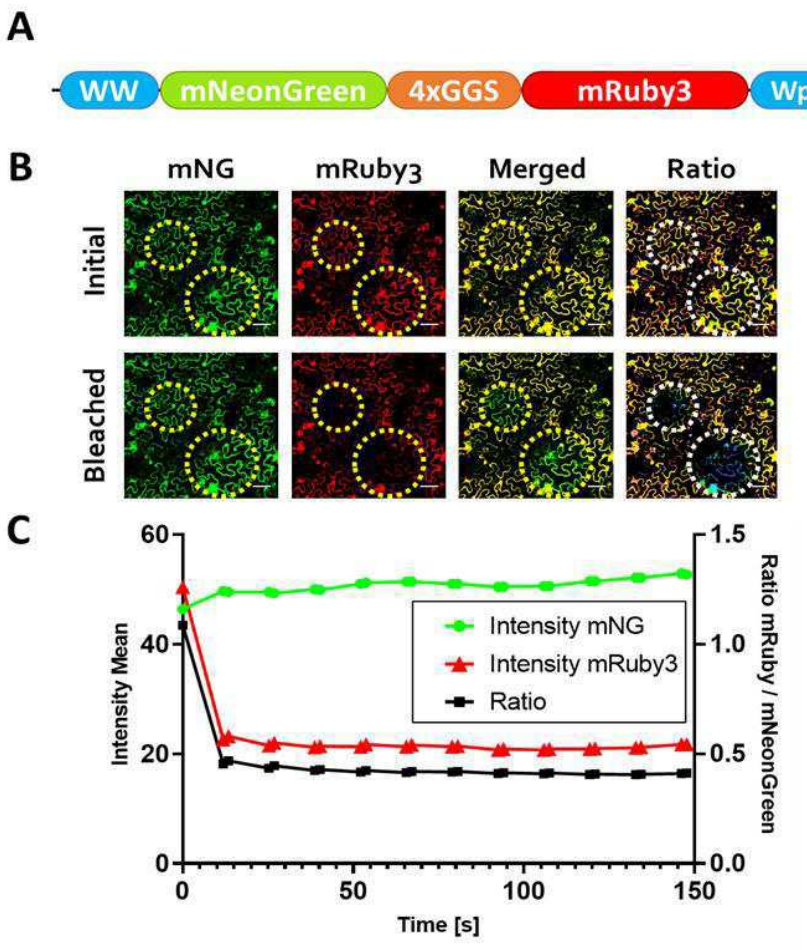


Figure 4. Bleaching experiment for the transiently expressed version of the biosensor FRET^{x6GS}. A) Design of the biosensor with a flexible 4xGS linker. **B)** Photo-bleaching experiments of the areas indicated with yellow and white circles lead to decrease of the red channel derived from the FRET signal when comparing initial to bleached conditions, and a subsequent slight increase in the mNeonGreen signal. Differences are observable in the merged channels as well as in the ratio composed image (Bars represent 50 µm). **C)** Quantification over time of the signal intensity

changes following bleaching of mRuby3 for the different channels (mNeonGreen emission and FRET mRuby3) and the ratio of mRuby3/mNeonGreen.

For our first trial, we included the caspase-3 minimal substrate sequence DEVD in the protease site (Stennicke et al., 2000) and named the resulting probe FRET^{DEVD} (Fig. 5A). We imaged at day two after infiltration. Confocal imaging showed a reduction of the mRuby3 FRET signal when caspase-3 was co-expressed in tobacco leaves (Fig. 5B). Quantification of the FRET displayed a four-fold decrease in the mRuby3/mNeonGreen FRET ratio(Fig. 5C). These results were similar to those obtained using the helper peptides fused to a Turquoise2 and mCitrine FRET pair with a caspase-3 recognition site in mammalian cells (Grunberg et al., 2013). Anecdotally, when the FRET^{DEVD} sensor was transiently co-expressed with caspase-3 in the leaves, the cells experienced a cell death like response. When we expressed separately the sensor and the caspase-3, severe cell death was only observed in the caspase-3 infiltrated sectors (Fig 5D). In relation to several references in the literature to caspase-3 like activity in plant cell death (Sueldo and van der Hoorn, 2017; Buono et al., 2019), here we found that ectopic expression of mammalian caspase-3 in plants can induce cell death. In conclusion, the results showed that the FRET sensor can be used for the detection of proteolytic cleavage in plants and we further explored this by modification of the protease RS.

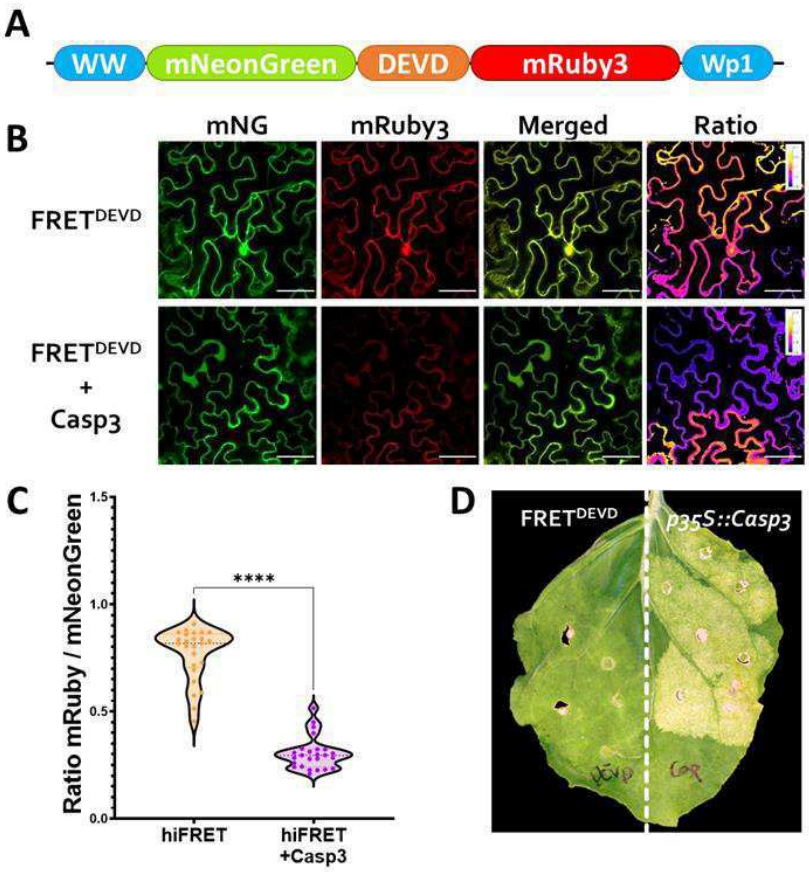


Figure 5. The caspase-3 biosensor FRET^{DEVD} is capable of detecting DEVDase activity in plants. **A)** FRET^{DEVD} design showing DEVD tetrapeptide as linker. **B)** Confocal images showing separated channels, merged and ratio picture (Bars represent 50 µm). **C)** Ratio quantification of FRET^{DEVD} and the sensor with caspase-3 (orange and purple respectively; Student's t-test, p-value < 0.05; n=25 per sample). **D)** Cell death observed in *N. benthamiana* induced by transient expression of caspase-3.

Following, we modified the protease RS with the TEV cleavage recognition motif (ENLYFQS), also used in ZipGFP and NLS^{TEVrs}NESGFP. Similar to the caspase-3 reporter, the FRET^{TEVrs} sensor showed a four-fold reduction of FRET ratio at two days after infiltration (Fig. 6A and B). Interestingly, five to seven days after infiltration, cell death lesions could be observed,. Trypan blue staining of the *N. benthamiana* plants confirmed that the TEVp cassette can induce cell death by an unknown mechanism (Fig. 6C-F). The positive results obtained from co-expression of the FRET^{DEVD} and FRET^{TEVrs} sensors with their respective proteases prompted us to use FRET-based probes to investigate endogenous plant and viral protease activity.

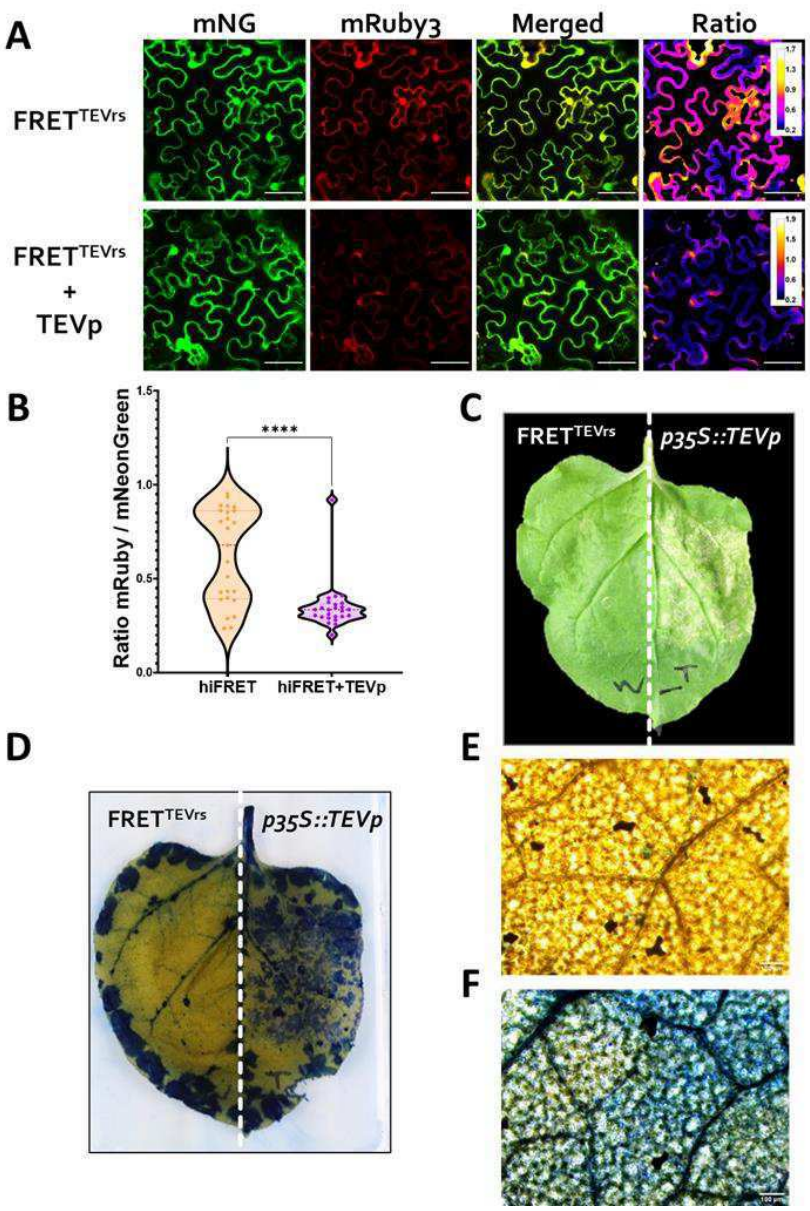


Figure 6. FRET^{TEVrs} works efficiently in the detection of TEV proteolytic activity. A) Panel showing fluorescence of mNeonGreen, FRET derived signal of mRuby3, merged and ratio panels showing clear drop in the TEVp expressing conditions (Bars represent 50 µm). **B)** Quantification of the acceptor/donor level ratio (mRuby/mNeonGreen) of FRET^{TEVrs} confocal images expressing only FRET^{TEV} (orange) and FRET^{TEV} together with the TEVp (purple). (Student's t-test , p-value < 0.05, n=25 per condition). **C)** A leaf seven days after *A. tumefaciens* transient infiltration with only the sensor or only the TEVp showing mild signs of cell death. **D)** Photography of trypan blue treated leaves described in (C). **E)** Microscopy image of trypan blue staining of cells *N. benthamiana* leaf sectors for mock plants and **F)** cells from a leaf expressing TEVp (Bars represent 100 µm).

FRET-based proteolysis reporters can detect viral infections in plants

We assessed whether the FRET^{TEVrs} sensor can work as a reporter of viral infection. The commonly used TEVp is an optimized, and in some cases engineered, protein from the Tobacco Etch Virus Nuclear-inclusion-a endopeptidase (NIa). TEV affects plant development and growth (Bedhomme and Elena, 2011) and the TEV protease NIa is required for cleavage and functionality of the original viral polyprotein chain into seven functional proteins (Carrington and Dougherty, 1988). In order to test this, we expressed the sensor in 6-8 week old plants of *N. tabacum* plants inoculated with TEV

or non-infiltrated. Leaves with obvious signs of infection (Suppl. Fig. 1) were recalcitrant to *Agrobacterium tumefaciens* transient infiltration. The FRET^{TEVrs} sensor infiltrated in healthy-looking leaves of control and virus inoculated plants displayed slight differences in the FRET ratios (Fig. 7B). The fluorescent signals were heterogeneous, with patches of adjacent cells varying strongly in their FRET ratios (Fig. 7A). Nevertheless, the ratios were significantly different between the healthy and virus inoculated plants (Fig. 7B). The difference in extent of FRET ratio compared to TEVp overexpression could be explained by the heterogeneity of the virus infection or by the decreased activity of the viral NIa protease in comparison to TEVp that has been engineered for stability and efficiency. Nevertheless, we managed to distinguish infected from control plants through FRET ratio changes of FRET^{TEVrs} before any visible signs of viral infection.

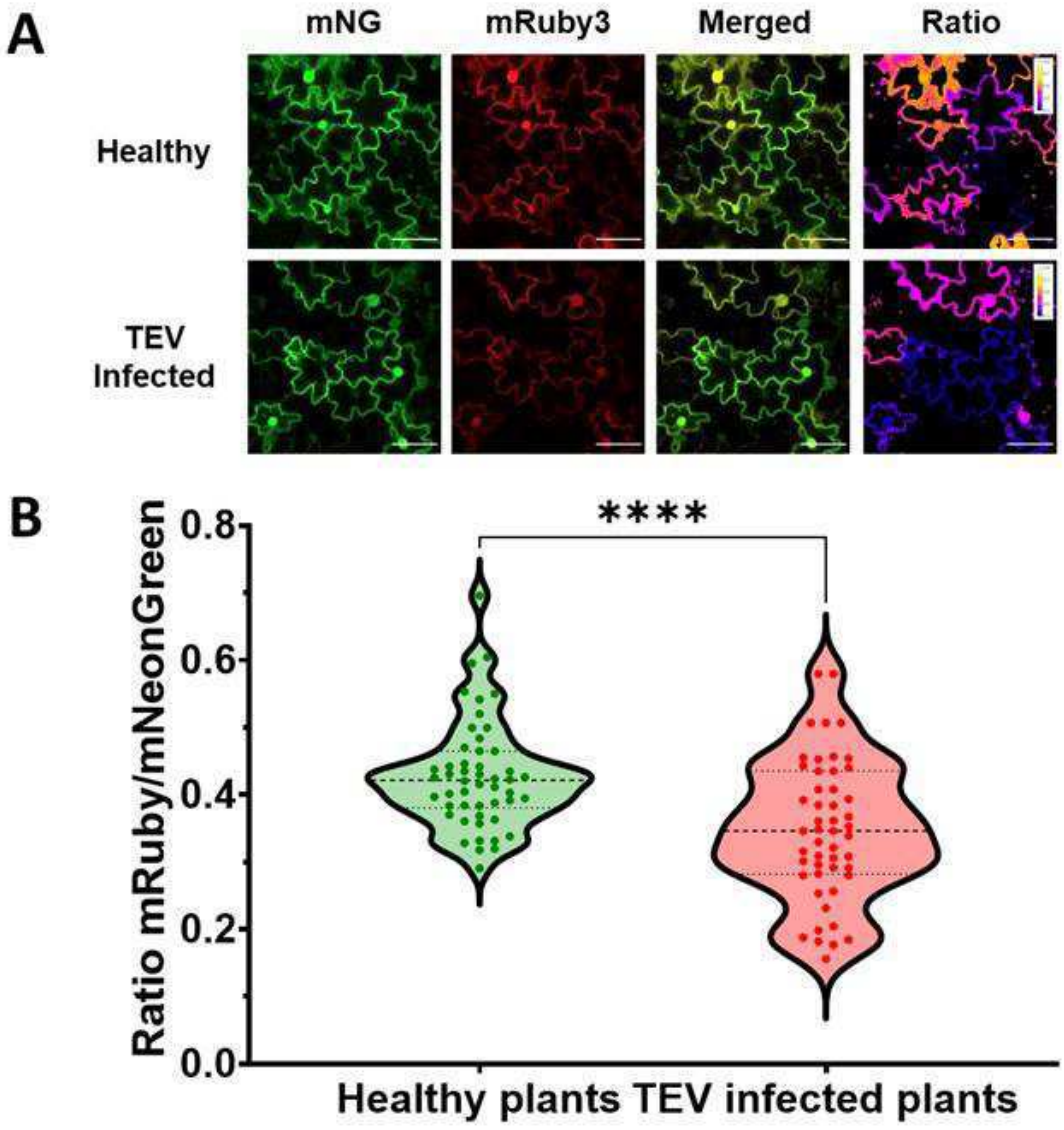


Figure 7. FRET^{TEVrs} can detect TEV infection by its proteolytic activity. A) Panel showing confocal images for the fluorescent proteins of the FRET^{TEVrs} biosensor, mNeonGreen, FRET signal from mRuby3, merged channels and ratio (Bars represent 50 µm). **B)** Quantification of the ratio signal acceptor/donor (mRuby3 FRET/mNeonGreen emission) for healthy plants (green) and infected plants (red) showing small but significant differences between samples (Student's t-test, p-value < 0.05, n=50 per condition).

FRET sensors shows promising results for the detection of metacaspase activity

Lastly, we have engineered FRET sensors with metacaspase cleavage motifs in the protease RS. The composition of the amino acid cleavage sequences were obtained from previous studies that represent the specificity of AtMC4 and AtMC9 (Vercammen et al., 2004; Tsiatsiani et al., 2013; Hander et al., 2019).The sequences contain differences in the amino acid lengths, but we expected that FRET efficiency would be similar to some of the reporters described above. For instance, the fluorescent proteins in the FRET^{6xGS} design are linked by 12 amino acids. In addition, the nature of the amino acids in the linker should be considered. For example, GS or GGS repeats are likely flexible and will not compromise the FRET efficiency by displacing the fluorescent proteins, while rigid linkers could change the distance or relative position between FRET pairs thereby decreasing FRET efficiency. For our purpose, we attempted to keep small linkers, but we did not evaluate differences in FRET between reporters with different linker sizes.

The sequences included in the protease RS were metacaspases minimal recognition motifs VRPR, EKAKD, EKTKD, EKTRD obtained frowm AtMC4 and AtMC9 substrate preference and VTSRATKV and VTSRATKVKAKQRG based on PROPEP1 cleavage site and named PCS1 and PCS2 respectively for simplification (Fig. 8A, only FRET^{VRPR} is shown as representative example). We stably transformed these sensors and selected *Arabidopsis thaliana* plants. We found that in the FRET^{VRPR} lines, there was a ratio change at specific locations in the root corresponding to outer layers of lateral root emergence during early stages of lateral root emergence (Fig. 8B) and external sites at the middle height of the root cap in the primary root (Fig. 8C and D). The locations where the FRET ratio levels reach minimums correspond to the expression patterns described for *AtMC9* in the root cap (Fig. 8E) and in the endodermal cells on top of growing lateral root primordia (Escamez et al., 2020). While

AtMC9 is also expressed in the vascular bundle in roots proto-xylem cells (Fig. 8F), we could not detect any FRET difference in this area with our set-up.

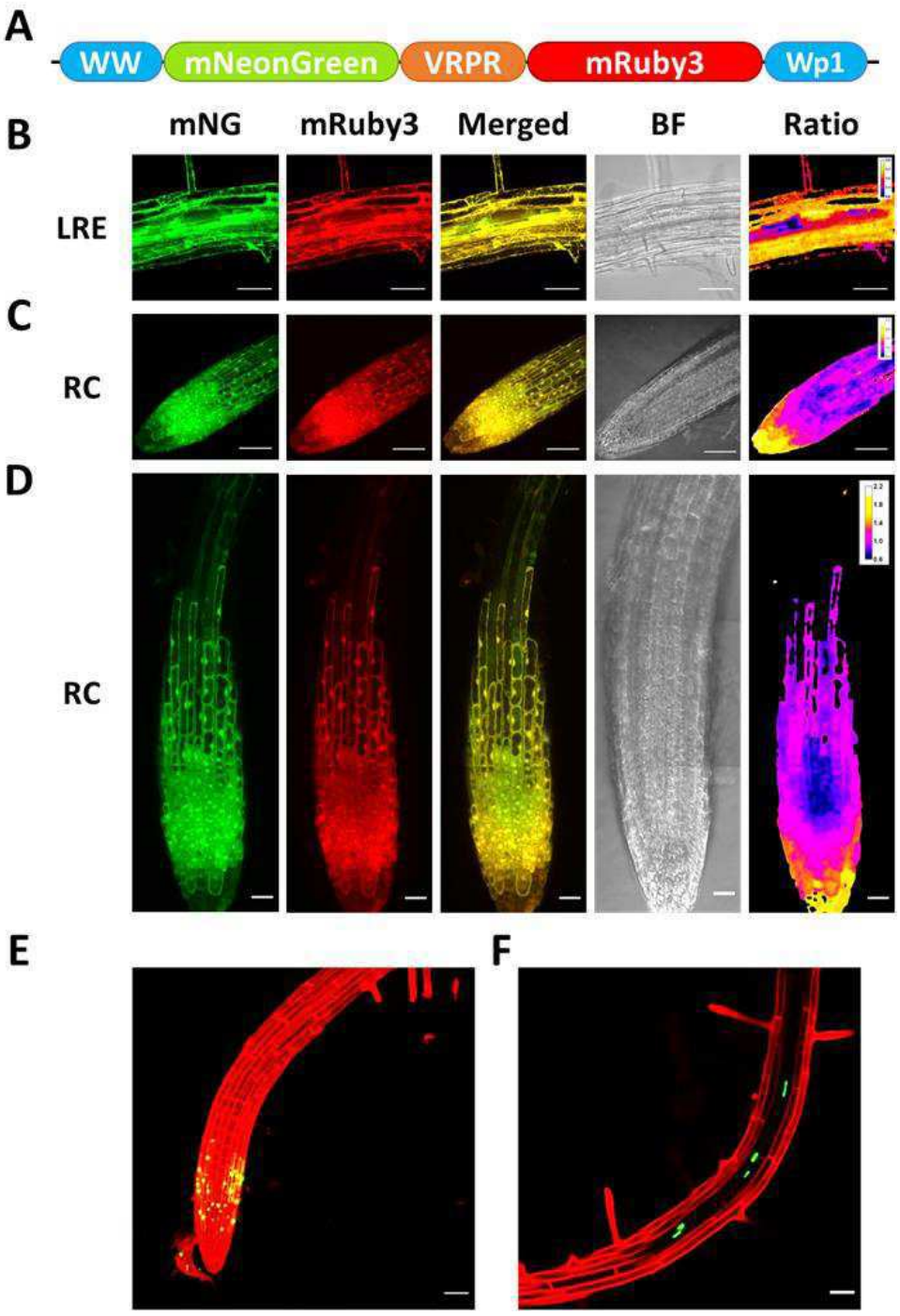


Figure 8. FRET^{VRPR} allows the detection of proteolytic changes in regions where AtMC9 is epxressed in *Arabidopsis thaliana*. A) Design of the FRET^{VRPR} sensor. **B-D)** Panels showing Z-stack projections of confocal images for the different channels of the FRET^{VRPR} sensor. mNeonGreen signal (mNG), FRET derived signal of mRuby3 (mRuby3), merged imaged of mNeonGreen and mRuby3, bright field (BF) and composed FRET ratio images (mRuby3 FRET/mNeonGreen emission). **B)** Representative image of the FRET^{VRPR} sensor in the lateral root emergence (LRE) areas (Bars represent 50 µm) **C)** Representative image of the FRET^{VRPR} sensor in the root cap (RC; Bars represent 50 µm), and **D)** Composed image by stitching of 3 adjacent Z-stacks images of the root tip in *Arabidopsis* thaliana (Bars represent 20 µm). **E-F)** Merged images of the fluorescent expression pattern of the AtMC9 shown by the promoter fusion line *pATMC9::NLS-3xGFP* (green) in PI stained Arabidopsis root

seedlings (red). pAtMC9 derived signal is localized mainly in root cap cells of the primary root (**E**) and xylem cells at higher tissues (**F**) of 10 days-old Arabidopsis plants (Bars represent 50 μm).

Additionally, we infiltrated the sensor FRET^{PCS1} (Fig. 9A) in *N. benthamiana* plants and performed laser wounding experiments, as we previously found that AtMC4 is activated by this method (Hander et al., 2019). To induce a higher response we co-infiltrated AtMC4 together with the sensor. We observed that by laser wounding the cells targeted were bleached, similarly to the results observed in the FRET^{6xGS} as described in figure 4C. However for the FRET^{PCS1}, we detected changes in the surroundings of the damaged cell for a time of 5 minutes (Fig. 9B). this observations were cleared when we observed the nuclei of the damaged cell, undergoing immediate bleaching (Fig. 9C) and the ratio of the adjacent cell which decreasing gradually overtime (Fig. 9D). When FRET values were quantified a rapid FRET decrease was detected in the wounded cell and a lower but constant and gradual response in the adjacent cell (Fig. 9E).

In this chapter we have shown that proteolytic reporters are efficient tools that can be used in plants for the detection of proteolytic activity and that they can be engineered for specific purposes, from viral infections to the detection of plant proteases in their native conditions. We expect further efforts to confirm our results in order to be able to characterize proteolytic activity of plant proteases in vivo.

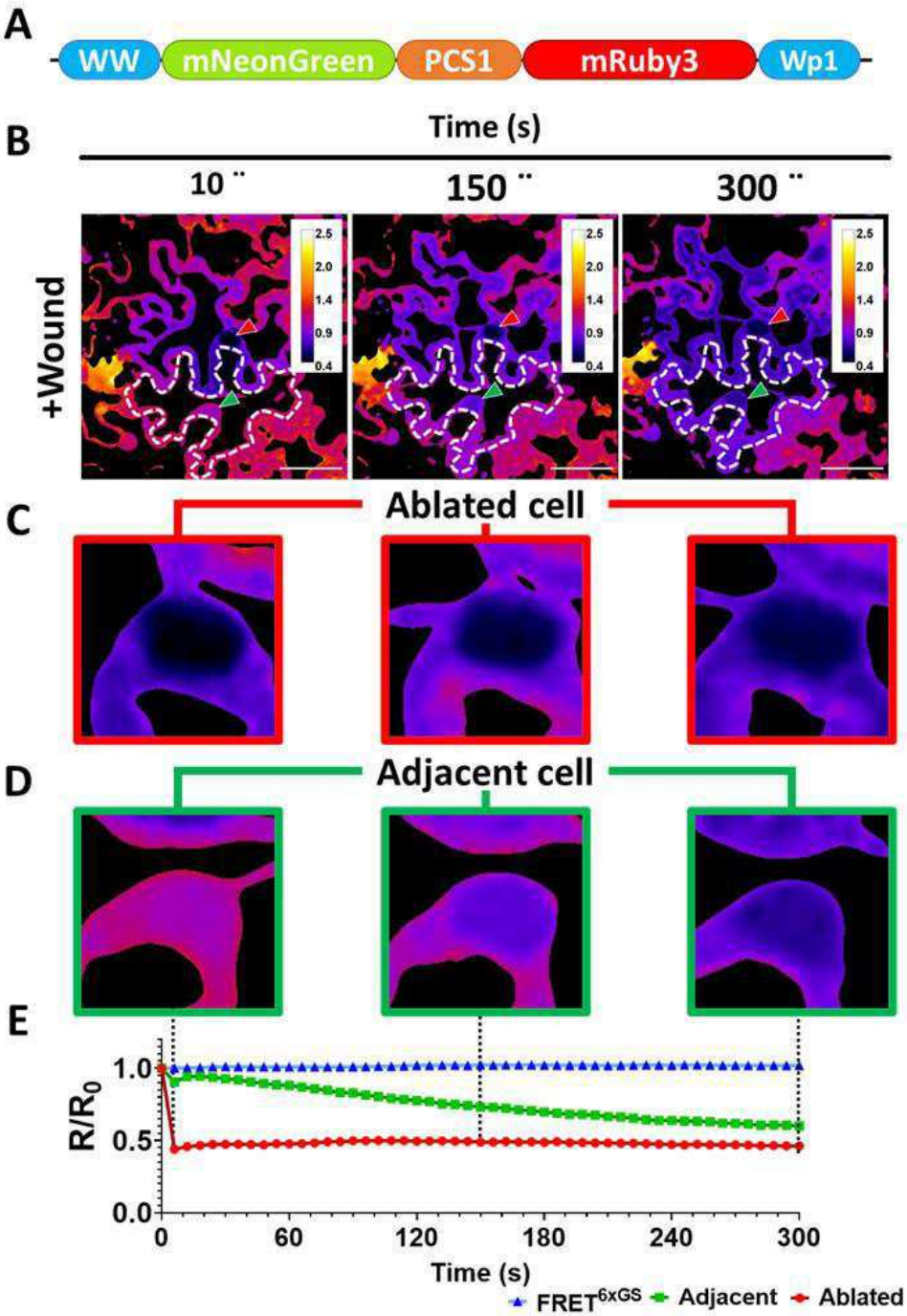


Figure 9. FRET^{PCS1} shows a drop of fluorescence in proximal cells to the wounded area when co-expressed with AtMC4. A) Representation of the design of the FRET^{PCS1} reporter bearing a PROPEP1 derived motif. **B)** Ratio images of the FRET signal after laser ablation at timepoints 1, 150 and 300 seconds after ablation. The nuclei of the ablated cell and the adjacent cell are indicated with a red and a green arrow respectively. The FRET ratio of the cells around the ablated cells decreases over time, illustrated by the cell below, delimited with a dashed white line. Plants were co-infiltrated with FRET^{PCS1} and a cassette overexpressing AtMC4 (pB7WG2,0-AtMC4s). Images were taken two days after infiltration. (Bars represent 50 µm). **C)** Enlargement of the FRET ratio for the ablated (delimited by a red frame) and **D)** enlargement of the nuclei of the adjacent cell (delimited with a green frame) at 10, 150 and 300 seconds after wounding. **E)** Quantification of the normalized ratio (R/R_0) of the FRETP^{CS1} mRuby/mNG signal for the nuclei of the ablated cell (red) and the nuclei of the adjacent cell (green). Additionally a negative control using the FRE^{6xGS} (blue) is shown. Measurement of the fluorescent intensity elapsed to a period of 300 seconds showing initial decrease on the FRET value of the red and green line due to the ablation, following dynamics changes in the values over time as indicated.

DISCUSSION

In this chapter we tested a set of genetically encoded fluorescent proteolysis reporters in plants and assessed their use for answering plant-specific biological questions. We verified that proteolytic sensors can be successfully used in plants and that they can become useful tools to assist in determining proteolytic dynamics in time and space. Genetically encoded versions have the advantage to be cheaply produced in organisms where genetic transformation is possible. Genetically encoded sensors are also easily adaptable, as the amino acid sequence of the protease RS can be adjusted to the substrate recognition motifs of a protease or a protease family to study. Furthermore, genetically encoded sensors can be addressed to sub-cellular compartments with the use of targeting signals. Therefore, genetically encoded sensors allow their development and accommodation for specific purposes. Additionally, genetically encoded reporters can be generally purified by the addition of affinity tags allowing to deeper characterize their properties and to contrast *in vivo* results with results obtained *in vitro*.

Each sensor in this study has different advantages and drawbacks. ZipGFP, similarly to newer protease sensors like FlipGFP (Zhang et al., 2019), is based on the complementation of split parts of GFP. Activatable sensors are amiable to visualize and to detect due to the gain of fluorescence in specific locations. Here, we also correlate increase of activity with increase of fluorescence, while for FRET and BRET proteolytic sensors increase of enzymatic activity is reported by a loss of signal. Moreover, ZipGFP and FlipGFP contain a permanent fluorescence (mCherry) in their design, turning them into ratiometric sensors. Importantly, in steady state conditions, ZipGFP displayed low background signal when expressing the two uncleaved parts ($ZipGFP^{1-10}$-2A-mCherry and $ZipGFP^{11}$; Fig.1D). For example, this signal was not significantly different when compared to expressing only the bigger fragment $ZipGFP^{1-10}$-2A-mCherry. Also, our results show that both halves of the reporter require TEVp activity for being able to complement. We observed small background signal levels in the green channel as it could be expected from reports self-complementation of the split fluorescent parts. These results often lead to false positives and artifacts as seen by other methods using split fluorescent proteins complementation (Kudla and Bock, 2016). However, complementation based systems require time to assemble and

shape a mature complex that delays the real-time response. This property would complicate the study of fast-acting proteolysis dynamics, for example, of metacaspase activity within minutes of damage. Because of its superior background and ratiometric capacity, ZipGFP would still be an excellent tool, for example to label cell types where proteolysis occurred, like recently reported for viral proteases in mammalian cell culture (Guerreiro et al., 2020).

Sensors based on localization changes could, in theory, immediately report proteolysis by the changes in the cellular environment. In our hands, we could detect changes in localization between the control lines expressing only the reporter NLS^{TEVrs}NES-GFP and the reporter together with TEVp. We carefully used the same conditions for the expression of the reporter like bacterial days of incubation, bacterial OD_{600} concentration, pre-infiltration cultivation, infiltration in plants at same developmental status and comparison of images by opposite sectors in the same plant leaf. Nevertheless, it might be that for less trained users the results would not be so clear and we identified some drawbacks in our first conception. The first is that expression signal by transient expression in *N. benthamiana* can be variable and some cells could express the reporter at higher levels than others giving the false impression of more cytosolic signal. This could be interpreted as a sign of higher levels of proteolysis and therefore one must be cautious with conclusions. Another is the above-mentioned detection of nuclear signals even in plants expressing the TEVp. Both complications could be solved with the optimization of the sensor into a ratiometric sensor and careful quantification of nucleus versus cytosolic signal. By this approach, we could detect changes on the reporter localization in contrast to a second signal, which would be insensitive to protease action. Similarly, other molecular tools, such as Jas9-Venus and rBIFC 2-in-1, have used the expression of free or nuclear tagged red fluorescent proteins as ratiometric signal (Grefen and Blatt, 2012; Larrieu et al., 2015). Preferably, both fluorescent proteins (NLS^{TEV}NES-GFP and the proteolytic insensitive fluorescent protein) should be included in the same T-DNA using the same genetic components by use of the same promoter or gene elements with similar expression levels.

Our FRET proteolysis reporters showed efficient transfer of the energy when the sensor remained uncleaved and favorable results by FRET loss when reporting proteolytic activity. We could use specific FRET sensors with their corresponding proteases obtaining similar results. This shows the reliability of the sensor when

adapting its modules, and potential for expanding its applicability to other proteases with known minimal substrates.

Here we exclusively tested the functioning of the reporter bearing a protease recognition site with its own protease. Given the flexibility on the design, the use of the reporter with a variety of proteases and their corresponding recognition sites could be attempted. For instance the study of several plant viral proteases with different and known specificity like the Nuclear Inclusion a protease from Turnip Mosaic Virus or Potato Yellow Virus that cleave after VXXE/VXXQ where X has a flexibility for different amino acids (Hidaka et al., 1992; Han et al., 2010).

We were able to identify FRET variations depending on metacaspase expressing locations. Low levels of FRET in root cap and lateral root emergence sites coincides with AtMC9 expression sites. It would be interesting to estimate the pH levels in those regions from the surroundings, using for instance genetically encoded pH-sensor that relate possible activation of AtMC9 (Moseyko and Feldman, 2001; Vercammen et al., 2004). In parallel, we could see differences in the FRET dynamics obtained from a laser wounding experiment. This activity could be accounted for by the calcium activation of AtMC4 in proximity to a wound site. In the transient approaches shown, AtMC4 under the promoter 35S was co-expressed in *N. benthamiana* leaves in order to obtain a more stringent response an rate the FRET sensor. The goal of the sensor is the detection of wounding in roots, but having detected general expression of AtMC4 and conservation of metacaspases in other species, it might be useful for the detection of metacaspase activity in other tissues or other model species. Also, additional controls would be required to claim metacaspase specificity. Ideally, the sensors would be tested in plants with depleted metacaspase activity with mutants of high order, or rather the application of metacaspase specific inhibitors like VRPR-fmk or agents depleting type-II metacaspase activity such as calcium chelators could back up this hypothesis. A parallel effort should be done in order to characterize *in vitro* the FRET biosensor in calcium and pH ranges that resemble those detected in plants in wounding and AtMC9 expressing locations.

Moreover, we could detect FRET differences by the proteolytic activity in virus-infected plants. We found this experiment extremely challenging and this may open new

venues in further stages of the sensor development. However, only one biological replicate was performed and more repeats are needed to obtain more consistent data. Viruses are one of the biggest causes of crop losses worldwide (Jones, 2021) and proteolytic reporters could become useful tools to investigate them. The sensors could be used as readouts in screenings for chemical inhibitors that prevent protease activity, haltering their assembly and plant spread. This application could be translated to other systems like blockage of viral protease affecting humans with known substrate affinities as done for SARS-CoV-2 main protease (Ullrich and Nitsche, 2020; van de Plassche et al., 2020).

In the future, the FRET reporter could be further optimized. In the development phase, we chose mNeonGreen and mRuby3 as a FRET pair due to their high brightness and photo-stability (Bajar et al., 2016). The use of a green/red pair is beneficial, as most of the laboratories are currently using green/red filters in their fluorescent settings. Moreover, due to their spectral shift, the FRET pair could report better the signal originated in deeper tissues than FRET blue/yellow pairs. This might become useful, for example, in the detection of AtMC9 from xylem cells. A downside of green/red pairs is the overlap of emission from red FRET acceptors with chlorophyll auto-fluorescence of green tissues. This property requires the use of filters that might reduce the FRET signal. In our case, we were not concerned due to the predominant expression of type-II metacaspases in roots (Bollhoner et al., 2013). Furthermore, additional optimization of the FRET pairs could be done. Recently, the FRET pair mNeonGreen and mScarlet-I reported good energy transfer and might be wise to test the dynamic range of this FRET (McCullock et al., 2020). For leaf expressing proteases the use of other FRET pairs like mTurquoise2 and mNeonGreen could be an alternative. Localization of the sensor to specific location could be interesting when a specific protease is located in an organelle.

We have tested different protease reporters acknowledging the different advantages and drawbacks for each sub type of fluorescent reporter. We could confirm that even they are not spread through the plant molecular biology community in plants, protease reporters can be used in plants. We tested the functions of specific reporters for caspase-3 activity in plants and of the plant-colonizing virus TEV. In addition, genetically encoded sensors allow the modulation of their parts and we tested metacaspase activity with sensors including motifs that are preferred by

metacaspases. Thus, specific sensors could give us information that before remain obscure and determine the time and location of proteolytic activity in plant organisms. We expect that the usage of these sensors will spread among different groups allowing plant community to reveal the temporal and spatial dynamics of their protease of interest.

MATERIAL AND METHODS

Plant material and growth conditions

7 to 10 days-old *Arabidopsis thaliana* (Col-0) plants were used for root imaging. Plants were sterilized using the gas chlorine method and placed in squared plates containing ½ MS solid media, covered with aluminum foil and stratified during 3-4 days at 4 °C in the darkness. After stratification, plants were placed for 10 days in a growing chamber at 21 °C with 16-8 hours of light/dark cycle at a light photo-intensity of 100 μmol · m^{-2} · s^{-1}. Four to six weeks old *Nicotiana benthamiana* (Nb1) plants were used for transient expressions. For the TEV infected plants, TEV inoculation was ordered at the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ) Institute. Freeze-dried plant material was placed in a precooled mortar and a few drops of inoculation buffer was added (50 mM sodium/potassium phosphate buffer adjusted to pH 7.0, 1 mM EDTA, 5 mM diethyldithiocarbamate, 5mM thioglycolic acid) and ground to a thick paste. Inoculation buffer is added to a final volume of 4 ml and rubbed in a maximum number of three plants. *Nicotiana tabacum* were used as host plants for the TEV inoculation, grown at 22-25°C in the greenhouse at a 14-10 hours of light/dark cycle. Healthy and infected plants were grown in the same chamber.

Cloning and transformation

We used GreenGate collection constructs (Lampropoulos et al., 2013). ZipGFP and TEVp original sequences were ordered from addgene (ZipGFP1-10_TEV:Plasmid #81242; ZipGFP11_TEV; Plasmid #81243; pcDNA3.1 TEV: Plasmid #64276). The DNA fragment corresponding to WW sequence was custom synthesized (Integrated DNA Technologies). Original caspase-3 plasmid (Casp3p30) was obtained from Prof. Savvas Savvides containing the p12 and p18 caspase 3 subunits rendering a total p30 size protein that is constitutively activated. To generate GreenGate compatible entry clones we used collection compatible entry plasmids or CloneJET PCR Cloning Kit

(Thermo Scientific) according to the manufacturer manual. Primer names and sequences used in this study can be found in Suppl. Table 1. Final plasmid construction by restriction and ligation was performed according to Lampropoulos *et al.*, (2103) using periods of 2 minutes for the ligation at 16 °C followed by enzymatic restriction during 2 minutes at 37°C for a total of 25 cycles. Ligation product was directly used for *E.coli* transformation and growing colonies were verified by colony PCR and Sanger sequencing. All the plasmid details and names are indicated in Suppl. Table 2. Plasmids were transformed in *Agrobacterium tumefaciens* C58C1 and selecting using the appropriate resistances.

N.benthamiana transient transformation

Agrobacterium tumefaciens strains were grown in flasks containing 10-20 ml YEB media with their corresponding resistances (Gentamicin, Rifampicin and Spectinomycin) at 28 ºC for 2-3 days with shaking. Bacterial concentration was estimated and the culture volume for a final OD_{600}=1.5 spinned down and resuspended in the infiltration buffer (10 mM $MgCl_2$, 10 mM MES buffer, 10 µM freshly added acetosyringone). Resuspended cells were incubated at 28 ºC for 2-4 hours and infiltrated in 4 weeks old leaves of *Nicotiana benthamiana* or *Nicotiana tabacum*. Confocal imaging and tissue sampling was done at day 2 or 3 after infiltration unless specified on a Zeiss 710 confocal setup.

Photo images

Photo images were taken in an adjustable custom-made stand with light bars using a Nikon Eos 650D photocamera.

Confocal imaging

Seedlings with good expression levels were selected to T3 generation and imaged using a Zeiss LSM 710 confocal microscope. Cell membranes of *Arabidopsis thaliana* root seedlings were stained with 10 µM propidium iodide (PI) five minutes prior to imaging. GFP expressing plants were imaged using a 488 nm laser and adjusting the filters for GFP emission and PI staining. In our hands, the Zeiss 488 nm laser induced less noise ratio when imaging mNeonGreen, than the 512 nm laser and therefore we used this laser for FRET detection. Detection of the signal and gain for each channel were maintained constant to compare images. Gain changes in one channel were corresponded for the equivalent change in the next channel. Images were analyzed

and edited using ImageJ free software. FRET ratio images were done as indicated for the donor and acceptor channels in Hander *et al.*, (2019).

DEVDase visualization with CR(DEVD)$_2$

For the validation of the caspase activity we used CR(DEVD)$_2$ CV-Caspase 3 & 7 detection kit (Enzo Life Sciences, BML-AK118-0001). CR(DEVD)$_2$ powder was soluted in 100 µl dimethyl sulfoxide and diluted 1:5 in milliQ H_2O prior to making a staining solution. Working solution was made by diluting 1:50 in milliQ H_2O. Samples were incubated 30 minutes at room temperature before imaging. For confocal imaging, an excitation filter of 550 nm (540-560 nm) and a long pass >610 nm emission/barrier filter were used.

Immunoblotting

Protein samples were loaded in a 12,5% SDS-PAGE gels in Tris-Glycine-SDS (TGS) buffer at a voltage of 200 V for 45 minutes approximately. Gels were transferred to a precast TransBlot Turbo membrane for 3 minutes at 2.5 amperes and voltages up to 25 V in a BioRad membrane transfer system. Membranes were blocked in 5% skim milk in PBS-T (0,2% Tween) for 2 hours. The blotting of GFP fluorescent-tagged proteins was done by 2 hours incubation in 1% skim milk with single step antibody (1:1,000) rabbit anti-GFP conjugated to HRP (Miltenyi) followed three washing steps for 10 minutes in PBS-T. Detection of the TEVp tagged with HA was blotted for 4 hours with anti-HA produced in rat (1:1,000; Roche) in 1% skim milk in PBS-T follow by overnight incubation at 4 ºC with anti-rat fused to HRP (1:10,000; GE Healthcare) in 1% PBS-T and 3 washes in PBS-T for 10 minutes duration. Detection was performed by exposition to a (1:1) solution of Western Lightning® Plus-ECL reagents (Perkin Elmer) for 1 minute. In the cases membranes were stripped and re-probed, we incubated the membrane for 10 minutes with a fresh solution of 10 mM Glycine pH 2,7 and SDS 10% (1:1), following 6 times washings for 5 minutes and re-starting by blocking for 2 hours using 5% skim milk PBS-T solution.

Trypan blue

Trypan blue staining was performed by immersing the plant material in a solution containing 0.5% trypan blue (Sigma-Aldrich) in ddH_2O, Glycerol ≥ 99% (Sigma-Aldrich) and lactic acid 90% (w/w) in equal relative volumes (1:1:1). Plant tissue in a

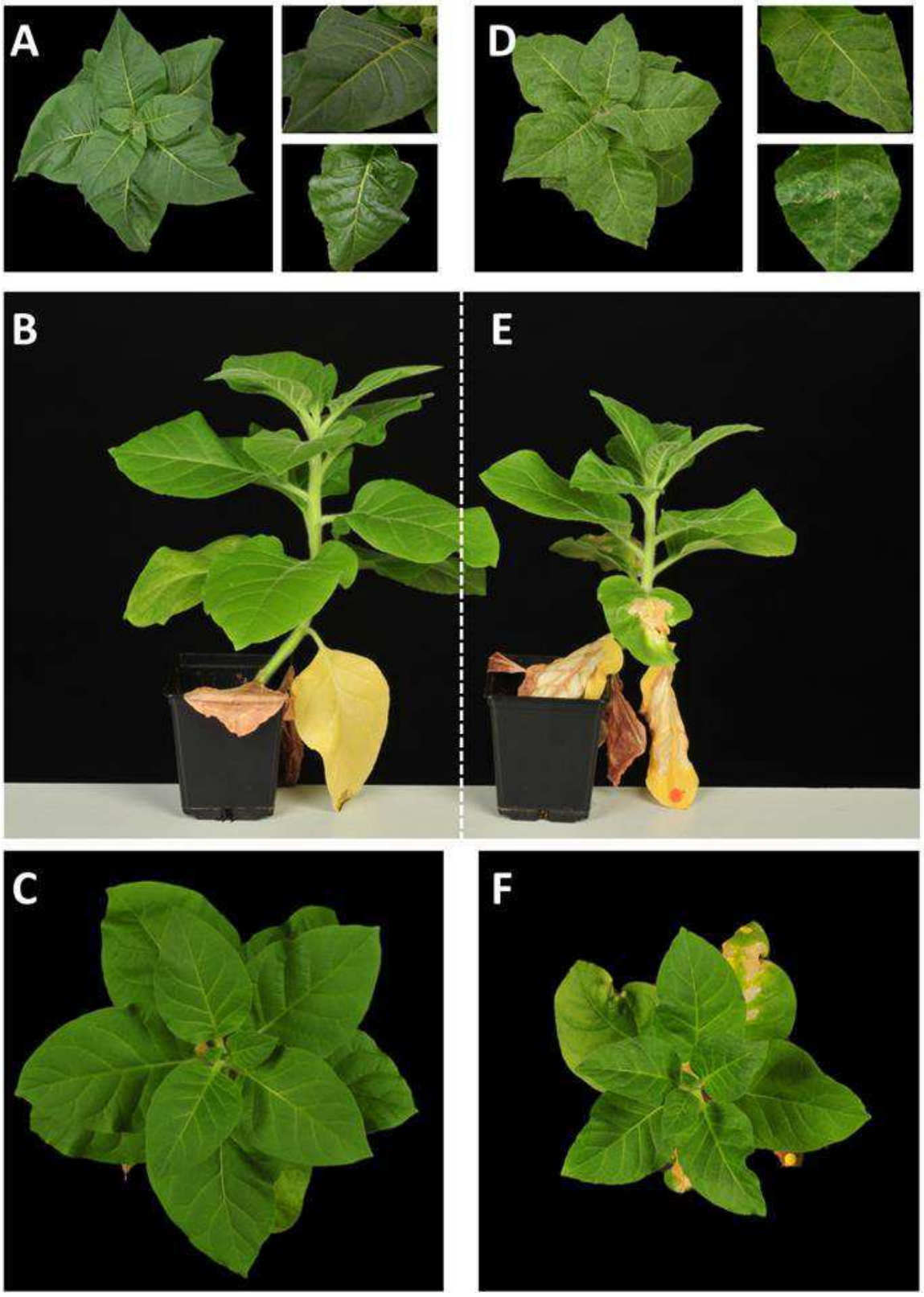
heat resistant container was heated in a conventional lab microwave for one minute following repetitive washing with PBS buffer.

Statistical Analysis

Statistical analysis was performed using GraphPad Prism 9.0 for Windows.

Acknowledgements

We thank Prof. Dr. Savvas Savvides for sharing with us the plasmid containing the His-Casp3p30 sequence. The original *Nicotiana benthamiana* Nb1 seed stock was a gift of Prof. Gregory B. Martin.



Suppl. Figure 1. Tobacco plants used for assessing TEV infection. A-C) Tobacco healthy plants. **D-F)** *Nicotiana tabacum* plants inoculated with TEV virus.

Oligo name	Sequence
GG-B-HA-FW	TTTTGGTCTCAAACACCACCATGGCATACCCTTACGATG
GG-C-HA-RV	TTTTGGTCTCAAGCCACTAGCATAATCAGGAACATCATAAGG
GG-TEV-FW	TTTTGGTCTCAGGCTCCACCATGGGCGAGAGCCTT
GG-D-TEV-ns-RV	TTTTGGTCTCTCTGAAAGCTGAGTGGCTTCCTTA
GG-E-TEV-s-RV	TTTTGGTCTCTGCAGTCAAAGCTGAGTGGCTTC
GG-C-6His-Casp3p30-FW	CCCCTTTTCCTCTCGGGCTCCATGGATCCACATCACCATCAC
GG-E-Casp3p30-s-RV	TTTTGGTCTCTGCAGTCAGTGATAAAAGTACAGTTCTTTCGTGA
GG-C-ZipGFP1-10-FW	TTTTGGTCTCAGGCTCCACCATGGAAGTGAGCGCGCTGGAAAAAGAAGTCTCCGCTCTT
GG-E-ZipGFP1-10-s-RV	TTTTGGTCTCTGCAGTTATTCTTTCAGCGCGCTCACTTTTTCCTTAAGGGCACTGAC
GG-E-mCherry-s-RV	TTTTGGTCTCTGCAGTTACTTGTACAGCTCGTCCATGCCGCCGG
GG-C-Zip11-FW	TTTTGGTCTCAGGCTCCACCATGGGCGGCAGCAAGG
GG-E-Zip11-s-RV	TTTTGGTCTCTGCAGTTATTTTTCAAGCGCAGAGACTTCCTT
GG-C-TEVrs-NES-GFP-FW	TTTTGGTCTCAGGCTCCGGAGAAAATTTGTATTTTCAAGGTATGTTGGAAAAGGATACTTTGTTGATTGATTTGCATGGAATGGTGAGCAAGGGCGAG
GG-E-GFP-s-RV	TTTTGGTCTCAGCAGCTACTTGTACAGCTCGTCCAT
GG2-WW-B-FW	CCCCTTTTGGTCTCAAACACCACCATGTTTGAAATTCCTGATGATGTTCCT
GG2-WW-C-RV	CCCCTTTTGGTCTCAAGCCCATTTGAGACAACATAGCCTTTCTA
GG-B-mNG-FW	CCCCGGTCTCAAACACCACCATGGTGTCCAAGGGCGA
GG-D-mNG-RV	CCCCGGTCTCTCTGAGGTGAAGGCTTTCTGCC
GG2-mNG-GS-linker-D-RV	CCCCGGTCTCTCTGATCCAGAGCCGCTGCCGGAACCACTCCCACTTCCGCCGCTTCCGGAACCACTCCCT
GG-D-mNG-DEVD-RV	TTTTGGTCTCTCTGATCCAGAGCCGCTGCCGGAACCACTCCCATCAACTTCATCGCTTCCGGAACCACTCCCT
OPT-GG-D-mNG-1xTEVrs-RV	TTTTGGTCTCTCTGATCCAGAGCCGCTGCCGGAACCACTTTGGAAGTACAAGTTTTCGCTTCCGGAACCACTCCCT
GG-D-mNG-VRPR-RV	TTTTGGTCTCTCTGATCCAGAGCCGCTGCCGGAACCACTCCCTCTAGGTCTAACGCTTCCGGAACCACTCCCT
GG-D-mNG-EKAKD-RV	TTTTGGTCTCTCTGATCCAGAGCCGCTGCCGGAACCACTCCCATCCTTAGCCTTTTCGCTTCCGGAACCACTCCCT
GG-D-mNG-EKTKD-RV	TTTTGGTCTCTCTGATCCAGAGCCGCTGCCGGAACCACTCCCATCCTTAGTCTTTTCGCTTCCGGAACCACTCCCT
GG-D-mNG-EKTRD-RV	TTTTGGTCTCTCTGATCCAGAGCCGCTGCCGGAACCACTCCCATCTCTAGTCTTTTCGCTTCCGGAACCACTCCCT
GG-D-mNG-PP1-PCS1	TTTTGGTCTCTCTGATCCAGAGCCGCTGCCGGAACCACTCCCGACCTTTGTGGCTCTACTCGTCACGCTTCCGGAACCACTCCCT
GG-D-mNG-PP1-PCS2	TGGTCTCTCTGATCCAGAGCCGCTGCCGGAACCACTCCCCCCCCTTTGCTTTGCCTTGACCTTTGTGGCTCTACTCGTCACGCTTCCGGAACCACTCCCT
GG-D-mRuby3-FW	TTTTGGTCTCATCAGGCATGGTGTCTAAGGGCGAAGAGCTGA
GG2-mRuby3-Wp1-s-GG-E-RV	CCCCGGTCTCAGCAGCTATCCAGTTCCAACAGTATAAGGAGGAGGAGGAGTTCCAGATCCCTTGTACAGCTCGTCCATGCCACC
mRuby3-Wp2-s-GG-E-RV	TTTTGGTCTCAGCAGCTATCCAGTATAAGTAGGAGGCAGAGGAGATCCCTTGTACAGCTCGTCCATGCCACC
GG-B-proAtMC9-RV	TTTTGGTCTCAAACTTGAGTTTCTGGATATGATTTCTTAA
GG-A-proAtMC9-FW	TTTTGGTCTCAAACTTTTTGGATGCTTCTCGTGTAAAATT

Supplementary table 1. Oligo names and sequences used in this study.

Construct name	Use	Resistance	Named on the chapter
p35S::3xHA-TEVp::G7t-HygR-pGGAG	Protease expression	Plant: HygR; Bact:Spec	TEVp
p35S::6xHis-Casp3p30::G7t-HygR-pGGAG	Protease expression	Plant: HygR; Bact:Spec	Casp3
pUbi10::ZipGFP1-10TEVrs-2A-mCherry::35St-SdR-pGGAG	Biosensor 1/2	Plant: SdR; Bact:Spec	ZipGFP1-10^{TEVrs}-2A-mCherry
pUbi10::ZipGFP11TEVrs::35St-BarR-pGGAG	Biosensor 2/2	Plant: BarR; Bact:Spec	ZipGFP11^{TEVrs}
pUbi10::NLSTEVrsNES-GFP::35St-SdR-pGGAG	Protease expression	Plant: SdR; Bact:Spec	NLS^{TEVrs}NES-GFP
p35S::WW-mNG-4xGGS-mR3-WP1::35St-KmR-pGGAG	Protease expression	Plant: KmR; Bact:Spec	$FRET^{6xGS}$
p35S::WW-mNG-DEVD-mR3-WP1::35St-KmR-pGGAG	Protease expression	Plant: KmR; Bact:Spec	$FRET^{DEVD}$
p35S::WW-mNG-1xTEVrs-mR3-WP1::35St-KmR-pGGAG	Protease expression	Plant: KmR; Bact:Spec	$FRET^{TEVrs}$
p35S::WW-mNG-VRPR-mR3-WP1::35St-KmR-pGGAG	Protease expression	Plant: KmR; Bact:Spec	$FRET^{VRPR}$
p35S::WW-mNG-EKAKD-mR3-WP1::35St-KmR-pGGAG	Protease expression	Plant: KmR; Bact:Spec	$FRET^{EKAKD}$
p35S::WW-mNG-EKTKD-mR3-WP1::35St-KmR-pGGAG	Protease expression	Plant: KmR; Bact:Spec	$FRET^{EKTKD}$
p35S::WW-mNG-EKTRD-mR3-WP1::35St-KmR-pGGAG	Protease expression	Plant: KmR; Bact:Spec	$FRET^{EKTRD}$
p35S::WW-mNG-PP1-PCS1-mR3-WP1::35St-KmR-pGGAG	Protease expression	Plant: KmR; Bact:Spec	$FRET^{PCS1}$
p35S::WW-mNG-PP1-PCS2-mR3-WP1::35St-KmR-pGGAG	Protease expression	Plant: KmR; Bact:Spec	$FRET^{PCS2}$
pAtMc9::NLS-3xGFP::35St-KmR-pGGAG	Expression reporter	Plant: KmR; Bact:Spec	*pATMC9*::NLS-3xEGFP

Supplementary table 2. Constructs used in this study. Entry modules were obtained from the GreenGate collection or cloned (See Suppl. table 1). HygR: indicates plant resistance to Hygromycin; BarR: indicates plant resistance to BASTA; SdR: indicates plant resistance to Sulfadiazine; KmR: indicates plant resistance to Kanamycin; Spec: indicates bacterial resistance to Spectinomycin.

REFERENCES

Ai, H.W., Baird, M.A., Shen, Y., Davidson, M.W., and Campbell, R.E. (2014). Engineering and characterizing monomeric fluorescent proteins for live-cell imaging applications. Nat Protoc **9:** 910-928.

Ast, C., Foret, J., Oltrogge, L.M., De Michele, R., Kleist, T.J., Ho, C.H., and Frommer, W.B. (2017). Ratiometric Matryoshka biosensors from a nested cassette of green- and orange-emitting fluorescent proteins. Nat Commun **8:** 431.

Bajar, B.T., Wang, E.S., Zhang, S., Lin, M.Z., and Chu, J. (2016). A Guide to Fluorescent Protein FRET Pairs. Sensors (Basel) **16**.

Bedhomme, S., and Elena, S.F. (2011). Virus infection suppresses Nicotiana benthamiana adaptive phenotypic plasticity. PLoS One **6:** e17275.

Bollhoner, B., Zhang, B., Stael, S., Denance, N., Overmyer, K., Goffner, D., Van Breusegem, F., and Tuominen, H. (2013). Post mortem function of AtMC9 in xylem vessel elements. New Phytol **200:** 498-510.

Bosch, M., and Franklin-Tong, V.E. (2007). Temporal and spatial activation of caspase-like enzymes induced by self-incompatibility in Papaver pollen. Proc Natl Acad Sci U S A **104:** 18327-18332.

Buono, R.A., Hudecek, R., and Nowack, M.K. (2019). Plant proteases during developmental programmed cell death. J Exp Bot **70:** 2097-2112.

Carrington, J.C., and Dougherty, W.G. (1988). A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing. Proc Natl Acad Sci U S A **85:** 3391-3395.

De Col, V., et al. (2017). ATP sensing in living plant cells reveals tissue gradients and stress dynamics of energy physiology. Elife **6**.

DeFalco, T.A., Toyota, M., Phan, V., Karia, P., Moeder, W., Gilroy, S., and Yoshioka, K. (2017). Using GCaMP3 to Study Ca2+ Signaling in Nicotiana Species. Plant Cell Physiol **58:** 1173-1184.

Denay, G., Schultz, P., Hansch, S., Weidtkamp-Peters, S., and Simon, R. (2019). Over the rainbow: A practical guide for fluorescent protein selection in plant FRET experiments. Plant Direct **3:** e00189.

Duvall, S.W., and Childers, W.S. (2020). Design of a Histidine Kinase FRET Sensor to Detect Complex Signal Integration within Living Bacteria. ACS Sens **5:** 1589-1596.

Escamez, S., Andre, D., Sztojka, B., Bollhoner, B., Hall, H., Berthet, B., Voss, U., Lers, A., Maizel, A., Andersson, M., Bennett, M., and Tuominen, H. (2020). Cell Death in Cells Overlying Lateral Root Primordia Facilitates Organ Growth in Arabidopsis. Curr Biol **30:** 455-464 e457.

Exposito-Rodriguez, M., Laissue, P.P., Yvon-Durocher, G., Smirnoff, N., and Mullineaux, P.M. (2017). Photosynthesis-dependent H2O2 transfer from chloroplasts to nuclei provides a high-light signalling mechanism. Nat Commun **8:** 49.

Grefen, C., and Blatt, M.R. (2012). A 2in1 cloning system enables ratiometric bimolecular fluorescence complementation (rBiFC). Biotechniques **53:** 311-314.

Grossmann, G., Krebs, M., Maizel, A., Stahl, Y., Vermeer, J.E.M., and Ott, T. (2018). Green light for quantitative live-cell imaging in plants. J Cell Sci **131**.

Grunberg, R., Burnier, J.V., Ferrar, T., Beltran-Sastre, V., Stricher, F., van der Sloot, A.M., Garcia-Olivas, R., Mallabiabarrena, A., Sanjuan, X., Zimmermann, T., and Serrano,

L. (2013). Engineering of weak helper interactions for high-efficiency FRET probes. Nat Methods **10:** 1021-1027.

Guerreiro, M.R., Fernandes, A.R., and Coroadinha, A.S. (2020). Evaluation of Structurally Distorted Split GFP Fluorescent Sensors for Cell-Based Detection of Viral Proteolytic Activity. Sensors (Basel) **21**.

Han, H.E., et al. (2010). The nuclear inclusion a (NIa) protease of turnip mosaic virus (TuMV) cleaves amyloid-beta. PLoS One **5:** e15645.

Hander, T., et al. (2019). Damage on plants activates Ca(2+)-dependent metacaspases for release of immunomodulatory peptides. Science **363**.

Hidaka, M., Yoshida, Y., Masaki, H., Namba, S., Yamashita, S., Tsuchizaki, T., and Uozumi, T. (1992). Cloning and sequencing of the 3' half of a potato virus Y (O strain) genome encoding the 5k protein, protease, polymerase and coat protein. Nucleic Acids Res **20:** 3515.

Jones, A.M., Danielson, J.A., Manojkumar, S.N., Lanquar, V., Grossmann, G., and Frommer, W.B. (2014). Abscisic acid dynamics in roots detected with genetically encoded FRET sensors. Elife **3:** e01741.

Jones, R.A.C. (2021). Global Plant Virus Disease Pandemics and Epidemics. Plants (Basel) **10**.

Kostyuk, A.I., Demidovich, A.D., Kotova, D.A., Belousov, V.V., and Bilan, D.S. (2019). Circularly Permuted Fluorescent Protein-Based Indicators: History, Principles, and Classification. Int J Mol Sci **20**.

Kudla, J., and Bock, R. (2016). Lighting the Way to Protein-Protein Interactions: Recommendations on Best Practices for Bimolecular Fluorescence Complementation Analyses. Plant Cell **28:** 1002-1008.

Lam, A.J., St-Pierre, F., Gong, Y., Marshall, J.D., Cranfill, P.J., Baird, M.A., McKeown, M.R., Wiedenmann, J., Davidson, M.W., Schnitzer, M.J., Tsien, R.Y., and Lin, M.Z. (2012). Improving FRET dynamic range with bright green and red fluorescent proteins. Nat Methods **9:** 1005-1012.

Lampropoulos, A., Sutikovic, Z., Wenzl, C., Maegele, I., Lohmann, J.U., and Forner, J. (2013). GreenGate---a novel, versatile, and efficient cloning system for plant transgenesis. PLoS One **8:** e83043.

Larrieu, A., et al. (2015). A fluorescent hormone biosensor reveals the dynamics of jasmonate signalling in plants. Nat Commun **6:** 6043.

Liao, C.Y., Smet, W., Brunoud, G., Yoshida, S., Vernoux, T., and Weijers, D. (2015). Reporters for sensitive and quantitative measurement of auxin response. Nat Methods **12:** 207-210, 202 p following 210.

Lin, Z., Xie, F., Trivino, M., Karimi, M., Bosch, M., Franklin-Tong, V.E., and Nowack, M.K. (2020). Ectopic Expression of a Self-Incompatibility Module Triggers Growth Arrest and Cell Death in Vegetative Cells. Plant Physiol **183:** 1765-1779.

McCullock, T.W., MacLean, D.M., and Kammermeier, P.J. (2020). Comparing the performance of mScarlet-I, mRuby3, and mCherry as FRET acceptors for mNeonGreen. PLoS One **15:** e0219886.

Monshausen, G.B., Messerli, M.A., and Gilroy, S. (2008). Imaging of the Yellow Cameleon 3.6 indicator reveals that elevations in cytosolic Ca2+ follow oscillating increases in growth in root hairs of Arabidopsis. Plant Physiol **147:** 1690-1698.

Morimoto, K., and van der Hoorn, R.A. (2016). The Increasing Impact of Activity-Based Protein Profiling in Plant Science. Plant Cell Physiol **57:** 446-461.

Moseyko, N., and Feldman, L.J. (2001). Expression of pH-sensitive green fluorescent protein in Arabidopsis thaliana. Plant Cell Environ **24:** 557-563.

Nietzel, T., et al. (2019). The fluorescent protein sensor roGFP2-Orp1 monitors in vivo H2 O2 and thiol redox integration and elucidates intracellular H2 O2 dynamics during elicitor-induced oxidative burst in Arabidopsis. New Phytol **221:** 1649-1664.

Oliveira-Silva, R., Sousa-Jeronimo, M., Botequim, D., Silva, N.J.O., Paulo, P.M.R., and Prazeres, D.M.F. (2020). Monitoring Proteolytic Activity in Real Time: A New World of Opportunities for Biosensors. Trends Biochem Sci **45:** 604-618.

Salvesen, G.S., Hempel, A., and Coll, N.S. (2016). Protease signaling in animal and plant-regulated cell death. FEBS J **283:** 2577-2598.

Shaner, N.C., Steinbach, P.A., and Tsien, R.Y. (2005). A guide to choosing fluorescent proteins. Nat Methods **2:** 905-909.

Stennicke, H.R., Renatus, M., Meldal, M., and Salvesen, G.S. (2000). Internally quenched fluorescent peptide substrates disclose the subsite preferences of human caspases 1, 3, 6, 7 and 8. Biochem J **350 Pt 2:** 563-568.

Sueldo, D.J., and van der Hoorn, R.A.L. (2017). Plant life needs cell death, but does plant cell death need Cys proteases? FEBS J **284:** 1577-1585.

To, T.L., Schepis, A., Ruiz-Gonzalez, R., Zhang, Q., Yu, D., Dong, Z., Coughlin, S.R., and Shu, X. (2016). Rational Design of a GFP-Based Fluorogenic Caspase Reporter for Imaging Apoptosis In Vivo. Cell Chem Biol **23:** 875-882.

Toyota, M., Spencer, D., Sawai-Toyota, S., Jiaqi, W., Zhang, T., Koo, A.J., Howe, G.A., and Gilroy, S. (2018). Glutamate triggers long-distance, calcium-based plant defense signaling. Science **361:** 1112-1115.

Tsiatsiani, L., Timmerman, E., De Bock, P.J., Vercammen, D., Stael, S., van de Cotte, B., Staes, A., Goethals, M., Beunens, T., Van Damme, P., Gevaert, K., and Van Breusegem, F. (2013). The Arabidopsis metacaspase9 degradome. Plant Cell **25:** 2831-2847.

Ullrich, S., and Nitsche, C. (2020). The SARS-CoV-2 main protease as drug target. Bioorg Med Chem Lett **30:** 127377.

van de Plassche, M.A.T., Barniol-Xicota, M., and Verhelst, S.H.L. (2020). Peptidyl Acyloxymethyl Ketones as Activity-Based Probes for the Main Protease of SARS-CoV-2*. Chembiochem **21:** 3383-3388.

van der Hoorn, R.A. (2008). Plant proteases: from phenotypes to molecular mechanisms. Annu Rev Plant Biol **59:** 191-223.

Vercammen, D., van de Cotte, B., De Jaeger, G., Eeckhout, D., Casteels, P., Vandepoele, K., Vandenberghe, I., Van Beeumen, J., Inze, D., and Van Breusegem, F. (2004). Type II metacaspases Atmc4 and Atmc9 of Arabidopsis thaliana cleave substrates after arginine and lysine. J Biol Chem **279:** 45329-45336.

Waadt, R., Hitomi, K., Nishimura, N., Hitomi, C., Adams, S.R., Getzoff, E.D., and Schroeder, J.I. (2014). FRET-based reporters for the direct visualization of abscisic acid concentration changes and distribution in Arabidopsis. Elife **3:** e01739.

Ward, B.M., and Lazarowitz, S.G. (1999). Nuclear export in plants. Use of geminivirus movement proteins for a cell-based export assay. Plant Cell **11:** 1267-1276.

Woo, J., Park, E., and Dinesh-Kumar, S.P. (2014). Differential processing of Arabidopsis ubiquitin-like Atg8 autophagy proteins by Atg4 cysteine proteases. Proc Natl Acad Sci U S A **111:** 863-868.

Zaman, N., Seitz, K., Kabir, M., George-Schreder, L.S., Shepstone, I., Liu, Y., Zhang, S., and Krysan, P.J. (2019). A Forster resonance energy transfer sensor for live-cell imaging of mitogen-activated protein kinase activity in Arabidopsis. Plant J **97:** 970-983.

Zhang, L., Xu, Q., Xing, D., Gao, C., and Xiong, H. (2009). Real-time detection of caspase-3-like protease activation in vivo using fluorescence resonance energy transfer during plant programmed cell death induced by ultraviolet C overexposure. Plant Physiol **150:** 1773-1783.

Zhang, Q., et al. (2019). Designing a Green Fluorogenic Protease Reporter by Flipping a Beta Strand of GFP for Imaging Apoptosis in Animals. J Am Chem Soc **141:** 4526-4530.

Discussion and perspectives

Plant proteases – Exploring a one-way post translational modification

Post-translational modifications (PTMs) can be conceived as a set of mechanisms that enable to enlarge the diversity of proteins. Through PTMs, the functions of proteins can be rapidly adjusted to the requirements of the cell (Walsh et al., 2005; Arsova et al., 2018). PTMs can convey changes in the protein function, localization and interaction partners. Moreover, many of the PTMs can be very dynamic, by action of a "writer" and an "eraser", that allows addition or removal of moieties. Many PTMs have been studied, being phosphorylation and ubiquitination the most studied. On the other hand, many PTMs remain to be studied or even discovered. Proteolysis and its main actors, proteases, are in between two fronts, on one side, they are mainly studied due to their importance in apoptosis and disease signalling in mammalians, but there are plenty of proteases with unknown functions, especially in other fields than animal cell death (Sukharev et al., 1997; Solary et al., 1998).

Proteases are hydrolytic enzymes able to process peptidyl bonds in a thermodynamically favourable and unidirectional way. Traditionally, proteases have been perceived mostly as enzymes that destroy other proteins. However, they can be fine modulators of multiple cellular activities. In plants, proteases are able to contribute to epidermal cell fate, signal abscission processes, contribute to plant pathogen interaction, or aid during fertilization (Chapter 2). Due to the irreversibility of protease substrate cleavage, proteolytic activity is finely controlled. Among plant proteases, metacaspases are arguably the best studied and reaching several milestones in respect to understand their activation mechanisms (Vercammen et al., 2004; Watanabe and Lam, 2011a, b; Hander et al., 2019; Zhu et al., 2020). The implementation of omics tools have helped the identification of their substrates. For instance, the use of COFRADIC on AtMC9 functions resulted in the identification of substrates with importance in seed imbibition and subsequent early stages of seedling growth (Tsiatsiani et al., 2013; Rojas et al., 2020).

Wound activated proteases – Exploring a one-way post-translational modification

In Arabidopsis, AtMC4 is the most abundant metacaspase according to its expression patterns (Watanabe and Lam, 2011b; Bollhoner et al., 2013) and it was shown to be activated *in vitro* by calcium availability (Watanabe and Lam, 2011a; Hander et al., 2019). In chapter 4 of this thesis, we characterized the activation mechanisms of AtMC4 *in vivo* using wounding as a triggering mechanism. Wounding has been known for some time to be an stimuli able to induce calcium waves in plants (Leon et al., 2001) and PROPEP1 was reported an elicitor peptide in plant damage conditions (Huffaker et al., 2006). We portrayed the events following wound-

induced calcium released, including the subsequent metacaspase activation and PROPEP processing. These consecutive events occur almost immediately, and despite of the simplicity, it has been the first proteolytic mechanism illustrating proteolysis from its direct activation to substrate processing in plants. Our observations anticipated later studies on the PROPEP processing (Shen et al., 2019) and on a more detailed explanation of the molecular changes of AtMC4 by calcium presence, based on direct visualizations of calcium binding and structural rearrangements in a AtMC4 crystal structure (Zhu et al., 2020). The local study of wound signal, calcium availability and tonoplast release of PROPEP1 in surrounding regions to the damage has opened new questions. One would be what is the effect of each metacaspase, and if they have compensatory activity. For instance, knock-outs of the tandem repeat of type-II metacaspases showed a phenotype in Arabidopsis plants infected with *Botrytis cinerea* (Shen et al., 2019). Other questions to frame are “How are the calcium dynamics affecting AtMC4 activation?”; “Are there other stimuli different to wounding able to induce metacaspase activity?”; “How does this influence to PROPEP family members processing and their biological importance?”.

It is also possible that other proteases are modulated during wounding events and that those might work in parallel or synergistically with metacaspases. In damaged cells, multiple signals are elicited in a local or systemic manner (Tanaka and Heil, 2021). For example, calcium and ROS travel at different speeds and through different tissues in wounded plants (Farmer et al., 2020; Vega-Munoz et al., 2020). Screening studies with the use of activity-based probes would reveal the class of proteases active at different time points after wounding. Additionally, peptidomics and N-terminomic studies could reveal processed peptides in wound stresses. Therefore, I predict that these unanswered questions on the dynamics of activated proteases will go hand-in-hand with studies combining damage and biotic stresses. In fact, later studies have focused on the crosstalk of wounding with hormone pathways like jasmonate and ethylene in herbivore feeding (Marhavy et al., 2019; Zhou et al., 2019), bacteria colonization (Zhou et al., 2020) and crossroads with root regeneration (Canher et al., 2020; Hoermayer et al., 2020).

Hunting the targets - omics techniques for the detection of protease substrates

Understanding the substrates and activation of proteases are the main challenges of the field. Traditionally, substrate identification could be achieved by gel based methods (Agard and Wells, 2009). Several high throughput proteomic techniques have emerged in order to assist in the identification of protease substrates (Van Damme et al., 2005; Bogaert and Gevaert, 2020). In plants, a modest number of N-terminomic experiments have been published recently.

For instance, COFRADIC has been used for the detection of AtMC9 substrates in Arabidopsis seedlings (Tsiatsiani et al., 2013), CHAFRADIC has revealed a number above two thousand N-terminal peptides (Venne et al., 2015) or TAILS was exploited in the identification of N-End Rule Pathway components (Zhang et al., 2018). The identification of some protease substrates has led to further investigation on the targets. For example, initial proteolytic processing of PEPCK1 identified in (Tsiatsiani et al., 2013) later has been studied in detail in relation to nutrient remobilization in seeds (Rojas et al, 2020). This illustrates that N-terminomic experiments can serve to neighbour fields studying proteostasis and shows the potential value to employ the data in further research.

In chapter 5, I used an adapted version of COFRADIC for the identification of AtMC4 substrates in Arabidopsis roots using wounded versus control plants in wild type or mutant background. Additionally, we combined the N-terminomics with shotgun proteomics, in four biological repeats for each sample. This multi proteomic study is the first study combining both techniques in plants. In numbers, it allowed the identification of a number close to two thousand N-terminal peptides and eight thousand peptides for the shotgun proteomics. We mainly focused on the study of peptides processed after arginine or lysine, due to their relevance as possible metacaspase substrates. Some of the AtMC4 candidate substrates are currently under investigation for assessment of their processing in metacaspase triggering conditions. Among candidate substrates we identified PYK10 and NAI2, ER localized proteins that are processed after wounding. PYK10 displays an archetypical ER retaining signal KDEL at its C-termini and processing upstream this motif may affect its localization. On the other hand NAI2 share localization with PYK10 but does not possess a KDEL signal. AtMC4 localization was predicted to be nucleo - cytoplasmic, similar to AtMC9, and cleavage of ER located substrates is not totally understood. PROPEP1 is also allocated in a membranous organelle of the cell, and this substrate – protease partial segregation might work as a mechanism to minimize their encounter and processing. Ideally, confirmation of NAI2 and PYK10 processing, and destiny of their processed halves after AtMC4 processing should be examined in detail. Also, considering the involvement of PYK10 and NAI2 in plant defence processes (Yamada et al., 2011), experiments with insect feeding in parallel with measurements of glucosinolate levels seem as future research direction. Processing of substrates at different dynamics was observed like GRP7 cleavage by AtMC4. While AtMC4 is expected to be activated at early times after wounding it cannot be discarded that the protease – substrate affinities can change depending on their interaction.

The presence of some proteinaceous inhibitors of proteases results interesting due to their regulatory functions on proteases. A heavily labelled peptide matching Kunitz Trypsin Inhibitor (KTI1) protein sequence was found significant in the 2-way ANOVA and in wound versus

control of wild type samples. However, the peptide sequence did not follow the typical metacaspase footprint of displaying an arginine or lysine at P1. Also, peptides for cystatin 1 (CYS1) and cystatin B (CYSB) were found, but did not reach statistical significance. Cystatin are ubiquitous cysteine inhibitors with different roles across plant species (Benchabane et al., 2010) and CYS1 was identified in the AtMC9 N-terminome. Individual studies on their processing and if they can bind or affect activity of AtMC4 will be performed in the future.

The application of the N-terminome techniques has resulted useful for the identification of protease substrates or degradation targets (Tsiatsiani et al., 2013; Zhang et al., 2018). Knowing the triggering conditions modifying the N-terminome is essential for the successful identification of the protease targets. For instance, and as mentioned above, for AtMC4 wounding suffices for its proteolytic activation. The use of mutants of different order or activity depletion by inhibitors can also provide good control conditions to compare to proteolytic active samples. In addition, the use of recombinant proteases (*in vitro*) in proteome extracts can also be useful and might help to minimize the variability found by the mutant rather than by the protease activity. Targets identified in more than one approach (for example *in vivo* and *in vitro*) are likely more robust substrates. Enrichment of neo N-termini combined with mass spectrometry identification of substrates has many advantages, from the identification of multiple substrates in one run, to the identification of low abundant peptides. Nevertheless, it also has some drawbacks. For instance, it is likely that not all the protease substrates can be identified using this approach. This is exemplified by the lack of Pep1 or its homologs in the AtMC4 N-terminome and could be explained by unsuitable properties of the digested peptide in the mass spectrometer. In order to enhance the identification of substrates, different digesting enzymes could be used, generating different peptide versions but revealing the same processing position. This could enhance the variability of peptides and their suitability for detection.

Spotting the cut – spatial dynamics of proteolysis

Another major unexplored aspect of proteolysis in plants is the detection of the enzymatic activity *in situ*. This task can be performed with the use of individualized biosensors. However, it is necessary to obtain a minimal knowledge on the enzymatic activity of proteases and substrate preference to be able to engineer the protease substrate cleavage site of the biosensor. Thus, N-terminomics, substrate profiling assays and inhibitor evaluation are complementary approaches to biosensor development. MS-based techniques provide information for substrate preference and might suffice for the engineering of specific probes. The use of biosensors, *in vivo* tissues or fixated samples *in situ* (See chapter 3) can deliver

information that cannot be obtained with other techniques. For instance, biosensors can reveal substrate processing providing with real time information of the spatial-temporal proteolysis. Moreover, they allow to screen for proteolysis-inducing conditions, and allow for a detail characterization of the events. In chapter 6, I implemented and generated some tools for the detection of protease activity in plants. I aimed to adapt the FRET sensor for the detection of metacaspase activity using previous information of the known substrates, resulting in drop of ratio in location and conditions where metacaspases are active. I predict that change on the fluorescent protein and use of other processing sites known for plant proteases will generate highly helpful tools in the characterization of proteolytic activity in plants.

In summary, during my PhD I have been able to experiment and learn about plant proteases, focusing on metacaspase activity. I have acquired knowledge of techniques used for the identification of protease substrates and imaging techniques that can be complementary used in the study of proteolysis in plants. PROPEP1 identification as a biological substrate of AtMC4 in wounding, resulted in a better understanding of the dynamics of AtMC4 opening new paths for the detection of other metacaspase substrates. Additionally, I have been able to perform proteomic techniques with the integration of N-terrminomics and shotgun proteomics. This study allowed to identify novel AtMC4 candidate substrates in wound events. Lastly, I have implemented, optimized and engineered fluorescent reporters for the detection of proteolysis *in vivo*. I have adapted the sensors for the detection of metacaspase activity and followed their dynamics showing very promising results. Altogether, I expect that my contribution to the field of plant proteases, concretely in the metacaspase topic, will help to decipher novel activation mechanisms, relevant substrates and will inspire colleague researcher to further investigate this exciting topic.

REFERENCES

Agard, N.J., and Wells, J.A. (2009). Methods for the proteomic identification of protease substrates. Curr Opin Chem Biol **13:** 503-509.

Arsova, B., Watt, M., and Usadel, B. (2018). Monitoring of Plant Protein Post-translational Modifications Using Targeted Proteomics. Front Plant Sci **9:** 1168.

Benchabane, M., Schluter, U., Vorster, J., Goulet, M.C., and Michaud, D. (2010). Plant cystatins. Biochimie **92:** 1657-1666.

Bogaert, A., and Gevaert, K. (2020). Protein amino-termini and how to identify them. Expert Rev Proteomics **17:** 581-594.

Bollhoner, B., Zhang, B., Stael, S., Denance, N., Overmyer, K., Goffner, D., Van Breusegem, F., and Tuominen, H. (2013). Post mortem function of AtMC9 in xylem vessel elements. New Phytol **200:** 498-510.

Canher, B., Heyman, J., Savina, M., Devendran, A., Eekhout, T., Vercauteren, I., Prinsen, E., Matosevich, R., Xu, J., Mironova, V., and De Veylder, L. (2020). Rocks in the auxin stream: Wound-induced auxin accumulation and ERF115 expression synergistically drive stem cell regeneration. Proc Natl Acad Sci U S A **117:** 16667-16677.

Farmer, E.E., Gao, Y.Q., Lenzoni, G., Wolfender, J.L., and Wu, Q. (2020). Wound- and mechanostimulated electrical signals control hormone responses. New Phytol **227:** 1037-1050.

Hander, T., et al. (2019). Damage on plants activates Ca(2+)-dependent metacaspases for release of immunomodulatory peptides. Science **363**.

Hoermayer, L., Montesinos, J.C., Marhava, P., Benkova, E., Yoshida, S., and Friml, J. (2020). Wounding-induced changes in cellular pressure and localized auxin signalling spatially coordinate restorative divisions in roots. Proc Natl Acad Sci U S A **117:** 15322-15331.

Huffaker, A., Pearce, G., and Ryan, C.A. (2006). An endogenous peptide signal in Arabidopsis activates components of the innate immune response. Proc Natl Acad Sci U S A **103:** 10098-10103.

Leon, J., Rojo, E., and Sanchez-Serrano, J.J. (2001). Wound signalling in plants. J Exp Bot **52:** 1-9.

Marhavy, P., Kurenda, A., Siddique, S., Denervaud Tendon, V., Zhou, F., Holbein, J., Hasan, M.S., Grundler, F.M., Farmer, E.E., and Geldner, N. (2019). Single-cell damage elicits regional, nematode-restricting ethylene responses in roots. EMBO J **38**.

Rojas, B.E., Hartman, M.D., Figueroa, C.M., and Iglesias, A.A. (2020). Proteolytic cleavage of Arabidopsis thaliana phosphoenolpyruvate carboxykinase-1 modifies its allosteric regulation. J Exp Bot.

Shen, W., Liu, J., and Li, J.F. (2019). Type-II Metacaspases Mediate the Processing of Plant Elicitor Peptides in Arabidopsis. Mol Plant **12:** 1524-1533.

Solary, E., Eymin, B., Droin, N., and Haugg, M. (1998). Proteases, proteolysis, and apoptosis. Cell Biol Toxicol **14:** 121-132.

Sukharev, S.A., Pleshakova, O.V., and Sadovnikov, V.B. (1997). Role of proteases in activation of apoptosis. Cell Death Differ **4:** 457-462.

Tanaka, K., and Heil, M. (2021). Damage-Associated Molecular Patterns (DAMPs) in Plant Innate Immunity: Applying the Danger Model and Evolutionary Perspectives. Annu Rev Phytopathol.

Tsiatsiani, L., Timmerman, E., De Bock, P.J., Vercammen, D., Stael, S., van de Cotte, B., Staes, A., Goethals, M., Beunens, T., Van Damme, P., Gevaert, K., and Van Breusegem, F. (2013). The Arabidopsis metacaspase9 degradome. Plant Cell **25:** 2831-2847.

Van Damme, P., Martens, L., Van Damme, J., Hugelier, K., Staes, A., Vandekerckhove, J., and Gevaert, K. (2005). Caspase-specific and nonspecific in vivo protein processing during Fas-induced apoptosis. Nat Methods **2:** 771-777.

Vega-Munoz, I., Duran-Flores, D., Fernandez-Fernandez, A.D., Heyman, J., Ritter, A., and Stael, S. (2020). Breaking Bad News: Dynamic Molecular Mechanisms of Wound Response in Plants. Front Plant Sci **11:** 610445.

Venne, A.S., Solari, F.A., Faden, F., Paretti, T., Dissmeyer, N., and Zahedi, R.P. (2015). An improved workflow for quantitative N-terminal charge-based fractional diagonal chromatography (ChaFRADIC) to study proteolytic events in Arabidopsis thaliana. Proteomics **15:** 2458-2469.

Vercammen, D., van de Cotte, B., De Jaeger, G., Eeckhout, D., Casteels, P., Vandepoele, K., Vandenberghe, I., Van Beeumen, J., Inze, D., and Van Breusegem, F. (2004). Type II metacaspases Atmc4 and Atmc9 of Arabidopsis thaliana cleave substrates after arginine and lysine. J Biol Chem **279:** 45329-45336.

Walsh, C.T., Garneau-Tsodikova, S., and Gatto, G.J., Jr. (2005). Protein posttranslational modifications: the chemistry of proteome diversifications. Angew Chem Int Ed Engl **44:** 7342-7372.

Watanabe, N., and Lam, E. (2011a). Calcium-dependent activation and autolysis of Arabidopsis metacaspase 2d. J Biol Chem **286:** 10027-10040.

Watanabe, N., and Lam, E. (2011b). Arabidopsis metacaspase 2d is a positive mediator of cell death induced during biotic and abiotic stresses. Plant J **66:** 969-982.

Yamada, K., Hara-Nishimura, I., and Nishimura, M. (2011). Unique defense strategy by the endoplasmic reticulum body in plants. Plant Cell Physiol **52:** 2039-2049.

Zhang, H., Gannon, L., Hassall, K.L., Deery, M.J., Gibbs, D.J., Holdsworth, M.J., van der Hoorn, R.A.L., Lilley, K.S., and Theodoulou, F.L. (2018). N-terminomics reveals control of Arabidopsis seed storage proteins and proteases by the Arg/N-end rule pathway. New Phytol **218:** 1106-1126.

Zhou, F., Emonet, A., Denervaud Tendon, V., Marhavy, P., Wu, D., Lahaye, T., and Geldner, N. (2020). Co-incidence of Damage and Microbial Patterns Controls Localized Immune Responses in Roots. Cell **180:** 440-453 e418.

Zhou, W., Lozano-Torres, J.L., Blilou, I., Zhang, X., Zhai, Q., Smant, G., Li, C., and Scheres, B. (2019). A Jasmonate Signaling Network Activates Root Stem Cells and Promotes Regeneration. Cell **177:** 942-956 e914.

Zhu, P., Yu, X.H., Wang, C., Zhang, Q., Liu, W., McSweeney, S., Shanklin, J., Lam, E., and Liu, Q. (2020). Structural basis for Ca(2+)-dependent activation of a plant metacaspase. Nat Commun **11:** 2249.

Acknowledgements

I had the pleasure to do my PhD in PSB for five years surrounded by great colleagues and friends. I have tried to compile most of you but it is possible I forgot some of you. A Spanish saying dictates “Ni están todos los que son, ni son todos los que están”, and if I am forgetting someone or someone included does not feel represented, you can complain in headquarters. I would like to render thanks to many OSSers members: to Barbara and Valerie for their patience from the first day and for showing me where the green labgoats should go (n>5), to Huaming for introductory lessons to Chinese culture, to Takanori for the nice talks over 1,000 beers, to the hot pot dragon team (Xiaopeng, Xingliang and Pangfeng) and Robin for the time in the lab and numerous hours in the kicker. To Pavel for showing me the Canary-island way to sweeten “academia” on Friday evenings. Also to the rest of OSS members for helping so much day after day and for the collaborative environment while working so hard: Kai, Su, Inge, Jonas, Bo, Xi, Amna, Patrick, JingJing, Ola, Zhicheng and Zeya… and to my supervisors Frank and Simon for giving me the opportunity to do this PhD and share your ideas and thanks to my students for reminding me of the spark of science.

To the PSB Spanish community members: a la pequeña y vieja Armada Española, al capitán Palací portador del estandarte de la Santa paella, la comandante Cubría gobernadora del pica-pica a la hora del comino y a la almirante Triviño, protectora del secreto de la fina loncha de ibérico extremeño. También mis reconocimientos al incansable marinero de agua dulce el teniente van der Meer, que a pesar de su pasado y presente en las provincias es ahora aceptado entre los Grande de España. A las incorporaciones de media temporada y refuerzos latinoamericanos, algunos de los que se acaban de ir y a los que no saben lo que les queda: a Javi, Nacho, Mario, Rubén, Christian, Jose Antonio y Jhon Mario. Gracias por hacerme la vida más fácil, por las comidas y los viajes; a Piti y Paloma, al ya mítico Wilson por sus memorables lemas como “demasiadas reglas debería estar regulado”. There are many more people I would like to thank for their help and kindness and constructive conversations, writing your names here won´t honour your impact to my time here: Jonathan Dragwidge, Matthias Van Durme and others I really enjoyed talking about science in the corridor or in the cubes when this was possible. Thanks to Alexander Jones for meeting me in

the ASPB mentoring program, your words and advice really kept me afloat when I most needed it and to Pepe Gutiérrez Marcos for introducing me to the plant models.
Last to the members of CAMPUS team that shared many sport evenings with me, from our ascension from the hells of the 2nd division to fight for the 1st league football title: to Federico “*Il Dottore*”, Paulo “Neymar JR”, Mattias “*The Gigant*”, Carlos “*El Pipa*” and our favourite referee for watching our stylish *catenaccio* football every week.

A mi familia y en especial a mis padres Ana y Vicente, por darme la oportunidad de estudiar y poder desarrollarme como persona, además de por todo el apoyo incondicional a pesar de los momentos difíciles y a mi hermana Laura por introducirme en el mundo de la biología y por estar siempre disponible a pesar de la distancia. Gracias a Lis y su familia por acogerme en Bélgica y sobre todo a Lis, que junto a sus *lap assistants* Junie y Kous, me han escuchado y aguantado durante este arduo camino. También gracias por escuchar todas las charlas sobre ciencia y sobre los innumerables problemas que acaban teniendo solución.

CURRICULUM VITAE

PERSONAL INFORMATION

Name: Álvaro Daniel
Surname: Fernández-Fernández
Work address: OSS Lab, VIB-UGent Plant Systems Biology, Gent (Belgium)
Twitter @alvaroffer
e-mail: alvaro.ffer@gmail.com

EDUCATION

2015 – *2021*	PhD candidate in Plant Biotechnology OSS Lab, Plant Systems Biology VIB-University of Ghent (Belgium)
2014 – 2015	Master (ManaMa) in Plant Biotechnology (Graduated with *hons*) University of Ghent (Belgium)
2011 – 2012	Master in Education: Teaching of Science (Graduated with *hons*) University of León, León (Spain)
2010 – 2012	MSc in Biotechnology (Graduated with *hons*) University of Granada, Granada (Spain)
2004 – 2010	Bachelor degree in Molecular Biology (300 ECTS) University of León, León (Spain)

WORK EXPERIENCE AND INTERNSHIPS

Jan 2016 – *Present*	Junior Research Assistant, PSB VIB-University of Ghent (Belgium)
Apr 2014 – June 2014	Field Research Assistant Laboratory for General Botany, Vrije University of Brussels, (Belgium)
May 2013 – Nov 2013	Internship in Quality Assurance pharma production laboratories Laboratorios SYVA, León (Spain)
July 2012 – Sep 2012	Biochemical and Microbiological analyst in food, water and soil A-Logos, Abrantes (Portugal) Leonardo da Vinci Programme
Jan 2011 – Jan 2012	"Identification and characterization of 6-PGDH in *A. thaliana* peroxisomes" Master thesis: *Del Río* Lab. CSIC-EEZ, Granada (Spain)
July 2010 – Sep 2010	Plant Development Department-Warwick Crop Centre, U. of Warwick, (UK). *Gutiérrez-Marcos* Lab. Genetic and epigenetics in *A. thaliana* and maize.
Feb 2010 – June 2010	Internship in the Plant Biotechnology Department. UTAD, Vila Real (Portugal) "*In vitro* approach to study *Sesamoides purpurascens* in salinity gradients"

GRANTS AND FUNDING

- American Society Plant Biologists Travel Grant PlantBio2020 (Washington, USA).
- EMBO and FWO Short Term Research Stay Grant (Münster University, Germany), 2020-pending.
- FWO Travel Grant 2019 for course attendance (Upsala, Sweden), June 2019.

CURRICULUM VITAE

PUBLICATIONS

Willems, P., Van Ruyskenvelde, V., Maruta, T., Pottie, R., **Fernández-Fernández, Á. D.**, Pauwels, J., Gevaert, K., Van Breusegem, F. and Van der Kelen, K. (*In preparation*) Genetic and chemical perturbation of the plant spliceosome increases oxidative stress tolerance

Vega-Muñoz, I., Durán-Flores, D., **Fernández-Fernández, Á.D.**, Heyman, J., Ritter, A., Stael, S. (2020) Breaking bad news - dynamic molecular mechanisms of wound response in plants. *Frontiers in Plant Science. In press*

Hander, T.*, **Fernández-Fernández, Á. D. *,** Kumpf, R. P., Willems, P., Schatowitz, H., Rombaut, D., … Stael, S. (2019). Damage on plants activates Ca $^{2+}$ -dependent metacaspases for release of immunomodulatory peptides. *Science.* (*) Authors contributed equally.

Fernández-Fernández, Á. D., Van der Hoorn, R. A. L., Gevaert, K., Van Breusegem, F., & Stael, S. (2019). Caught green-handed: methods for in vivo detection and visualization of protease activity. *Journal of Experimental Botany.*

De Smet, B., Willems, P., **Fernandez-Fernandez, A. D.**, Alseekh, S., Fernie, A. R., Messens, J., & Van Breusegem, F. (2019). In vivo detection of protein cysteine sulfenylation in plastids. *Plant Journal.*

Fernández-Fernández, Á. D., & Corpas, F. J. (2016). *In Silico* Analysis of Arabidopsis thaliana Peroxisomal 6-Phosphogluconate Dehydrogenase. *Scientifica.*

POSTERS AND ORAL PRESENTATIONS

Oral presentation: "Metacaspases: plant proteases as switches to wound response" XXVII Workshop Advances in Molecular Biology by Young Researchers Abroad. CNB-CSIC, Madrid 19th December 2019.

Oral presentation: "Arabidopsis METACASPASE 4 is a highly regulated protease controlling plant responses to wounding". 12th Fall Symposium BCRB. Brussels (Belgium), December 2019.

Oral presentation: "Getting to know plant Metacaspases: cues above; substrates identification and downstream signaling". *New Phytologist* Next Gen Scientists Symposium. Dublin (Ireland), July 2019:

Oral presentation: "Plant Cysteine Metacaspases: Sharpening the knowledge of plant scissors". 39th Proteases and Inhibitors Winter School. Tiers (Italy), March 2019.

Poster presentation: "*In vivo* reporters for plant protease activity". Plant Proteases and Programmed Cell Death Meeting. 3rd poster prize. Ghent (Belgium), September 2018.

Oral presentation: "Reducing unwanted proteolysis in Molecular Farming using a CRISPR-dCas strategy". CRISPR User Meeting. Leuven (Belgium), September 2016.

Poster presentation: "Versatile tools for in vivo detection of protease activity: Fluorophore Quencher probes". Plant Proteases and Programmed Cell Death Meeting. 2nd prize. Oxford (UK), April 2016.

FORMATION, TEACHING AND EXTRA CURRICULAR ACTIVITIES

Attendance to courses: "How to write a winning grant proposal", "Scientific Writing & Oral Presentations", "Basic statistics in R", "MS Data Processing" and "Computer Optimized Microscopy".
Student guidance of master (2), bachelor (6) and professional bachelor (1) educational programs.
Science communication activity: "Ciencia no Verão", BiotechDay, Skype-a-scientist, PlantingScience.
Member of American Society of Plant Biologists, Plantae mentorship-mentee program (2019 - *present*).

REFERENCES

Prof. Dr. Frank Van Breusegem.
VIB-UGent Center for Plant Systems Biology
Technologiepark 71, 9052 Gent (Belgium)
Phone: +32 9 331 39 20
e-mail: frank.vanbreusegem@ugent.vib.be

Prof. Jose Gutierrez-Marcos.
Life Sciences, University of Warwick
Coventry, CV4 7AL (UK)
Phone +44 2476 575077
e-mail: J.F.Gutierrez-Marcos@warwick.ac.uk

the plant journal

SEB Society for Experimental Biology

The Plant Journal (2019) **97,** 765–778 doi: 10.1111/tpj.14146

RESOURCE

In vivo detection of protein cysteine sulfenylation in plastids

Barbara De Smet[1,2,3,4,5], Patrick Willems[1,2,6,7], Alvaro D. Fernandez-Fernandez[1,2], Saleh Alseekh[8,9], Alisdair R. Fernie[8,9], Joris Messens[3,4,5,*] and Frank Van Breusegem[1,2,*]

[1]*Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium,*
[2]*VIB-UGent Center for Plant Systems Biology, VIB, 9052 Ghent, Belgium,*
[3]*VIB-VUB Center for Structural Biology, VIB, 1050 Brussels, Belgium,*
[4]*Brussels Center for Redox Biology, 1050 Brussels, Belgium,*
[5]*Structural Biology Brussels, Vrije Universiteit Brussel, 1050 Brussels, Belgium,*
[6]*Department of Biochemistry, Ghent University, 9000 Ghent, Belgium,*
[7]*Center for Medical Biotechnology, VIB, 9000 Ghent, Belgium,*
[8]*Max Planck Institute of Molecular Plant Physiology, 14476 Potsdam-Golm, Germany, and*
[9]*Centre of Plant Systems Biology and Biotechnology, 4000 Plovdiv, Bulgaria*

Received 10 July 2018; revised 23 October 2018; accepted 26 October 2018; published online 5 November 2018.
*For correspondence (e-mails joris.messens@vib-vub.be and frank.vanbreusegem@psb.vib-ugent.be).

SUMMARY

Protein cysteine thiols are post-translationally modified under oxidative stress conditions. Illuminated chloroplasts are one of the important sources of hydrogen peroxide (H_2O_2) and are highly sensitive to environmental stimuli, yet a comprehensive view of the oxidation-sensitive chloroplast proteome is still missing. By targeting the sulfenic acid YAP1C-trapping technology to the plastids of light-grown Arabidopsis cells, we identified 132 putatively sulfenylated plastid proteins upon H_2O_2 pulse treatment. Almost half of the sulfenylated proteins are enzymes of the amino acid metabolism. Using metabolomics, we observed a reversible decrease in the levels of the amino acids Ala, Asn, Cys, Gln, Glu, His, Ile, Leu, Lys, Phe, Ser, Thr and Val after H_2O_2 treatment, which is in line with an anticipated decrease in the levels of the glycolysis and tricarboxylic acid metabolites. Through the identification of an organelle-tailored proteome, we demonstrated that the subcellular targeting of the YAP1C probe enables us to study *in vivo* cysteine sulfenylation at the organellar level. All in all, the identification of these oxidation events in plastids revealed that several enzymes of the amino acid metabolism rapidly undergo cysteine oxidation upon oxidative stress.

Keywords: sulfenic acid, plastid, amino acid metabolism, PTM, oxidative stress, redox signaling, *Arabidopsis thaliana*.

INTRODUCTION

Environmental changes, such as varying light intensities and temperatures, are rapidly perceived in chloroplasts, as these changes directly affect photosynthetic efficiency (Dietz and Hell, 2015; Kleine and Leister, 2016). Singlet oxygen, superoxide radicals and hydrogen peroxide (H_2O_2) are produced along the photosynthetic electron transport chain when excessive energy input exceeds the reducing potential. H_2O_2 is an especially important signaling molecule during plant developmental processes and stress responses (Waszczak *et al.*, 2018). H_2O_2 reacts with cysteine thiols on proteins, forming sulfenylated cysteines. This modification can affect the conformation, activity and interactivity of these proteins (Couturier *et al.*, 2013; Huang *et al.*, 2018). After this first oxidation step, the sulfenic acid (-SOH) formed can react with protein thiols or low-molecular-weight thiols forming disulfides. These disulfides can be recycled by the thioredoxin (Trx) or glutaredoxin (Grx) pathways. Alternatively, -SOH can be further oxidized to sulfinic ($-SO_2H$) and sulfonic ($-SO_3H$) acids (Roos and Messens, 2011). The Grx- and Trx-dependent reversibility together with the high sensitivity of thiols make H_2O_2-dependent sulfenic acid formation an ideal switch to trigger or shut-off downstream signaling responses (Waszczak *et al.*, 2015).

Journal of Experimental Botany, Vol. 70, No. 7 pp. 2125–2141, 2019
doi:10.1093/jxb/erz076 Advance Access Publication 25 February 2019

REVIEW

Caught green-handed: methods for *in vivo* detection and visualization of protease activity

Álvaro Daniel Fernández-Fernández[1,2], Renier A. L. van der Hoorn[3], Kris Gevaert[4,5], Frank Van Breusegem[1,2,*], and Simon Stael[1,2,4,5]

[1] Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium
[2] VIB Center for Plant Systems Biology, 9052 Ghent, Belgium
[3] The Plant Chemetics Laboratory, Department of Plant Sciences, University of Oxford, Oxford OX1 3RB, UK
[4] Department of Biomolecular Medicine, Ghent University, B9000 Ghent, Belgium
[5] VIB Center for Medical Biotechnology, B9000 Ghent, Belgium

* Correspondence: frbre@psb.vib-ugent.be

Received 11 November 2018; Editorial decision 28 January 2019; Accepted 29 January 2019

Editor: Peter Bozhkov, Swedish University of Agricultural Sciences, Sweden

Abstract

Proteases are enzymes that cleave peptide bonds of other proteins. Their omnipresence and diverse activities make them important players in protein homeostasis and turnover of the total cell proteome as well as in signal transduction in plant stress responses and development. To understand protease function, it is of paramount importance to assess when and where a specific protease is active. Here, we review the existing methods to detect *in vivo* protease activity by means of imaging chemical activity-based probes and genetically encoded sensors. We focus on the diverse fluorescent and luminescent sensors at the researcher's disposal and evaluate the potential of imaging techniques to deliver *in vivo* spatiotemporal detail of protease activity. We predict that in the coming years, revised techniques will help to elucidate plant protease activity and functions and hence expand the current status of the field.

Keywords: Activity-based probes, fluorescence, *in vivo* imaging, luminescence, plants, protease, proteolysis, reporter, sensor.

Introduction

Proteases exert a tight control on cellular functions by breaking the polypeptide chain of their substrate proteins. Substrate cleavage is the result of recognition between the target amino acid sequence of a substrate and the binding pocket of a protease. This action can lead to changes in the localization, biomolecular interactions, turnover, or enzymatic activity of the substrate. The importance of proteases across biological kingdoms and viruses cannot be overstated as they intervene in most developmental processes and responses to environmental cues (Turk, 2006; van der Hoorn, 2008). Proteases are also quite numerous; for example Arabidopsis has a reviewed number of around 600, totalling 2% of the protein-coding genes, similar in number and family conservation to other plant species like rice and poplar (García-Lorenzo *et al.*, 2006; van der Hoorn, 2008; Lallemand *et al.*, 2015).

Abbreviations: ABP, activity-based probe; BRET, bioluminescence resonance energy transfer; CA-GFP, caspase-activatable green fluorescent protein; CFP, cyan fluorescent protein; eGFP, enhanced green fluorescent protein; FRET, Förster resonance energy transfer; GFP, green fluorescent protein; HCV, hepatitis C virus; mRFP, monomeric red fluorescent protein; NLS, nuclear localization signal; RFP, red fluorescent protein; TEVp, tobacco etch virus protease; UAS, upstream activating sequence; UV-C, ultraviolet C; VC3AI, Venus caspase-3 activation indicator; YFP, yellow fluorescent protein.

RESEARCH ARTICLE SUMMARY

PLANT SCIENCE

Damage on plants activates Ca^{2+}-dependent metacaspases for release of immunomodulatory peptides

Tim Hander*, Álvaro D. Fernández-Fernández*, Robert P. Kumpf, Patrick Willems, Hendrik Schatowitz, Debbie Rombaut, An Staes, Jonah Nolf, Robin Pottie, Panfeng Yao, Amanda Gonçalves, Benjamin Pavie, Thomas Boller, Kris Gevaert, Frank Van Breusegem, Sebastian Bartels, Simon Stael†

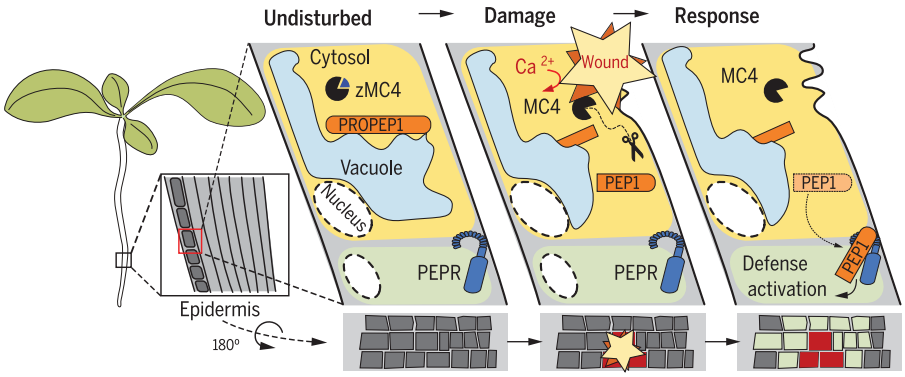
INTRODUCTION: Cellular damage caused by wounding triggers signals to alert the surrounding tissue. As a universal process in all multicellular organisms, these signals activate the immune system to prevent infection and promote tissue regeneration, eventually leading to wound healing, but they have to be secured because aberrant immune stimulation can negatively affect health and growth. Plants, as sessile organisms, are regularly subject to chewing or sucking insects and physical damage inflicted by metazoans or exposure to the environment. In plants, short protein fragments or peptides can have immunomodulatory functions, such as those derived from the plant elicitor peptide (Pep) gene family. Peps are part of precursor proteins and have been proposed to act as wound signals that bind and activate the extracellular Pep receptors (PEPRs) to initiate an immune-like response. How Peps are produced and released upon wounding and how far this response extends from the harm site remain unclear.

RATIONALE: Proteases can generate peptides from precursor proteins, but the genome of the model plant, *Arabidopsis thaliana*, encodes approximately 600 different proteases. To understand the molecular mechanisms that control the immune-activating signals, it is essential to identify the immune response-contributing proteases. Furthermore, wound formation has to be observed with great speed and precision to delineate the extent of the overall process. Pep1 was studied here as a representative member of the gene family.

RESULTS: Pep1 generation was detected within 30 s after the injury, peaking at 5 min and lasting up to 1 hour in two wound model experiments on *Arabidopsis* seedlings: pinching with forceps or mechanical disruption of tissue integrity through grinding. Screening with various protease inhibitors revealed the involvement of metacaspases in Pep1 formation. Metacaspases are cysteine proteases, conserved in plants, fungi, protists, and bacteria that were historically named after caspases because of a certain degree of structural homology. However, they differ in activity and substrate specificity. Metacaspases require low-millimolar amounts of calcium (Ca^{2+}) for in vitro activity and cleave their substrate proteins after the amino acids arginine and lysine. Physical damage activated the abundant METACASPASE4 (MC4) in both wound experiments. MC4 released Pep1 from its protein precursor PRECURSOR OF PEP1 (PROPEP1) by cleaving it behind a conserved arginine. A mutant lacking MC4 was unable to produce Pep1 in leaf tissue, whereas certain redundancy occurred with other metacaspase (or metacaspase-like) activities in root tissue. Inside the plant cell, PROPEP1 is attached to the cytosolic side of the vacuolar membrane. When undisturbed, cytosolic Ca^{2+} concentrations ($[Ca^{2+}]_{cyt}$) are too low to activate metacaspases that need unusually high $[Ca^{2+}]$ to function. When *Arabidopsis* root epidermis cells were damaged by means of multiphoton laser ablation, Pep1 release from the vacuolar membrane occurred only in the cytosol of directly hit cells, as observed with confocal microscopy. Loss of plasma membrane integrity in these cells led to a high and prolonged influx of extracellular $[Ca^{2+}]$ into the cytosol, sufficient to activate metacaspases, leading to the cleavage and release of Pep1 from PROPEP1. Application of exogenous PROPEP1 protein fragments, longer than the native Pep1, reduced root growth—a known negative effect of Pep1 overload—irrespective of MC4 cleavage. Accordingly, PROPEP1 cleavage to overcome retention of Pep1 at the vacuolar membrane seems more important than obtaining the mature Pep1 size.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aar7486



Metacaspases activate defense responses upon wounding. Undisturbed root epidermal cells contain inactive zymogen MC4 (zMC4) in the cytosol and PROPEP1 attached to the vacuolar membrane. Damage induces a prolonged influx of Ca^{2+} in the cytosol, triggering MC4 to cleave PROPEP1 and to release Pep1. Pep1 can then diffuse to neighboring cells and bind to the receptor, PEPR, to activate a defense response (bottom, light green cells). Pep1 release occurs only in cells that lose plasma membrane integrity (bottom, red cells).

CONCLUSION: Plants exploit highly conserved mechanisms—such as Ca^{2+}-partitioning across intact membranes and maturation of immunomodulatory peptides by proteases—to rapidly trigger defense responses against tissue damage. Pep1 is released by activation of MC4 upon prolonged high levels of $[Ca^{2+}]_{cyt}$ that occur only in directly damaged cells, and this response is safeguarded by subcellular retention of PROPEP1 at the vacuolar membrane in the absence of damage. Metacaspases, together with Peps and PEPRs, now emerge as potential targets for breeding and improving crop immunity. ■

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: simon.stael@psb.vib-ugent.be
Cite this article as T. Hander *et al.*, *Science* **363**, eaar7486 (2019). DOI: 10.1126/science.aar7486

frontiers in Plant Science

REVIEW
published: 08 December 2020
doi: 10.3389/fpls.2020.610445

Breaking Bad News: Dynamic Molecular Mechanisms of Wound Response in Plants

*Isaac Vega-Muñoz[1], Dalia Duran-Flores[1], Álvaro Daniel Fernández-Fernández[2,3], Jefri Heyman[2,3], Andrés Ritter[2,3] and Simon Stael[2,3,4,5]**

[1]*Laboratorio de Ecología de Plantas, CINVESTAV-Irapuato, Departamento de Ingeniería Genética, Irapuato, Mexico,* [2]*Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium,* [3]*VIB-UGent Center for Plant Systems Biology, Ghent, Belgium,* [4]*Department of Biomolecular Medicine, Ghent University, Ghent, Belgium,* [5]*VIB-UGent Center for Medical Biotechnology, Ghent, Belgium*

Recognition and repair of damaged tissue are an integral part of life. The failure of cells and tissues to appropriately respond to damage can lead to severe dysfunction and disease. Therefore, it is essential that we understand the molecular pathways of wound recognition and response. In this review, we aim to provide a broad overview of the molecular mechanisms underlying the fate of damaged cells and damage recognition in plants. Damaged cells release the so-called damage associated molecular patterns to warn the surrounding tissue. Local signaling through calcium (Ca^{2+}), reactive oxygen species (ROS), and hormones, such as jasmonic acid, activates defense gene expression and local reinforcement of cell walls to seal off the wound and prevent evaporation and pathogen colonization. Depending on the severity of damage, Ca^{2+}, ROS, and electrical signals can also spread throughout the plant to elicit a systemic defense response. Special emphasis is placed on the spatiotemporal dimension in order to obtain a mechanistic understanding of wound signaling in plants.

Keywords: wound response, damage, damage-associated molecular pattern, systemic signaling, herbivory, jasmonic acid, regeneration

OPEN ACCESS

Edited by:
Massimo E. Maffei, University of Turin, Italy

Reviewed by:
Ivan Galis, Okayama University, Japan
Lotte Caarls, Wageningen University and Research, Netherlands

****Correspondence:***
Simon Stael
simon.stael@psb.vib-ugent.be;
sista@psb.vib-ugent.be

Specialty section:
This article was submitted to Plant Pathogen Interactions, a section of the journal Frontiers in Plant Science

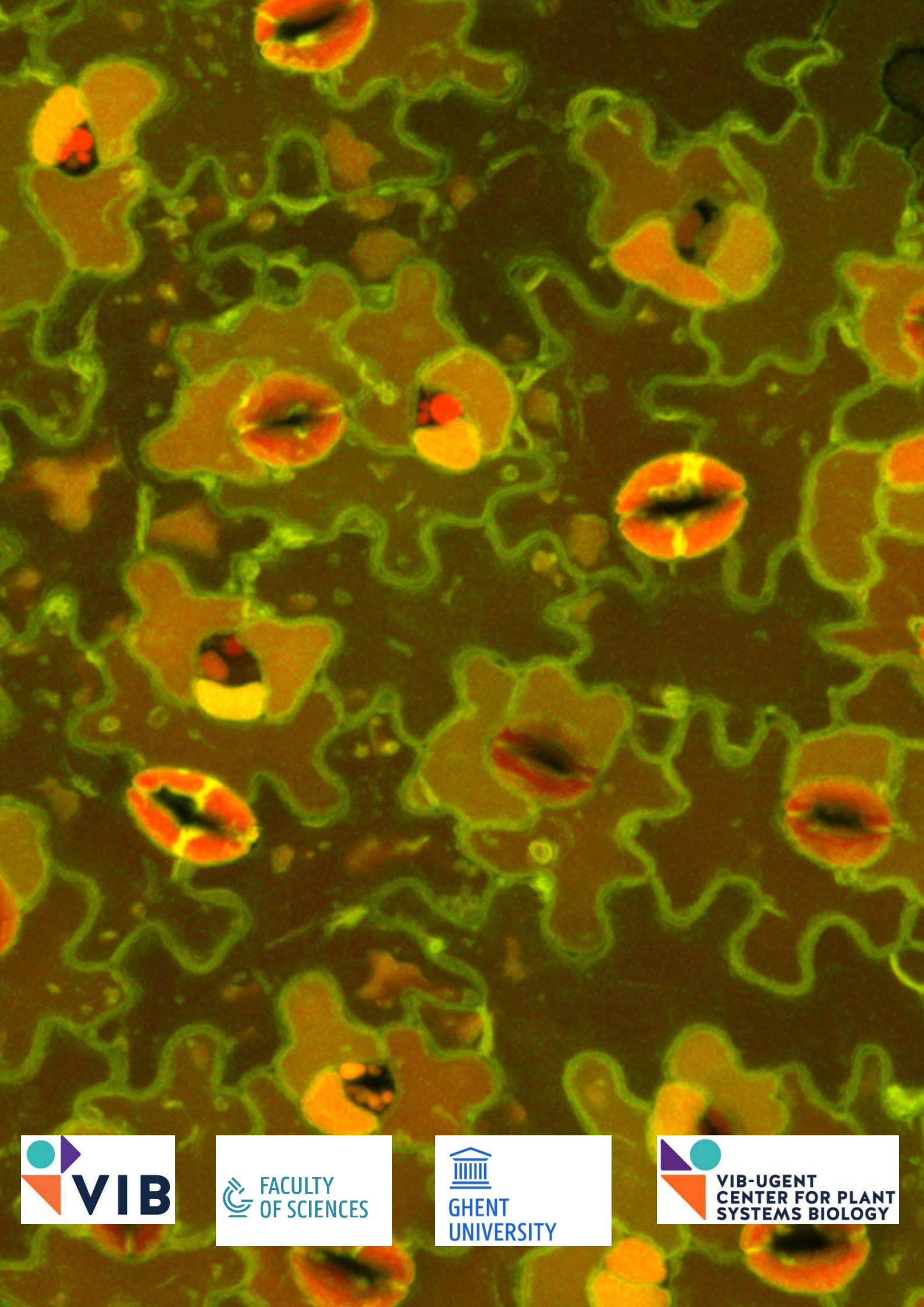
Received: *25 September 2020*
Accepted: *17 November 2020*
Published: *08 December 2020*

Citation:
Vega-Muñoz I, Duran-Flores D, Fernández-Fernández ÁD, Heyman J, Ritter A and Stael S (2020) Breaking Bad News: Dynamic Molecular Mechanisms of Wound Response in Plants. Front. Plant Sci. 11:610445. doi: 10.3389/fpls.2020.610445

INTRODUCTION

Plants are especially susceptible to damage as they are unable to run away when facing danger. Wounds can originate from harsh weather conditions (e.g., strong wind, hail, fire, and frost), physical damage (e.g., trampling), exposure to chemicals (e.g., DNA damage and toxic substances), or biotic attack (e.g., microbes and herbivores). Damage can range in severity from single cell death to complete removal of organs and in duration from single events to repeated injury, for example, from chewing insects. In the lab, mechanical damage can be rather "clean" as in cutting with a sharp razor blade, application of pin pricks, and laser-mediated wounding, or "messy" by bruising tissue with pinches of a forceps or hemostat. We define here "wound" (wounding, wound-induced, etc.) as a general term, while the type of damage that produced the wound can be further specified, such as mechanical- or herbivore-induced damage.

In contrast to metazoans, plants do not rely on a dedicated nerve system or mobile immune cells to sense or respond to wounds. Nevertheless, plants have evolved efficient mechanisms to


VIB
FACULTY OF SCIENCES
GHENT UNIVERSITY
VIB-UGENT CENTER FOR PLANT SYSTEMS BIOLOGY