Multifaceted roles of RING-type ubiquitin E3 ligases: reverse phenomic approaches

Mirko Pavicic Venegas

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Agriculture and Forestry of the University of Helsinki, for public examination.

HELSINKI, FINLAND 2019

DISSERTATIONES SCHOLA DOCTORALIS SCIENTIAE CIRCUMIECTALIS, ALIMENTARIAE, BIOLOGICAE 21/2019
Supervisors  Docent Kristiina Himanen  
Organismal and Evolutionary Biology Research Program  
University of Helsinki, Finland

Follow-up group  Professor Teemu Teeri  
Department of Agricultural Sciences  
University of Helsinki, Finland

Docent Kirk Overmyer  
Organismal and Evolutionary Biology Research Program  
University of Helsinki, Finland

Reviewers  Assistant Professor Lukáš Spíchal  
Centre of the Region Haná for Biotechnological and Agricultural Research  
Palacký University Olomouc, Czech Republic

Docent Esa Tyystjärvi  
Department of Biochemistry  
University of Turku, Finland

Opponent  Professor Thomas Roitsch  
Department of Plant and Environmental Sciences  
University of Copenhagen, Denmark

Custos  Professor Paula Elomaa  
Department of Agricultural Sciences  
University of Helsinki, Finland

ISSN 2342-5423 (print)  
ISSN 2342-5431 (online)  
ISBN 978-951-51-5482-8 (print)  
https://ethesis.helsinki.fi/  
Unigrafia Oy 2019
ABSTRACT

Ubiquitin proteasome system (UPS) is an ATP dependent pathway for targeted protein degradation. The role of UPS is to maintain a healthy protein balance in the cell and to mediate activation and repression of plant developmental processes, hormones and other signalling cascades as well as responses to environmental perturbations. The UPS is composed of several actors, the most important of them being the ubiquitin E3 ligases, which are responsible for providing the specificity for substrate recognition. About 5% of Arabidopsis thaliana genome encodes for ubiquitin E3 ligase genes (~1,400), classified in seven different subgroups, among which the second most abundant group is the RING-type ubiquitin E3 ligase family with nearly 500 members. More than half of the RING-type ubiquitin E3 ligase genes are uncharacterised. Those that are characterized demonstrate their multi-target ability which implies additional roles and cross-reactivity with other pathways. With the emergence of high throughput sequencing, improved Arabidopsis genome assemblies are available and there is a constantly growing amount of transcriptomics data available for the RING-type ubiquitin E3 ligase genes that link them to different developmental stages and perturbations. However, only few of these genes have been associated phenotypically with these processes. Our first aim was to use reverse genetics approach to rescreen Arabidopsis genome in order to update the number of annotated RING-type ubiquitin E3 ligase genes. We further aimed to develop a set of image-based phenotyping methods to systematically assign them in their signalling cascades and developmental pathways, and to functionally characterize the identified molecular networks of the RING-type ubiquitin E3 ligases and their substrates. This study revealed 50 new RING-type ubiquitin E3 ligase genes, while 31 earlier annotated genes were excluded, giving a total new count of 509 RING-type ubiquitin E3 ligase genes. RING-type ubiquitin E3 ligases were then assigned to different developmental, hormonal and/or perturbation related pathways, based on their gene expression profiles. To allow systematic and efficient functional confirmation of these associations in plants, protocols for image-based high throughput phenotypic assays were established. In these assays, the associated knockout lines were studied for rosette shape and growth, cotyledon emergence as a proxy for germination analysis, and Botrytis cinerea symptom progression. These phenotypic screens confirmed 36 flower enriched RING-type ubiquitin E3 ligases, 11 of which were associated to flowering, three rosette and one sepal development. Four of them responded negatively and one positively to ABA treatment at germination. Furthermore, ten RING-type ubiquitin E3 ligases were associated with Botrytis responses, with one characterised at molecular level. Collectively, the results of this study demonstrated the versatility and pleiotropy of RING-type ubiquitin E3 ligases and set the foundation for a systematic screening of phenotypes regulated by UPS components.
ACKNOWLEDGEMENTS

This work was carried out in the Department of Agriculture and Forestry at the Faculty of Agriculture and Forestry at the University of Helsinki. I would like to thank the funding institutions that supported this research: The Academy of Finland (Suomen Akatemia #283138, #256094, #250972), The National Fund for Scientific and Technological Development (FONDECYT, #11060072) from the National Commission for Scientific and Technological Research (CONICYT) Chile, Becas Chile, The Ella and Georg Ehrnrooth Foundation, The European Plant Phenotyping Network and NordPlant network.

I would like to express my deepest gratitude to my supervisor Docent Kristiina Himanen for all the advices, support, mentoring and for not losing faith that this PhD was possible. I want to thank my Professor advisor Paula Elomaa and my dissertation pre-examiners Assistant Professor Lukáš Spíchal and Docent Esa Tyystjärvi for all your critical comments and help during the thesis submission process. I would like to thank also my thesis follow-up group members Professor Teemu Teeri and Docent Kirk Overmyer for their guidance and advices.

I want to thank all technicians and gardeners, specially Eija Takala, Leena Grönholm and Daniel Richterich, without your support this work would not have been possible. I would like to thank everybody who were part of FUPS group and now they are long gone.

Warm thanks to my lunch team Katrin Artola, Sylvain Poque, Feng Wang and Javier Jiménez, university life is always better with friends and deep talks. Also, I want to thank the whole Askelten Palo Folk dance group, specially my good friend Milton Untiveros for all the beers, friendship and be like a brother all these years. I would like to thank Karen Sims-Huopaniemi for all her help in my daily life as PhD student and also many thanks to Katriina Mouhu for all her contributions to my research.

Special thanks to my best friend Tanja Paasela, your support and help during all these years were highly valuable and I think without them I would not be graduating today. You will always have a special place in my memory and in my heart, many thanks again.

Many thanks to my Chilean friends Nico “El longa”, Matías “The fox” and Rodrigo “El panda”. Special thanks to my special friend Soledad, who despite the distance has always been all these years next to me. Lastly I want to thank my dad and my sister for believing in me and supporting me all these years far away from home.

I would like to finish this epistle with the following quote:

... “nunca esta demasiado darse un tiempo para aquellos que son parte importante de tu vida, así como nunca está de más hacerles saber a éstos que lo son”...

-Matías The Fox-
# TABLE OF CONTENTS

List of Publications ......................................................................................................................... 7

Author’s contribution ..................................................................................................................... 8

Abbreviations ................................................................................................................................ 10

1  Introduction .......................................................................................................................... 13
   1.1 Ubiquitination ............................................................................................................... 14
   1.2 Ubiquitin activation enzymes (E1s) ............................................................................. 15
   1.3 Ubiquitin conjugating enzymes (E2s) .......................................................................... 16
   1.4 Ubiquitin ligase enzymes (E3s) .................................................................................... 16
   1.5 The ubiquitin E3 ligase RING domain ......................................................................... 17
   1.6 26S proteasome ............................................................................................................. 18
   1.7 UPS regulated pathways ............................................................................................... 20
       1.7.1 The role of UPS in hormonal signalling ............................................................... 20
       1.7.2 UPS in Plant growth and development ................................................................. 21
       1.7.3 Abiotic stress .......................................................................................................... 23
       1.7.4 Biotic stress ............................................................................................................ 23
   1.8 Reverse genetics approach to assign genes to their functions ................................... 25
   1.9 Image-based automated phenotyping ......................................................................... 26

2  Aims of the study .................................................................................................................... 28

3  Materials and Methods .......................................................................................................... 29

4  Results and discussion ........................................................................................................... 31
   4.1 Reverse genetics approach............................................................................................ 31
   4.2 Rescan of RING domain in all annotated Arabidopsis CDS ........................................ 31
   4.3 Transcriptomic analysis ................................................................................................. 33
   4.4 High throughput in vitro seed germination screen identified two novel ABA responsive RING-type ubiquitin E3 ligases in Arabidopsis thaliana .............................................. 34
   4.5 Phenomic screens of flower related RING-type ubiquitin E3 ligases confirmed their pleiotropic roles ............................................................................................................. 36
RING-type ubiquitin E3 ligases are involved in both susceptibility and resistance to *Botrytis*. .................................................................38

4.6.1 AIP2 interacts with multiple proteins. .................................................................39

4.7 Molecular mechanisms of ubiquitin E3 ligase activity ...........................................40

4.8 Image processing and phenotyping assays: challenges and potential solutions ......42

5 Conclusions and future perspectives ...........................................................................44

6 References....................................................................................................................45
LIST OF PUBLICATIONS

This thesis is based on one peer reviewed article and three manuscripts. The article and manuscripts are cited in the text by their roman number. The articles I and II have been reprinted with permission of the publishers.


IV. **Pavicic, M.,** Garretón, V., Himanen, K. ABI3 INTERACTING PROTEIN 2 confers resistance to *Botrytis cinerea* and regulates PHR1 LIKE 2 protein. Manuscript.
AUTHOR’S CONTRIBUTION

Genomic and Phenomic Screens for Flower Related RING Type Ubiquitin E3 Ligases in Arabidopsis

Mirko Pavicic (MP), Katriina Mouhu (KM), Feng Wang (FW), Marcelina Bilicka (MB), Erik Chovancek (EC) and Kristiina Himanen (KH).

MP conducted the genomic screen of RING domain in coding sequences of Arabidopsis genes and contrasted with previous publication. MP conducted all datamining using gene expression databases. MP performed the experimental design for phenomic screens and the strategies for statistical analysis of these data. KM performed the qPCR analysis and participated in phenotyping and statistical data analysis of leaf at bolting and days to bolting, FW designed and revised genotyping of the mutant collection. MB designed and revised the flower phenotype analysis. EC participated in the phenotyping assays and conducted flowering time experiments. MP designed all artwork for this publication, built, and assembled all figures. MP contributed with the text writing process in all sections related to the tasks before mentioned where MP participated. KH designed the project as a whole, approved the data and wrote the manuscript.

High throughput in vitro seed germination screen identified two novel ABA responsive RING-type ubiquitin E3 ligases in Arabidopsis thaliana

Mirko Pavicic (MP), Feng Wang (FW) and Kristiina Himanen (KH).

KH, MP and FW designed the experiments, MP and FW performed the bioinformatic analysis. FW and KH characterized the transgenic lines, KH designed the phenotyping platform and MP and FW performed the HTP germination assays. MP performed the image processing and statistical analysis of data, MP performed the manual germination assays and designed all artwork presented in this article. KH designed and wrote the manuscript together with FW and MP.
A high throughput method to screen *Botrytis cinerea* symptoms in leaves of the genetic model plant, *Arabidopsis thaliana*

Mirko Pavicic (MP), Kirk Overmyer (KO) and Kristiina Himanen (KH)

MP developed the assay, created the scripts for automated image analysis and data extraction, and did the statistical analysis. KO contributed with materials and training for Botrytis culture and handling, as well as Arabidopsis infection protocols. KH guided the project and wrote the manuscript together MP with the advice of KO.

**ABI3 INTERACTING PROTEIN 2 confers resistance to *Botrytis cinerea* and regulates PHR1 LIKE 2 protein**

Mirko Pavicic (MP), Virginia Garreton (VG) and Kristiina Himanen (KH)

MP performed the gene expression data mining using public databases. MP obtained and curated knockout lines used in this study as well as the generation of over expression lines in Arabidopsis and their verification by RT-qPCR. MP did all plasmid recombinations used in this study and performed transient expression experiments as well as western blots. MP conducted all infection assays and their respective statistical analysis. VG designed the yeast two-hybrid screening performed in this study. KH guided the whole project and wrote the manuscript together with MP.
ABBREVIATIONS

ABA  Abscisic acid
ABI3  ABA INSENSITIVE 3
ABI5  ABA INSENSITIVE 5
AGR  Absolute growth rate
AIP2  ABI3 INTERACTING PROTEIN 2
AIRP2  ABA INSENSITIVE RING PROTEIN 2
AMP  Adenosine monophosphate
ANK  Ankyrin repeats
AP3  APETALA3
APC  Anaphase Promoting Complex
ARA11  Araport11
ASK  Arabidopsis Skp1-related
ATH1  Arabidopsis thaliana 1
ATP  Adenosine triphosphate
Aux/IAA  Repressor of auxin respond factors ARF
BAK1  BRASSINOSTEROID-INSENSITIVE1-ASSOCIATED KINASE1
BB  BIG BROTHER
BIK1  BOTRYTIS-INDUCED KINASE1
C  Cysteine
Cas9  CRISPR-associated protein 9
CBF  C-repeat binding factors
cDNA  Complementary DNA
CDS  Coding sequence
CERK1  CHITIN ELICITOR RECEPTOR KINASE1
ChlF  Chlorophyll fluorescence
CO  CONSTANS
COI1  CORONATINE INSENSITIVE 1
COP1  CONSTITUTIVE PHOTOMORPHOGENIC1
CP  Core particle
CPR1  CONSTITUTIVE EXPRESSOR OF PR GENES1
CRISPR  Clustered regularly interspaced short palindromic repeats
CRY1  CRYPTOCHROME 1
CRY2  CRYPTOCHROME 2
CSU1  COP1 SUPPRESSOR1
CUL  CULLIN
CUL4  CULLIN 4
CUL4-DDB  CULLIN4-Damaged DNA-Binding protein
D  Aspartic acid
DELLA  Abbreviation for the DELLA motif using amino acid symbol of aspartate-glutamate-leucine-leucine-alanine
DNA  Deoxyribonucleic acid
DRE  Drought response elements
DREB2A  DEHYDRATION-RESPONSIVE ELEMENT BINDING PROTEIN2A
DRIP1  DREB2A-INTERACTING PROTEIN1
DRIP2  DREB2A-INTERACTING PROTEIN2
DWA1  DWD HYPERSENSITIVE TO ABA 1
DWA2   DWD HYPERSENSITIVE TO ABA 2
EBF1 /2 EIN3-BINDING F BOX PROTEIN 1 / 2
EIN3/EIL1 ETHYLENE INSENSITIVE 3/ EIN3-LIKE 1
ETI Efector triggered immunity
flg22 Flagellin peptide 22
FLS2 FLAGELLIN SENSITIVE2
G Glycine
GA Gibberellic acid
GEO Gene expression omnibus
GID1 GA INSENSITIVE DWARF1
GRAS Domain name derived from three transcription factors GAI, RGA, and SCR
H Histidine
HECT Homology to E6-AP C-Terminus
HOS1 HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENE1
HS Hyperspectral
HTP High throughput
HY5 LONG HYPOCHOTYL5
IBR In between RING finger domain
ICE1 INDUCER OF CBF EXPRESSION 1
IR Infrared
JA Jasmonic Acid
JAZ JASMONATE ZIM DOMAIN PROTEIN
LFY LEAFY
LisH LIS1 homology domain
LPS Lipopolysaccharide
LRR Leucine-rich repeats
LYK5 LYSM RECEPTOR KINASE5
MAPK Mitogen-activated protein kinase
NAC1 NAM/CUC1
NIR Near infrared
NLR Nucleotide-binding domain and leucine rich repeat
PAM Pulse amplitude modulated
PAMP Pathogen associated molecular patterns
PCR Polymerase chain reaction
PHD Plant homeodomain
PHL2 PHR1-LIKE 2
PP2Cs PROTEIN PHOSPHATASE 2C
PRA Projected rosette area
PRR Pattern recognition receptors
PSII Photosystem II
PTI PAMP triggered immunity
PUB Plant U-box
PYR/PYL/R PYRABACTIN RESISTANCE 1/PYR1 LIKE/REGULATORY COMPONENT
CAR OF ABA RECEPTOR 1
QY max Maximum quantum yield
R proteins Resistance proteins
RGB Red-Green-Blue
RGR Relative growth rate
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RHA1A</td>
<td>RING-H2 FINGER A1A</td>
</tr>
<tr>
<td>RIFP1</td>
<td>RCAR3 INTERACTING F-BOX PROTEIN 1</td>
</tr>
<tr>
<td>RING</td>
<td>Really interesting new gene</td>
</tr>
<tr>
<td>RING-v</td>
<td>RING variant</td>
</tr>
<tr>
<td>RLCK</td>
<td>Receptor like cytoplasmic kinase</td>
</tr>
<tr>
<td>RLK</td>
<td>Receptor like kinase</td>
</tr>
<tr>
<td>RNAseq</td>
<td>RNA sequencing</td>
</tr>
<tr>
<td>ROI</td>
<td>Region of interest</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive oxygen species</td>
</tr>
<tr>
<td>RP</td>
<td>Regulatory particle</td>
</tr>
<tr>
<td>RPN</td>
<td>26S PROTEASOME REGULATORY SUBUNIT</td>
</tr>
<tr>
<td>RPS2</td>
<td>RESISTANCE TOPSEUDOMONAS SYRINGAE2</td>
</tr>
<tr>
<td>RPT</td>
<td>REGULATORY PARTICLE TRIPLE-A</td>
</tr>
<tr>
<td>RT-qPCR</td>
<td>Quantitative reverse transcription PCR</td>
</tr>
<tr>
<td>S/T</td>
<td>Serine/Threonine</td>
</tr>
<tr>
<td>SCF</td>
<td>Skp, Cullin, F-box</td>
</tr>
<tr>
<td>SDIR1</td>
<td>SALT- AND DROUGHT-INDUCED RING FINGER1</td>
</tr>
<tr>
<td>SDIRIP1</td>
<td>SDIR1- INTERACTING PROTEIN1</td>
</tr>
<tr>
<td>SINAL7</td>
<td>SINA LIKE 7</td>
</tr>
<tr>
<td>SINAT5</td>
<td>SINA of <em>Arabidopsis thaliana</em> 5</td>
</tr>
<tr>
<td>SLY1</td>
<td>SLEEPY1</td>
</tr>
<tr>
<td>SnRK2s</td>
<td>subfamily 2 SNF1-related kinases</td>
</tr>
<tr>
<td>SPA1</td>
<td>SUPPRESSOR OF PHYA- 105 1</td>
</tr>
<tr>
<td>T-DNA</td>
<td>Transfer DNA</td>
</tr>
<tr>
<td>TEAR1</td>
<td>TIE1-ASSOCIATED RING-TYPE E3 LIGASE1</td>
</tr>
<tr>
<td>TIE1</td>
<td>TCP INTERACTOR-CONTAINING EAR MOTIF PROTEIN1</td>
</tr>
<tr>
<td>TIR1</td>
<td>TRANSPORT INHIBITOR RESPONSE 1</td>
</tr>
<tr>
<td>UBA</td>
<td>Ubiquitin activating enzyme</td>
</tr>
<tr>
<td>UBC</td>
<td>Ubiquitin conjugating enzyme</td>
</tr>
<tr>
<td>UFO</td>
<td>UNUSUAL FLORAL ORGANS</td>
</tr>
<tr>
<td>UPS</td>
<td>Ubiquitin proteasome system</td>
</tr>
<tr>
<td>XBAT32</td>
<td>XB3 ORTHOLOG 2 IN ARABIDOPSIS THALIANA 32</td>
</tr>
<tr>
<td>ZNF_ZZ</td>
<td>Zinc finger domain</td>
</tr>
</tbody>
</table>
1 INTRODUCTION

Proteins are major cellular macromolecules that carry out many tasks in cells, as catalytic enzymes, regulators, transporters and structural proteins. The availability and activity of the proteins are regulated at many levels, through biosynthesis, activating and repressing modifications and degradation. One mechanism by which cells maintain protein homeostasis is the ubiquitin signal. The main functions of ubiquitin signal are to control vesicle trafficking, cellular localization, protein interaction, and probably the most important one, protein degradation. Ubiquitin tagged proteins are degraded in a multimeric proteolytic complex called 26S proteasome and together they make up the ubiquitin proteasome system (UPS, Figure 1). The UPS, label and remove misfolded proteins, proteins which exceed their required amount and proteins whose function is not required anymore. UPS is a second checkpoint of protein regulation (after transcription) and plays important functions in the whole plant development, and their interactions with the environment.

Ubiquitin is a ubiquitous small peptide present in all eukaryotes and is used as a protein marker for degradation. The ubiquitin binding to proteins is not random and it requires the activity of three enzymes (Figure 1). The first enzyme of the pathway is the ubiquitin activating enzyme (E1) that uptakes ubiquitin from the medium and covalently attaches to itself. Then, ubiquitin is transferred to a second enzyme called the ubiquitin conjugating enzyme (E2). The third enzyme of the pathway is the ubiquitin ligase enzyme (E3). This enzyme acts as a platform that allows the interaction and final ubiquitination of the substrate target protein by the activated E2 enzyme, by providing specificity for substrate recognition. Typically, eukaryotes have one or two E1s, few E2s and several hundred E3s, showing the importance of a fine regulation of specific substrates. In Arabidopsis (Thale grass / Arabidopsis thaliana), for example, there are 2 E1s, 37 E2s and over thousand E3s (Hatfield et al., 1997; Kraft et al., 2005; Mazzucotelli et al., 2006). This is an overrepresentation of UPS components in plants in comparison to other organisms, showing the importance of this system for plants. In humans for instance there are about 600 of E3s, while in Arabidopsis there are about 1400 that corresponds to 3% and the 5% of their total genome respectively.
1.1 Ubiquitination

Ubiquitin is a conserved small peptide of 76-residues having identical sequence in every plant species analysed (Spremulli, 2000). Ubiquitination of a protein is a result of covalent binding of either one ubiquitin or a chain of ubiquitins. The covalent binding of only one ubiquitin to a target protein is called monoubiquitination, and when several monoubiquitinations occurs in the same protein it is called multiubiquitination (Fleury et al., 2007; Stone et al., 2005; Woloszynska et al., 2019). Polyubiquitination takes place when protein is ubiquitinated with a ubiquitin chain. Ubiquitin contains seven lysines K6, K11, K27, K29, K33, K48 and K63 indicating their position in its sequence. Ubiquitin can be attached to any of the seven lysine residues present in its own sequence or in its N-terminus, resulting in ubiquitin chain. The fate of the ubiquitinated target protein will depend on the chain type. For example, K6 chain has been associated with DNA repair, K11 with development and immunity, K27 with DNA repair and nuclear translocation, K29 with proteasomal and lysosomal and kinase modification, K33 with kinase modification, K48 with proteasomal degradation, and K63 with DNA repair and endocytosis (Zhou and Zeng, 2016). In the first step of ubiquitination, ubiquitin C-terminus is bound by a thioester bond to a cysteine in the catalytic site of an E1 enzyme by ATP hydrolysis. Afterwards, E1 interacts with E2 enzyme and transfers the ubiquitin to a sulfhydryl group in a cysteine residue in the E2. This transfer is a transthiolation reaction, and due to the final product, it is also a thioester linkage. In the last step, the ubiquitin is transferred from E2 to a lysine in the target protein. This reaction is facilitated by an E3 enzyme that brings together the E2 and the substrate so that the ubiquitination reaction can take place (Figure 1). Here the C-terminus of ubiquitin is covalently attached to ε-amino group of lysine by an isopeptide linkage. It is believed that the chain type is governed by the E2 enzymes (Ye and Rape, 2009). The most characterized ubiquitin chain is the K48-type and proteins tagged with four or more ubiquitins interlinked with K48 bonds are recognized by the proteasome for degradation. The proteasome hydrolyses the proteins, converting them into peptides. The released amino acids, as well as ubiquitins, are recycled to be reused in the cell (Figure 1).
1.2 Ubiquitin activation enzymes (E1s)

The first enzymes in the ubiquitin cascade are E1s, also known as ubiquitin activating enzymes (UBA). *Arabidopsis* has two UBAs, UBA1 and UBA2 and they are expressed in all tissues and cells of the plant. UBA1 and UBA2 share a high amino acid identity of 81.1%, but this identity declines quickly when compared with other species (Hatfield et al., 1997). UBAs are also very unspecific in that they can interact and transfer the ubiquitin to several different E2s (Hatfield et al., 1997). *Arabidopsis* UBAs are a single-chain polypeptides of about 1100 residues that contain a positionally conserved cysteine, which can bind activated ubiquitin and a nucleotide binding motif, which interacts with either ATP or AMP-Ubiquitin intermediates for ubiquitin activation (Glickman and Ciechanover, 2002; Hatfield et al., 1997).
1.3 Ubiquitin conjugating enzymes (E2s)

The second enzymes in the ubiquitination cascade are E2s or ubiquitin conjugating enzymes (UBC). These enzymes possess a characteristic UBC domain with a catalytic cysteine in its active site. This cysteine is used to accept the activated ubiquitin from the E1s. UBC domains are also used to interact with E3 enzymes and in some cases with the substrate directly (Glickman and Ciechanover, 2002). UBCs have variable expression patterns, ranging from some being ubiquitously expressed during plant development to others expressed in a specific temporal pattern, suggesting that they could regulate specific processes. Different UBCs have been associated with different cellular processes. For instance, E2s in yeast and animals have been associated with functions such as cell cycle, DNA repair and endocytosis (Pickart, 2001). In plants E2s have not been extensively studied beyond their function in the ubiquitin cascade. However, a subgroup of tomato E2s was found to be essential in plant immunity in *Pseudomonas syringae* pv. *tomato* responses (Zhou et al., 2017). *Arabidopsis* genome contains 37 seven different E2 genes, more than half of which have been demonstrated to bind ubiquitin by a thioester linkage, and to extend the ubiquitin chain (Kraft et al., 2005). They are also grouped in sixteen groups depending on their sequence similarity (Kraft et al., 2005). E2s are promiscuous enzymes which interact with a great variety of E3 ligases.

1.4 Ubiquitin ligase enzymes (E3s)

The third group of enzymes in the ubiquitination cascade are the ubiquitin E3 ligases. The E3 ligases are the most numerous ubiquitin signalling components and are responsible for substrate specificity and recognition. In *Arabidopsis*, there are over thousand ubiquitin E3 ligases classified in several groups depending on their structure and mode of action (Mazzucotelli et al., 2006; Sadanandom et al., 2012). According to Mazzucotelli et al. (2006), there are 1415 annotated ubiquitin E3 ligases distributed in seven classes in *Arabidopsis* genome and this number is expected to grow with better annotation tools and new versions of *Arabidopsis* genome. Ubiquitin E3 ligases can be classified as a single unit or a multimeric protein complex. In the single unit E3 ligases, the HECT (Homology to E6-AP C-Terminus) group has an active cysteine that uptakes ubiquitin prior to target ubiquitination (Downes et al., 2003). In this category there are also RING (Really Interesting New Gene, Figure 2) and PUB (Plant U-box)
ubiquitin E3 ligases that do not contain an active cysteine. They can act as monomers, homo or heterodimers, and function as an scaffold, facilitating the encounter of the E2 and the substrate (Azevedo et al., 2001; Stone et al., 2005). Components of the multimeric SCF ubiquitination protein complex are, CUL (CULLIN), ASK (Arabidopsis Skp1-related), and F-box (Cyclin F proteins). Other multimeric complexes are BTB (Bric as brac, Tramtrack and Broad complex), CUL4-DDB (CULLIN4-Damaged DNA-Binding protein) and the APC (Anaphase Promoting Complex, Mazzucotelli et al., 2006; Rojas et al., 2009).

1.5 The ubiquitin E3 ligase RING domain

The RING-type ubiquitin E3 ligases contain an arrangement of cysteine and histidine residues able to bind zinc ions in their tertiary structure, generating finger-like protrusions (Figure 2, Barlow et al., 1994). Initially, it was believed that RING fingers were able to bind DNA like other zinc finger structures, but they failed DNA binding in in vitro assays suggesting a different function (Everett et al., 1993). Today, it is known that RING finger structures allow protein-protein interactions. The three cysteine, one histidine and four cysteines arrange themselves in finger-like structure (Kosarev et al., 2002). This organization is called RING HC-type or C_3HC_4-type. In Arabidopsis, there are several variations of this canonical patterns such as RING-H2, RING-C_2, RING-v, RING-D, RING-S/T and RING-G, where one or more cysteines and histidines are replaced by an alternative amino acid (Stone et al., 2005). Besides the conserved cysteine and histidine, the sequences between them vary greatly from protein to protein, however the 3D fold remain similar (Barlow et al., 1994). This variation suggests that RING proteins may have vastly diverse functions. The RING domain exist alone, or it can be accompanied by other domains, such as plant homeodomain (PHD), LIS1 homology domain (LisH), in between RING finger (IBR), ankyrin repeats (ANK) and zinc finger domain (ZNF_ZZ), to name a few (Stone et al., 2005). Most of described RING proteins act as ubiquitin E3 ligases and their RING domain is used to interact with ubiquitin conjugating E2 enzyme (Sadanandom et al., 2012). Impairment of the cysteines in the RING domain, such as mutations of cysteine to serine (C/S), prevent its interaction with E2 enzyme, thereby also inhibits ubiquitination of the target (Zhang et al., 2005). In some cases, C/S mutations in the RING domain impair the interaction with the E2, but the E3 is still able to interact with the target. This creates a competition effect between the mutated and non-
mutated RING-type ubiquitin E3 ligase, preventing the target ubiquitination (Zhang et al., 2005). RING-type ubiquitin E3 ligases can also be unstable proteins through self-ubiquitination (de Bie and Ciechanover, 2011).

**Figure 2.** RING domain architecture. A. RING domain alignment of several A. thaliana RING proteins. Pairs of ion binding residues are denoted with brackets. B. schematic 2D representation of the RING domain fold and interaction with zinc ions modified from Borden and Freemont (1996). Pairs of ion binding residues are surrounded by dashed boxes. X is used to represent the number of potential residues between ion binding residues (Stone et al., 2005). C. 3D representation of a RING domain from Herpex virus regulatory protein. Red boxes are used to show the pairs of ion binding residues. In red, protein main structure; Green, carbon chains; Yellow, sulfhydryl groups, Black, amino group; Blue, zinc ions. Reprinted from Journal of Molecular Biology, 237, Barlow et al., Structure of the C3HC4 Domain by 1H-nuclear Magnetic Resonance Spectroscopy: A New Structural Class of Zinc-finger, 201-211, Copyright (1994), with permission from Elsevier.

### 1.6 26S proteasome

The 26S proteasome has two major roles: to maintain protein homeostasis within the cell and to remove misfolded proteins that can aggregate and become toxic for the cell. Once proteins are ubiquitinated they are recognized by the proteasome associated
receptors, unfolded and hydrolysed by the proteasome. This multimeric protein complex presents a barrel like structure with a core particle (CP) 20S and a regulatory particle (RP) 19S (Figure 3). The CP is responsible for the proteasome proteolytic activity and consists of four rings of seven subunits each. These rings called α and β gives the proteasome its barrel shape and they are organized in αββα order. The α rings are the gates to the core centre, while the β rings carry out the proteolytic activity. The RP is also a multimeric complex made of a base and a lid. The lid contains several 26S PROTEASOME REGULATORY SUBUNIT (RPN) proteins that recognize ubiquitinated proteins and help remove and recycle ubiquitin chains. The base, on the other hand, contains several REGULATORY PARTICLE TRIPLE-A (RPT) proteins that aligns with the CP core through conformational changes, which permits the degradation of unfolded proteins. The core pore is very narrow and only unfolded proteins can go through it. The ring in the RP base acts an ATPase motor that force protein unfolding mechanically and push ubiquitinated proteins to the CP. Proteins with ubiquitin chains of four or more ubiquitins are effectively degraded by the proteasome. However, it has been observed that monoubiquitinated proteins can also be degraded by the proteasome (Braten et al., 2016). For long time it was believed that only K48 type ubiquitin chain was associated with proteasome degradation. Today it is known that all chain types are able to be recognized by the RP and each of them has its own receptor (Bard et al., 2018). The biological role of the proteasome preference for the different chain types is unknown. Two requirements must be met for ubiquitinated protein to be degraded by the proteasome: to be recognized by the lid of RP and to possess an unstructured end which will be inserted in the RP base motor to start the unfolding and degradation. Even though the unstructured area in the protein increase its degradation efficiency, proteins without it can also be degraded by the proteasome. The proteasome 26S has been shown to play a direct role in plant development, abiotic stress, programmed cell death and pathogen responses (Kurepa and Smalle, 2008).
Figure 3. Proteasome 26S structure and protein degradation. A. Regulatory particle (RP) and core particle (CP). RP Base and lid are shown. Rings of α and β subunits are indicated in the CP. B. Ubiquitinated protein recognition and degradation.

1.7 UPS regulated pathways

1.7.1 The role of UPS in hormonal signalling

UPS acts at intracellular level, but its effect has a broader impact through hormone signalling, developmental transitions, and abiotic and biotic stress responses. In fact, most of the plant hormonal pathways are immediate targets of UPS mediated regulation. For instance, the auxin receptor TRANSPORT INHIBITOR RESPONSE 1 (TIR1) is an F-box ubiquitin E3 ligase that recruits SCF ubiquitination complex upon binding auxin and promotes the degradation of Aux/IAA transcriptional repressors, unleashing auxin responses (Dharmasiri et al., 2005). In an almost identical mechanism, CORONATINE INSENSITIVE 1 (COI1) is an F-box ubiquitin E3 ligase able to bind jasmonic acid (JA) and recruit SFC complex to promote the degradation of JASMONATE ZIM DOMAIN (JAZ) repressors, and activate JA signaling (Thines et al., 2007). Conversely, gibberellic acid (GA) do not bind directly to an E3, but to an independent receptor called GA INSENSITIVE DWARF1 (GID1). Once GID1 binds GA, its N-terminal extension (N-ex) surrounds the GA molecule creating a GID1 active form (Sun, 2011). This form is able to interact with GA repressors DELLA (symbols of
aspartate-glutamate-leucine-leucine-alanine motif) and induce conformational changes that expose their GRAS (GAI, RGA, and SCR) domain which in turn is recognized by SCFSLY1 (SLEEPY1) ubiquitination complex (Sun, 2011). DELLAs are then degraded and GA pathway is activated. On the other hand, ethylene acts by protecting their activators ETHYLENE INSENSITIVE 3/ EIN3-LIKE 1 (EIN3/EIL1) by promoting the degradation of the EIN3-BINDING F BOX PROTEIN 1 (EBF1) and EBF2 (An et al., 2010). Abscisic acid (ABA) is recognized by the PYR/PYL/RCAR (PYRABACTIN RESISTANCE 1/PYR1 LIKE/REGULATORY COMPONENT OF ABA RECEPTOR) receptor family and almost all of them can trigger transcriptions of ABA-responding elements (Bueso et al., 2014; Li et al., 2016; Zhang et al., 2018b). This binding allows the ABA receptors to interact and inhibit type 2C protein phosphatases (PP2Cs). The inhibition of PP2Cs allows the autophosphorylation of subfamily 2 SNF1-related kinases (SnRK2s), which in turn phosphorylate and activate ABA related transcription factors and thereby activate ABA signaling (Kline et al., 2010). UPS can shutdown ABA pathway directly by promoting the degradation of its receptors. Studies have shown that the multimeric complex SCFRIFP1 (RCAR3 INTERACTING F-BOX PROTEIN 1) promotes the degradation of ABA receptor RCAR3, while the single unit RING-type ubiquitin E3 ligase RSL1 can ubiquitinate and promote degradation of ABA receptors PYL4 and PYR1 (Bueso et al., 2014; Li et al., 2016). UPS can also regulate ABA pathway after its activation through RING-type ubiquitin E3 ligases that promote degradation or stabilization of two main ABA-related transcription factors ABA INSENSITIVE 3 (ABI3) and ABA INSENSITIVE 5 (ABI5). ABI3 INTERACTING PROTEIN 2 (AIP2) and KEEP ON GOING (KEG) regulate negatively ABI3 and ABI5 respectively, while SALT- AND DROUGHT-INDUCED RING FINGER1 (SDIR1) does it positively (Stone et al., 2006; Zhang et al., 2005, 2007).

1.7.2 UPS in Plant growth and development

UPS is involved in many transitions during plant life cycle. For example, the transcription factor ABI3 prevents seed germination in Arabidopsis. After stratification, AIP2 ubiquitin E3 ligase gene expression is induced and ABI3 is ubiquitinated and degraded by the proteasome, allowing seeds to germinate (Zhang et al., 2005). Knockout plants aip2-1 can germinate as wild-types under normal culture conditions, suggesting that there are additional germination regulators. After germination, the RING-type ubiquitin E3 ligase CONSTITUTIVE
PHOTOMORPHOGENIC1 (COP1), prevents seedling transition from skotomorphogenesis to photomorphogenesis while underground. The knockout mutant cop1-6 shows constitutive photomorphogenic phenotype with short hypocotyl and open cotyledons also in darkness (Xu et al., 2014). COP1 ubiquitinates and promotes degradation of key players in light responses including LONG HYPOCOTYL5 (HY5). To maintain COP1 homeostasis another RING-type ubiquitin E3 ligase COP1 SUPPRESSOR1 (CSU1), promotes COP1 ubiquitination and proteasomal degradation (Xu et al., 2014).

UPS is also involved in root development, for example the RING-type ubiquitin E3 ligase SINA OF ARABIDOPSIS THALIANA 5 (SINAT5) is able to attenuate auxin pathway by promoting NAM/CUC 1 (NAC1) degradation, thus impairing lateral root development (Xie et al., 2002). Plants overexpressing SINAT5 showed a dramatic reduction of lateral roots while the mutant SINAT5 (C49S) produces shorter main root and an increased number of lateral roots (Xie et al., 2002). Conversely, the RING-type ubiquitin E3 ligase XB3 ORTHOLOG 2 IN ARABIDOPSIS THALIANA 32 (XBAT32) is an activator of lateral root development. Mutant plants xbat32-1 showed reduced number of lateral roots due to impaired cell division in the lateral root initiation zone (Nodzon et al., 2004). Other examples of the role of UPS in development regulation are the RING-type ubiquitin E3 ligase BIG BROTHER (BB) that regulates organ size and TIE1-ASSOCIATED RING-TYPE E3 LIGASE1 (TEAR1) that controls leaf shape by promoting degradation of TCP INTERACTOR-CONTAINING EAR MOTIF PROTEIN1 (TIE1), causing a smooth leaf shape (Disch et al., 2006; Zhang et al., 2017).

UPS has also been shown to play a role in floral organ identity and in the regulation of flowering time in photoperiod perception. The F-box protein UNUSUAL FLORAL ORGANS (UFO) is required together with its co-activator LEAFY (LFY) for APETALA3 (AP3) activation (Risseeuw et al., 2013). AP3 is a MADS-box transcription factor that controls petal and stamen identities. Gain of function UFO plants generate ectopic floral meristems (Risseeuw et al., 2013). Regarding photoperiodic control of flowering time COP1 interacts with SUPPRESSOR OF PHYA- 105 1 (SPA1) and together they promote degradation of the main flowering regulator CONSTANS (CO) in darkness (Xu et al., 2016). During the day the blue light receptors CRYPTOCHROME 1 (CRY1) and CRYPTOCHROME 2 (CRY2) inhibit COP1-SPA1 action (Xu et al., 2016). This
mechanism induces early flowering during long day photoperiod and delayed flowering during short day photoperiod.

1.7.3 Abiotic stress

The most studied abiotic stresses in plants are osmotic, drought and freeze tolerance. UPS has been shown to play a role in abiotic stress responses. For instance, the transcription factor ABI5 is induced by osmotic stress and is controlled by UPS through DWD HYPERSENSITIVE TO ABA 1 (DWA1) and DWD HYPERSENSITIVE TO ABA 2 (DWA2) CUL4 ubiquitin E3 ligase complex (Skubacz et al., 2016). The RING-type ubiquitin E3 ligase XERICO is also upregulated under osmotic stress and promotes ABA synthesis (Ko et al., 2006). Another example of UPS in osmotic stress is SDIR1 that prevents seed germination in high salt by promoting SDIR1-INTERACTING PROTEIN1 (SDIRIP1) degradation (Zhang et al., 2007, 2015). Regarding drought stress, the DEHYDRATION-RESPONSIVE ELEMENT BINDING PROTEIN2A (DREB2A) transcription factor binds to drought response elements (DRE) part in the promoters of several cold and drought responsive genes and is regulated by the C3HC4 RING domain–containing proteins DREB2A-INTERACTING PROTEIN1 (DRIP1) and DRIP2 (Qin et al., 2008). Double knockout plants, the drip1/drip2, mutants show high drought resistance (Qin et al., 2008). Another example of drought is ABA INSENSITIVE RING PROTEIN 2 (AIRP2), which is a RING-type ubiquitin E3 ligase induced by ABA, cold and drought (Cho et al., 2011). AIRP2 overexpression plants are more sensitive to ABA keeping stomata closed, while knockout plants show more water loss through the leaves and are therefore more susceptible to drought (Cho et al., 2011). Furthermore, HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENE1 (HOS1) has an incomplete RING domain and it has been identified as an ubiquitin E3 ligase that interacts and ubiquitinates INDUCER OF CBF EXPRESSION 1 (ICE1) transcription factor, which regulates genes in the C-repeat binding factors (CBF) pathway involved in cold acclimation (Chinnusamy et al., 2003; Dong et al., 2006).

1.7.4 Biotic stress

Plants have evolved to recognize potentially harmful microorganisms by recognizing pathogen associated molecular patterns (PAMPs). Typical examples of PAMPs are bacterial lipopolysaccharide (LPS), the flagellin 22 amino acid conserved epitope flg22
and the fungal cell wall component chitin. PAMPs are recognized by pattern recognition receptors (PRR) in the plasma membrane level, which triggers plant immune responses through Ca^{2+} signalling, MAPK cascades, production of reactive oxygen species (ROS) and changes at transcriptional level. Collectively, these responses are called PAMP triggered immunity or PTI. Pathogens have evolved to avoid being recognized by PRR through different strategies. One strategy is to introduce proteins called effectors, which suppress PTI, inside the host. To deal with pathogen effectors plants have evolved resistance proteins (R proteins) which recognize effectors and trigger a more aggressive defence response called effector triggered immunity (ETI). This response typically includes programmed cell death, hypersensitive response and systemic responses (Sorel et al., 2019).

The UPS plays a dual role in pathogen perception and signalling, by helping to recycle receptors and as components of the signal transduction. Failure of either of these tasks could promote disease or trigger a constitutive defence state even when no pathogens are present (van Wersch et al., 2016). Probably the most studied example is the recognition of the bacterial epitope flg22 by the PRR FLAGELLIN SENSITIVE2 (FLS2). Once FLS2 recognise flg22, it interacts with the leucine-rich repeats (LRR) receptor like kinase (RLK) BRASSINOSTEROID-INSSENSITIVE1-ASSOCIATED KINASE1 (BAK1), which is autophosphorylated (Chinchilla et al., 2007). BAK1 then phosphorylates FLS2 and the receptor like cytoplasmic kinase (RLCK) BOTRYTIS-INDUCED KINASE1 (BIK1), which activates ROS production (Lin et al., 2014). The phosphorylated complex FLS2-BAK1 interacts and activates the plant U-box PUB12 and PUB13, through phosphorylation, which in turn ubiquitinates FLS2 promoting both endocytosis and proteasomal degradation of FLS2 (Lu et al., 2011). PUB12/13 can also interact and ubiquitinate the PRR for chitin recognition. This PRR is a heterotetramer consisting of two monomers of RLKs LYSM RECEPTOR KINASE5 (LYK5) and two of CHITIN ELICITOR RECEPTOR KINASE1 (CERK1, Cao et al., 2014). Upon chitin recognition, LYK5/CERK1 complex is autophosphorylated activating defence responses against fungal pathogens. PUB12/13 recognizes and ubiquitinates the activated form, promoting its degradation through the proteasome (Yamaguchi et al., 2017). In addition, UPS also regulates R proteins, for example the CONSTITUTIVE EXPRESSOR OF PR GENES1 (CPR1) which is an F-Box that
promotes degradation of CC-type nucleotide-binding domain and leucine rich repeat (NLR) RESISTANCE TOPSEUDOMONAS SYRINGAE2 (RPS2, Cheng et al., 2011).

Taken together, the UPS is involved in virtually all plant cellular processes acting as switches in the context of development transition and environmental responses, affecting the whole plant life cycle.

1.8 Reverse genetics approach to assign genes to their functions

To associate genes to their functions, there are a number of different strategies that can be generally described as forward and reverse genetics approaches. Forward genetics is a set of techniques that are used to associate phenotypes to genes or groups of genes. Screening different phenotypes can be achieved through the analysis of the offspring from the cross of two ecotypes of the same species, or by inducing random mutations in plants. Random mutations can be induced by mutagenic chemicals, ionizing radiation, Agrobacterium tumefaciens T-DNA insertions, transposons or CRISPR/Cas9 method. Afterwards, plants are screened for the phenotype of interest and the genes involved in the phenotype are mapped by molecular markers, T-DNA sequencing or whole genome sequencing. Contrary, reverse genetics aims to associate genes to phenotype by disrupting one particular gene and evaluate if the disruption caused any effect in the phenotype of interest. This can be achieved by generating mutant collection with T-DNA mapped, by gene silencing or by targeted gene editing using CRISPR/Cas9 method.

Arabidopsis is a model organism for flowering plants. Arabidopsis belongs to the mustard family (Brassicaceae), which includes many crop plants such as broccoli, cabbage, radish, cauliflower, oilseed rape, etc. Arabidopsis has a very short life cycle of about six weeks from germination to mature seeds, it grows well in limited space and it also generates a high yield of seeds, which makes it easy to propagate. All these features have made Arabidopsis the model organism for many plant researchers worldwide and over years it has become the most studied plant, with a myriad of resources that facilitate its study¹. Its genome is fully sequenced and profile pages with description of each annotated gene can be found across several online databases. It is not only possible to find gene information but also their transcriptional changes along

¹ The Arabidopsis Information Resource (TAIR), https://www.arabidopsis.org/
plant development and across a wide range of perturbations are available. Databases such as Genevestigator\(^2\) and Gene Expression Omnibus (GEO)\(^3\), collect experimental data from microarray experiments and in the recent years also from RNAseq experiments. This information can be used to quickly associate genes to a diverse set of processes to later use mutant plants to ultimately confirm the involvement of the genes in the process. Advances in computing made it relatively easy and fast to mine gene expression data and associate it to genes, and several large and mapped T-DNA insertional mutant line collections are already available for functional gene studies\(^4\). However, phenotyping of plants is a process that remains a bottleneck for gene to function association. This is due to the low throughput and labour intensity of manual phenotyping that slow down the whole process (Fiorani and Schurr, 2013).

### 1.9 Image-based automated phenotyping

Phenomics refers to capturing of multidimensional phenotypic data in a systematic manner (Houle et al., 2010). The fusion of reverse genetics strategies and phenomics is known as reverse phenomics (Lobos et al., 2017). In the recent years many image-based solutions have been developed to relief the phenotyping bottleneck (Fiorani and Schurr, 2013). These solutions range from small homemade prototypes with simple scripts to large infrastructures with automated plant care and delivery to the sensors or with moving sensors to image the plants. These two phenotyping strategies are called plant-to-sensor and sensor-to-plant, respectively. Regardless the size and technology, most of the phenotyping systems are equipped with Red-Green-Blue (RGB), pulse amplitude modulated (PAM) chlorophyll fluorescence (ChlF), near infrared (NIR), Thermal (IR) or hyperspectral (HS) cameras. Other imaging techniques such as X-Ray, laser and magnetic resonance are used for 3D imaging. RGB imaging is used to analyse plant growth, growth rate and plant architecture. IR, ChlF and HS imaging are commonly used to assess the physiological status of the plant. IR imaging is used to assess changes in plant temperature under several stresses, such as drought and osmotic (Munns et al., 2010; Vello et al., 2015). ChlF imaging is used as a proxy to assess the photosystem II (PSII) status, which can be used in turn to evaluate photosynthetic activity or different stresses such as biotic and

\(^2\) GENEVESTIGATOR, [https://genevestigator.com/gv/](https://genevestigator.com/gv/)
\(^4\) T-DNA express, [http://signal.salk.edu/cgi-bin/tdnaexpress](http://signal.salk.edu/cgi-bin/tdnaexpress)
abiotic. On the other hand, HS imaging is used to assess different plant components such as water and pigment content and can be also used to monitor different stress types (e.g. pathogen infection). These sensors can be coupled with automated plant management increasing the throughput, decreasing human work and decreasing variation due to uneven administration of nutrients (e.g. watering systems and plant delivery to the sensors by a conveyor belt). Due to the non-destructive nature of image-based phenotyping, plants can be measured repeatedly over time, which is also called longitudinal analysis. This type of analysis provides details of plant growth patterns and architecture dynamics, which cannot be obtained by manual analysis at few time points (Das Choudhury et al., 2018, 2019). Image-based phenotyping analysis allows also a long-lasting record of plants that can be reanalysed with different tools to extract different features than the originally intended. The analysis of the produced images can be scripted to automate plant-feature extraction reducing greatly the human input in the whole process (Das Choudhury et al., 2018). Further, probably the best characteristic of image-based phenotyping is an immense amount of temporal data that can be collected (high-throughput), which together with controlled environmental data, metabolomics, genomic and transcriptomic data can be used to build complex gene networks and ultimately to elucidate the function of a gene or cluster of genes (Großkinsky et al., 2018).

In this study reverse phenomics approach was undertaken, utilizing an image-based phenotyping system to associate RING-type ubiquitin E3 ligases to their developmental and signalling pathways.
2 AIMS OF THE STUDY

Ubiquitin E3 ligases are the main regulators of UPS since they confer specificity for the targets to be regulated. Thus, ubiquitin E3 ligases are the key players to understand biological processes where UPS is involved. The subgroup RING-type ubiquitin E3 ligases are the perfect study candidate due to their broad action in plant development and their mechanism of action as single unit or dimers. The aim of this study is to understand better the role of RING-type ubiquitin E3 ligases in plant development and their interaction with the environment. Because most of RING-type ubiquitin E3 ligase genes remain uncharacterised we aimed to perform reverse phenomics using genomic, transcriptomic and phenomic screenings to associate them to different plant processes. As study cases, we selected germination and ABA (Manuscript II), rosette and flower development (Article I) and Responses to Botrytis infections (Manuscripts III and IV). In each study case, a sensible methodology was developed to associate phenotypes to mutant lines of RING-type ubiquitin E3 ligases. Finally, one RING-type ubiquitin E3 ligase was chosen as an example and its molecular network was characterized through interactome studies. Taken together, the aims of this study are summarised as follows:

I) To systematically assign RING-type ubiquitin E3 ligase gene family members to their respective signalling pathways and developmental processes.

II) To develop a set of image-based phenotyping methods by combining reverse genetics and phenomics approaches to allow efficient reverse phenomics.

III) Functionally characterize a molecular network of an E3 ligase and its substrate.
3 MATERIALS AND METHODS

The materials and methods used in this research are summarized in the Table 1. A detailed description of the methods can be found in the enclosed articles.

Table 1. List of materials and methods.

<table>
<thead>
<tr>
<th>Materials</th>
<th>Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabidopsis plants</td>
<td>I, II, IV</td>
</tr>
<tr>
<td>Tobacco plants</td>
<td>I, II, IV</td>
</tr>
<tr>
<td>Botrytis cinerea</td>
<td>I, II, IV</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Methods</th>
<th>Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data mining</td>
<td>I, II, IV</td>
</tr>
<tr>
<td>InterproScan protein analysis</td>
<td>I</td>
</tr>
<tr>
<td>Genevestigator transcriptomic screen</td>
<td>I, II, IV</td>
</tr>
<tr>
<td>T-DNA express knockout search</td>
<td>I, II, IV</td>
</tr>
<tr>
<td>Primer design</td>
<td>I, II, IV</td>
</tr>
<tr>
<td>End point PCR</td>
<td>I, II, IV</td>
</tr>
<tr>
<td>RT-qPCR</td>
<td>I, II, IV</td>
</tr>
<tr>
<td>High throughput image-based plant phenotyping</td>
<td>I, II, III, IV</td>
</tr>
<tr>
<td>RGB images plant feature extraction and processing</td>
<td>I, III, IV</td>
</tr>
<tr>
<td>Chlorophyll fluorescence imaging, feature extraction and processing</td>
<td>II, III, IV</td>
</tr>
<tr>
<td>Non-linear mixed models for longitudinal data analysis</td>
<td>I, III, IV</td>
</tr>
<tr>
<td>Plant in vitro culture</td>
<td>I, II, IV</td>
</tr>
<tr>
<td>Gateway TM gene cloning</td>
<td>IV</td>
</tr>
<tr>
<td>Bacterial transformation</td>
<td>IV</td>
</tr>
<tr>
<td>Agrobacterium transient tobacco plants transformation (constitutive and inducible)</td>
<td>IV</td>
</tr>
<tr>
<td>Agrobacterium floral dip Arabidopsis transformation</td>
<td>IV</td>
</tr>
<tr>
<td>Plant protein isolation</td>
<td>IV</td>
</tr>
<tr>
<td>Western blotting</td>
<td>IV</td>
</tr>
</tbody>
</table>

In this study a novel phenotyping methodology was developed to effectively assess Botrytis cinerea symptoms in infected detached leaves of Arabidopsis. This approach utilizes ChlF imaging to detect the maximum quantum yield (QY max) of photo system II values. Based on the QY max values, the region of interest (ROI) affected by the disease can be dissected. A second approach to detect the ROI was developed using
RGB imaging, where a machine learning algorithm was trained to detect healthy, chlorotic and necrotic pixels and quantify them in infected leaves. Furthermore, sensible statistical modelling strategies were introduced to infer differences between controls and a knockout lines collection. Detailed description of these methods is presented in manuscript III.
4 RESULTS AND DISCUSSION

4.1 Reverse genetics approach

To systematically assign RING-type ubiquitin E3 ligase gene family members to their respective signalling pathways and developmental processes (Aim I), we took a gene-to-phenotype approach using image-based high-throughput phenotyping or, in other words, reverse phenomics approach. This strategy consisted of (i) identifying as many genes as possible containing a RING domain, (ii) analysis of their expression data in plant development and perturbations, (iii) obtaining knockout mutants, (iv) characterising their phenotypes at several levels (e.g. morphological and physiological) and (v) identifying their molecular networks (Figure 4).

![Diagram of reverse genetics approach](image)

4.2 Rescan of RING domain in all annotated Arabidopsis CDS

Gene annotation is a process by which gene location and sequence are identified in the genome. There are several steps in gene annotation that can be summarized in two big steps: structural annotation and functional annotation. Structural annotation involves
the identification of gene structure, coding regions (CDS) and regulatory elements (e.g. promoter, untranslated regions, etc). In turn, functional annotation is assignation of biological meaning to genes. This annotation includes biochemical and biological function, interaction networks and expression. Functional annotation may rely on in silico data or experimental data.

The first effort to identify and structurally annotate all genes containing RING domain in Arabidopsis genome was performed by Kosarev et al. (2002). In that study, 446 RING domain containing genes were identified and classified into RING-HC, RING-H2 and RING-variant for those that differed from the canonical RING domains. Few years later, Stone et al. (Stone et al., 2005) rescanned Arabidopsis genome and updated the count of genes containing RING domain to 470, including a more thorough analysis of non-canonical RING domains. Thus, RING-C2, RING-D, RING-S/T and RING-G domain variants were introduced. With the rise of next generation sequencing newer versions of the Arabidopsis genome have been released with better assembly and gene annotation but no update of genes containing RING domain was done in over 10 years. Therefore, the first step in this study was to verify if the already annotated RING domains in Arabidopsis genes remained unaltered in the newest genome release, and screen for potential new candidate genes. Using the latest version of Arabidopsis genome ARA11, all the annotated protein sequences were scanned using InterproScan software to identify all possible RING domain containing proteins. In this way, this study identified 50 new genes containing RING domain not previously included. The newer sequence of a few previously annotated genes did not contain the RING domain anymore and were not taken into the count, highlighting the importance of a constant annotation update. Thus, the count was updated from 470 to 509 (Article I, Figure 2). A recent study, performed a similar screening claiming a total count 506 RING genes with 55 new ones (Jiménez-López et al., 2018). These studies share an overlap of 479 gene accessions, meanwhile 30 were unique for Article I and 27 unique for Jiménez-López et al. (2018). These differences in gene annotation were likely due to the use of different domain scanning tools and selection criteria. In article I, InterProScan software was used, meanwhile Jiménez-López et al. (2018) used TAIR pattern matching tool. Second difference between both studies was the selection criteria, which in Jiménez-López et al. (2018) was less conservative than Article I, allowing for more residues between metal ligand cysteine pairs 1 – 2 and 6 – 7 (Figure
2). Collectively, the evolution of the RING domain annotation indicates the importance of a constant verification of the working material.

### 4.3 Transcriptomic analysis

The next step after genes were structurally verified was functional annotation. The fastest way to suggest a biological meaning to genes is to use gene expression data from microarray and RNAseq experiments. A multinational effort called AtGenExpress was designed to unravel the transcriptome of *Arabidopsis*\(^5\). In this initiative transcriptomic data was created for *Arabidopsis* development, hormone and chemical responses, abiotic and biotic stress (Goda et al., 2008; Kilian et al., 2007; Schmid et al., 2005). All these datasets are currently maintained at GEO and Genevestigator databases (Barrett et al., 2012; Edgar et al., 2002; Hruz et al., 2008). Most of the gene expression studies, however, were based on Affymetrix ATH1 gene chip that only contained probes for 392 out of the 509 RING genes, leaving an important amount of them out of the analysis. Newer datasets using RNAseq technology could shed light to the function of the remaining 117 RING genes, but currently they have been used in low amount of experimental conditions in comparison to microarray datasets. Thus, the 392 available RING genes were submitted for analysis of differential expression in three chosen study cases: plant development, hormonal and biotic stress responses (Figure 4). For plant development the focus was on germination, rosette and flower development and flowering time. For hormone responses the focus was on the effect of ABA in germination, and for biotic stress responses *Botrytis* (gray mold / *Botrytis cinerea*) infections were analysed. Consequently, 16 RING-type ubiquitin E3 ligase were chosen with differential expression for ABA responses (Manuscript II, Table 1). The expression patterns on ABA treatment were highly variable depending on the gene, treatment, organ and developmental stage analysed (Manuscript II, Figure 1A). For *Botrytis* responses, 16 upregulated RING genes were found (Manuscript IV). Since the transcriptomic data for developmental series was available as absolute expression, an artificial relative expression ranking was created. Accordingly, the gene expression ratio of flower to rosette was calculated to create a ranking score of flower expression enrichment (Article I, Figure 3). In this way, 122

---

\(^5\) [https://www.arabidopsis.org/portals/expression/microarray/ATGenExpress.jsp](https://www.arabidopsis.org/portals/expression/microarray/ATGenExpress.jsp)
differentially expressed RING genes were found to be flower organ or flowering time specific.

4.4 High throughput in vitro seed germination screen identified two novel ABA responsive RING-type ubiquitin E3 ligases in *Arabidopsis thaliana*

A complete functional annotation requires both *in silico* and experimental validation. In this study, knockout lines were used, in which RING-type ubiquitin E3 ligase genes were disrupted by a T-DNA insertion (Article I, Manuscript II and IV). T-DNA interrupts the functional messenger RNA production. Thereby, phenotypical changes were expected in those processes where these genes were differentially expressed. Thus, 16 RING-type ubiquitin ligase genes showed differential expression in ABA treatments for which 19 knockout lines were curated (Manuscript II, Table 1). Among many responses, ABA mediates stress signalling towards the germinating seedling. Consequently, the effect of ABA in germination was studied. Since germinating *Arabidopsis* seeds are too small to be efficiently scored without microscopical equipment, the emergence of cotyledons was used as a proxy for seed germination (stage 0.7, Boyes et al., 2001).

Several seed analysis methods have been developed using RGB and HS imaging. In these methods seeds are segmented from the background (usually blue or black) and then their morphological features such as viability, shape, size and germination, are analysed (Dumont et al., 2015; Jahnke et al., 2016; Joosen et al., 2010; Zhang et al., 2018a). These approaches work well for relatively large seeds, but they are not optimal for the analysis of small seeds like those of *Arabidopsis*. Joosen et al., (2010) established a method to detect germinated *Arabidopsis* seeds by subtracting the area of the seed coat to the area of the seed coat plus the germinated root. Thus, if the area of the root plus the area of the seed coat was bigger than the area of the seed coat alone the seed was considered to be germinated. Conversely, if the area of the root tip plus its seed coat was very close to the seed coat area alone, the seed was considered to be not germinated. All these methods together offer an excellent ensemble for the analysis of seed germination and seed quality. However, none of them is suitable for the analysis of growth regulator such as plant hormones. For *Arabidopsis*, one compatible option of growth regulator analysis using image-based phenotyping is *in vitro* plant
This type of analysis has been done for decades, but only recently automated high throughput approaches started to appear in the scientific literature (De Diego et al., 2017; Ugena et al., 2018). Ugena et al., (2018), proposed an in vitro set up using multi-well plates for biostimulants analysis in Arabidopsis. Here one seed per well is sowed in 48-well plates using media with and without biostmulants. The growth rate and colour parameters are then recorded and analysed using RGB imaging. Ugena et al., (2018) also established a method for seed germination using a colorimetric assay and spectrophotometric quantification with high throughput potential.

In this study, we proposed a method that merges both germination and hormone treatment, using cotyledon emergence as an indicator of germination, and in vitro culture for the hormone treatment (Aim II, Manuscript II). Due to the small size of Arabidopsis cotyledons it was difficult to differentiate them from the white media background. Imaging artefacts such as specular reflection and pixel size were the biggest drawbacks in this set up. Therefore, ChlF imaging was used instead to detect the emerged cotyledons. As long as the cotyledons contained chlorophyll, they were detectable even if their area was only few pixels. In this set up, each image contained nine plates with 30 seeds each. A script was created which enabled the selection of only cotyledons by thresholding pixels above certain value and counting the number of seedlings with emerged cotyledons located within a square drawn around each plate. This approach allowed processing hundreds of images in a few minutes. Most importantly, this protocol can be easily modified to assess RGB images, only if the pixel size is smaller e.g. putting plates closer to the camera. This process can also be modified to analyse RGB images of older seedlings.

Germinated seedlings of several knockout lines of genes induced by ABA were counted using their green cotyledons. Altogether, six mutant lines were identified with opposite responses to ABA treatment in germination (Manuscript II, Figure 3A and B). Among them, the mutant ring-h2 finger a1a-1 (rha1a-1) was the most insensitive to ABA treatment with higher percentage of cotyledon emergence, while wav3 homolog 1-3 (wavh1-3) was the most sensitive. These results suggest that RHA1A works as a repressor of germination, likely by degrading either a repressor of ABA signalling or an activator of germination (Manuscript II, Figure 5). Conversely, WAVH1 most likely acts as an activator of germination, possibly by promoting the degradation of ABA pathway components. Thus, the two E3s expanded the current models of ABA
signalling pathways. These results are supported by the expression patterns of the two 
genes, with \textit{RHA1A} highly expressed in dry seeds and low expressed in imbibed seeds, 
while \textit{WAVH1} is highly expressed in both conditions (ePlant\textsuperscript{6}, data not shown). \textit{RHA1A} 
has been originally detected in very young seedlings, and \textit{WAVH1} in leaf primordia of 
young seedlings (Jensen et al., 1998; Sakai et al., 2012).

\textbf{4.5 Phenomic screens of flower related RING-type ubiquitin E3 
ligases confirmed their pleiotropic roles}

Flower expression enrichment was studied in article I (Figure 3) and 122 differentially 
expressed RING-type ubiquitin E3 ligase genes were found, from which 43 knockout 
lines representing 30 unique gene accessions were subjected for phenotypic 
characterization. It was demonstrated that several of the RING-type ubiquitin E3 
ligases impact rosette growth and shape.

Rosette phenotyping is another important step to identify gene functions. In rosette 
producing plants like \textit{Arabidopsis}, the projected rosette area (PRA) is probably one of 
the most used parameters for phenotyping. PRA and parameters derived from it, 
which include absolute growth rate (AGR) and relative growth rate (RGR), are suitable 
to identify differences in diverse plant accessions, growth stimulators and stresses (De 
Diego et al., 2017; Tessmer et al., 2013; Ugena et al., 2018). Besides, there are available 
parameters for rosette geometry analysis, also called shape descriptors or 
morphological parameters. Shape descriptors are an excellent tool for identifying 
plant features that are not immediately evident such as leaf number, leaf serration, 
petiole length, etc. Despite their usefulness, shape descriptors remain widely 
derunderused and only few studies showing their applicability were found (Awlia et al., 
2016; Vasseur et al., 2018). A list of the published parameters for \textit{Arabidopsis} rosette 
can be found in Dobrescu et al. (2017). In this study, PlantScreen\textsuperscript{TM} parameters were 
used and their detailed description can be found in the article I, Figure 1. PRA and 
shape descriptors have proven to be complementary in dissecting plant phenotype and 
to establish differences between plant genotypes. For instance, the mutant \textit{rha1a-1}, 
showed almost identical PRA to the control lines, but they differed greatly in shape 
descriptors (Article I, Figures 5-7). These results show that use of growth parameters

\textsuperscript{6} ePlant: https://bar.utoronto.ca/eplant/
is not sufficient for relevant phenotype analysis and that shape descriptors should be used for accurate plant phenotyping.

In the phenotyping screening in Article I, the mutant csu1-4 showed smaller PRA, meanwhile the mutant sina like 7-2 (sinal7-2) was larger (Article I, Figure 5). These mutations seem to not affect the leaf size, but the timing of leaf emergence. The oscillating pattern of a fitted ellipse eccentricity around the plant rosette serves as a proxy for leaf emergence and expansion. These patterns were identical in both csu1-4 and sinal7-2 in terms of oscillation and amplitude, suggesting that organ development is not affected, but its timing (Article I, Figure 4 and 7). These results suggest that in this case CSU1 and SINAL7 works as development accelerator and brake, rather than developmental switches. RING-type ubiquitin E3 ligases also affect the leaf shape; the mutant line rha1a-1 generates more leaves than control plants and these leaves have serrated edges. The phenotype caused less amplitude in the oscillatory pattern of eccentricity (Article I, Figure 4 and 7). Later in the development, the mutant csu1-4 plants started their reproductive phase about 2-3 days earlier and with 4-5 fewer leaves than the control plants. CSU1 promotes degradation of COP1 and cop1-6 mutant plants show a smaller number of leaves at bolting and less days to bolting (Article I and Xu et al., 2014). Following this logic cop1-6 and csu1-4 should present opposite phenotypes, but they have similar days to bolting and leaves at bolting. This suggests that CSU1 supresses flowering in COP1 independent manner.

Even though most of the analyses performed in this study were automated, flower phenotyping is still a manual process. Phenotyping of Arabidopsis flowers offers mainly three challenges, the first one is flower size. Arabidopsis flowers are so small that a magnification tool is required before imaging. Secondly, Arabidopsis does not produce all flowers at the same time, which means that for a given day it is possible to encounter flowers at different stages of development. This issue could be solved by developing a classification tool to compare flowers in the same developmental stage. Thirdly, Arabidopsis flowers point in many different directions due to the plant branching patterns, which means that non-destructive approaches become hard to implement (e.g. a rotatory camera). Despite all these challenges, some efforts have been made for the analysis of larger flower petal shape and petal symmetry from an evolutionary perspective (Chacón et al., 2013; Savriama, 2018). Furthermore, Vasseur et al. (2018), developed a method to predict the fruit number of different Arabidopsis
accessions using images of the shoot skeleton. In our study, the mutant *sinal7-2* presented abnormal cavities in the flower buds produced by inward bending of two sepals (from a total of four). The bending pressed the petals inside the bud, restricting their growth generating the appearance of a “wrinkling” phenotype. The wrinkling phenotype persisted after the bud opened (Article I, Figure 9). All three mutants presented here affected at least two different plant traits suggesting that they have pleiotropic effects. Nonetheless, T-DNA insertional mutant lines may contain more than one T-DNA insertion (O’Malley et al., 2015). Therefore, gene complementation or the analysis of other knockout lines in the same gene are needed to ultimately associate these phenotypes to these RING-type ubiquitin ligase genes.

4.6 RING-type ubiquitin E3 ligases are involved in both susceptibility and resistance to *Botrytis*.

Following a reverse genetic strategy, we hypothesised that *Botrytis* induced RING-type ubiquitin E3 ligase genes would play a role in the fungus responses. Using public microarray databases, 16 upregulated RING-type ubiquitin ligase genes by *Botrytis* infection were identified, from which 12 knockout lines representing 10 independent genes were curated (Manuscript IV, Figure 1). Many of these genes are associated with ABA pathway, abiotic stress, programmed cell death, biotroph and necrotrophy pathogens (data not shown). To track *Botrytis* disease progression in these knockout lines an image-based method was developed (Aim II, Manuscript III).

The usage of images to assess plant diseases has been done for more than thirty years (Bock et al., 2010). Most of these approaches have been the use of RGB images to identify diseased areas in plant organs, whole plants and even clusters of plants from field images (Bock et al., 2010). As new imaging technologies become more accessible, the usage of ChlF and HS imaging for disease analysis have become more common (Chaerle et al., 2004; Rolfe and Scholes, 2010; Rousseau et al., 2013; Simko et al., 2017). Most of RGB and ChlF approaches have been implemented for plants with large leaves and fruits such as tobacco, wheat, bean and soybean, likely due to their economic importance and the ease to analyse larger diseased plant areas (Bock et al., 2010; Kuckenberg et al., 2009; Rolfe and Scholes, 2010; Simko et al., 2017). Using model organism such as *Arabidopsis* can boost the search for disease resistance genes. However, only few examples of high throughput (HTP) potential using this plant can
be found in the literature (Berger et al., 2007; Fordyce et al., 2018). In the manuscript III we propose a full experimental design to analyse plant pathogen interactions using *Botrytis/Arabidopsis* pathosystem (Aim II). This proposal consisted of an *in vitro* set up for symptom classification using RGB images and symptomatic area, and severity measurement using ChlF images. This set up uses detached leaves infected with *Botrytis* conidia placed in multiwell plates which allows longitudinal studies of disease progression of hundreds of leaves (Manuscript III, Figure 1 and Manuscript IV, Supplementary Table 2). Using the method described in the manuscript III, five different parameters were used to track *Botrytis* symptoms in detached leaves. Three of them corresponding to RGB imaging were proportion of necrotic, chlorotic and healthy pixels. The other two, corresponding to ChlF were symptomatic area and infection severity. All analysed lines showed differences in relation to control line in at least one or more parameters analysed (Manuscript IV, supplementary table 2). However, we found resistant and susceptible lines, and in some cases both susceptible and resistant depending on the parameter analysed. For instance, the line *ring domain ligase 3-1* (*rglg3-1*) showed more symptomatic area, but less severity than the control line. Besides, *rha1a-1* line showed less necrotic area, less symptomatic area and less severity showing a consistent resistant phenotype. On the other hand, *aip2-1* mutant line showed more chlorotic, necrotic and overall symptomatic area, but with no changes in severity, making it mildly susceptible to *Botrytis*. The variable results suggest that these RING-type ubiquitin E3 ligase genes are likely to affect different pathways, impacting different responses to *Botrytis*. Furthermore, these results indicate the need to assess infection responses from different perspectives (e.g. using different parameters), before classifying a gene as resistance- or susceptibility related.

### 4.6.1 AIP2 interacts with multiple proteins.

Once it is established that a gene is involved in a phenotype, the next step is to characterize the molecular phenotypes and the molecular networks (Aim III). In the case of RING-type ubiquitin E3 ligase genes, the protein they encode must interact with other proteins to achieve their function. Therefore, to ultimately unravel their role, the genes must be functionally characterized by establishing their interacting proteins and their ability to ubiquitinate and promote their degradation. As mentioned in the section 1.7.2, AIP2 has been characterized as a RING-type E3 ubiquitin ligase able to interact, ubiquitinate and promote degradation of the transcription factor
ABI3, and thus promote seed germination (Zhang et al., 2005). However, AIP2 is expressed in organs and at developmental stages where ABI3 is not expressed suggesting that AIP2 may regulate other proteins (Zhang et al., 2005). Using a cDNA expression library of 2-weeks old Arabidopsis seedlings, a yeast 2-hybrid screening was performed to identify potential interacting proteins of AIP2. A total of 24 potential interacting partners were found where four of them presented interaction scores high enough to be considered reliable working candidates (Manuscript IV, Table 2). From those four, we chose PHR1-LIKE 2 (PHL2) as working example due to its highly downregulated expression upon Botrytis infections (Manuscript IV, Figure 2). Using transient co-expression in Nicotiana benthamiana leaves it was established that both AIP2 and PHL2 are unstable proteins degraded by the UPS, and that AIP2 promotes degradation of PHL2 (Manuscript IV, figure 3).

Further characterization showed that overexpression of AIP2 and PHL2 confers resistance to Botrytis in several of the five parameters analysed (Manuscript IV, Figure 8 and Supplementary Table 2). AIP2 regulates negatively ABA pathway, while PHL2 activates Pi starvation pathway (Sun et al., 2016; Zhang et al., 2005). Both pathways have been linked to pathogen responses, but it is not clear how these two crosstalk (Baek et al., 2017; Cao et al., 2011; Lievens et al., 2017). More experiments using ABA, Pi starvation and Botrytis infection will be performed to establish the connection among the two proteins in the three environmental conditions.

Taken together, three different experimental setups were developed in this study, with sensible data analysis for rosette geometric changes along plant development, seedling count tracking as proxy for seed germination and ABA response, and symptom tracking in Botrytis infected leaves. Furthermore, seedling count and Botrytis symptom tracking methods where completely developed from raw images until the final data (Article and Manuscripts II-IV).

4.7 Molecular mechanisms of ubiquitin E3 ligase activity

As described in Section 1.1, ubiquitin signal has several roles, and in this study, we focused on the proteolytic function of the UPS. In this context, the way ubiquitin E3 ligases regulate different pathways depends on the function of the targets whose degradation they promote. For example, if a ubiquitin E3 ligase promotes degradation
of a seed germination repressor, then that E3 is a germination activator. Conversely, if it promotes degradation of an activator, then that E3 is germination repressor (Manuscript II, Figure 4). Another example is in pathogen perception, where UPS regulates the amount of pathogen receptors. An excess of receptors triggers a constitutive immune response, while a lack of them makes the plant extremely susceptible (Sorel et al., 2019). Pathogens have evolved effectors that can promote the degradation of these receptors to promote plant infection (Sorel et al., 2019). In a similar manner, the transition from vegetative to reproductive growth in Arabidopsis is regulated by promoting the degradation of flowering activators by UPS when days are short (Yu et al., 2008). When days are long the UPS cannot degrade the activator and the transition to reproductive phase can start (Yu et al., 2008). In other words, ubiquitin E3 ligases act as “molecular switches” that turn on and off developmental transitions as well as responses to environmental changes; a dynamic mechanism of action depicted in Figure 5. Furthermore, RING-type ubiquitin E3 ligases have pleiotropic effects, affecting several developmental stages by promoting degradation of multiple targets.

Figure 5. RING-type ubiquitin E3 ligase regulates different pathways by promoting the degradation of activators and repressors.
4.8 Image processing and phenotyping assays: challenges and potential solutions

Humans are very well trained to identify individual objects in a complex image. Computers instead, only know about numbers stored as pixels in an image, making image analysis challenging. A pixel is the smallest element in an image represented as a small square. An image is an array of pixels usually organized in a square or rectangular manner. Pixels are stored as a number or a mix of numbers depending on the imaging technique. For instance, in monochromatic imaging each pixel contains only one numeric value, RGB images contain three and multispectral imaging contains as many values as the number of wavelengths used to record the image. By manipulating these numbers, computers can identify and segment objects from their background in an image for further analysis. For instance, pixel thresholding is a technique that filters pixels based on their values, for example to select bright elements in a picture in a dark background or to select green pixels of a plant rosette from a brown or blue background, depending on the experiment set up (e.g. blue sand or blue mat to cover the soil). Other techniques are kernel convolution for edge detection or k-means for pixel classification. More advanced techniques of pixel classification involve supervised machine learning and artificial intelligence approaches, which are based on training a classification algorithm by manually pointing region of interest.

Systems which allow full automation of the plant care and data recording enable capturing and collection of information from several hundreds of plants in a time-efficient manner. Due to the non-destructive nature of image-based phenotyping the plant growth, architecture dynamics and physiology can be tracked over time, providing in-depth analysis of plant traits which would be invisible if plants were analysed manually one or few times in the given time. All these new techniques of HTP phenotyping serve as tools to relieve the phenotyping bottleneck and enable the analysis of large collections of mutant lines or multivariate analysis of combination of several treatments (e.g. drought, high temperature and high light). However, these protocols also open new challenges regarding the manner in which the data should be analysed, and the increased the complexity of the data structure. Therefore, sensible statistical approaches must be carefully chosen depending on the data type analysed. Thus, dependency of repeated measurements, non-linear patterns and non-homogeneous errors should be considered when modelling the data as well as appropriate...
probability distributions for different data types (e.g. proportional, positive values including zero, count, binary, zero or one inflated data, etc).

The commercially available HTP solutions come with associated software and servers to store and process images, and to extract plant features with automatic reporting. These tools not only help with the data capture but also reduce the analysis time. It has been about a decade since the boom of HTP begun and the field is still developing. Nowadays, phenotyping solutions are excellent to execute tasks for which they have been developed, but any slight deviation in the experimental set up can cause inaccuracy or deem the experiment useless. For instance, most of the available solutions have been developed to analyse low complexity plants like the flat Arabidopsis rosette or the thin structure of wheat. The available algorithms are not designed to analyse bushy or broad hanging leaf plants, leading to inaccuracies in the data. Similar case is the study of plant physiology, especially in the case of symptom track of plants infected by pathogens. Here, the software is programmed to analyse the whole plant instead the infected area only. Furthermore, HTP software offer only data visualization, but leaves out any sensible statistical modelling of the data. Hence, the current phenotyping bottleneck is the image post-processing, data extraction and statistical modelling.

In summary, the experimental design for each new plant species and application must be rigorously standardized and scripted to maintain the throughput and to avoid systematic errors.
5 CONCLUSIONS AND FUTURE PERSPECTIVES

The UPS is an essential regulator for plant development and plants ability to cope with biotic and abiotic stresses. The number of genes related to UPS is massive, thus HTP techniques must be used to untangle its complex regulation network. Reverse phenomics using gene knockout lines showed to be a very effective tool in combination with HTP to link genes to their functions. In this study we showed that the UPS components ubiquitin E3 ligases containing RING domain are versatile proteins, since several RING-type ubiquitin ligase genes were effectively associated to functions in plant development, hormone responses and pathogen responses (RHA1A, SINAL7, CSU1, WAVH1 and AIP2). RING-type ubiquitin E3 ligases involved in rosette development were also involved in vegetative to reproductive stage transition. Besides, opposite responses to hormone treatment suggests that RHA1A and WAVH1 could be involved in the degradation of activator and repressors of the same pathway. Furthermore, the RING-type ubiquitin E3 ligases were shown to mediate susceptibility or resistance to pathogens depending on their associated regulatory pathway. The same RING-type ubiquitin E3 ligase protein can regulate many plant processes likely by regulating different target proteins. For instance, besides the already known targets, AIP2 was able to promote degradation of PHL2 through UPS. Many other roles of AIP2 are yet to be discovered since the interactome study showed three more interacting candidates to be characterized.

In this study, we contributed to the analysis of rosette development, seed germination and symptom progression using HTP image-based strategies. In an attempt to alleviate the phenotyping bottleneck, the whole pipeline from data acquisition to data analysis was provided. The ability of the automated imaging systems to track several plant features overtime is without doubt a great advantage, unlocking growth and physiological patterns that otherwise would be omitted. There are hundreds of uncharacterized ubiquitin E3 ligases and HTP approaches greatly alleviate this tremendous quest. The future challenge of the phenotyping community is to provide scientists with standardized and reproducible phenotyping assays and analysis pipelines to tackle the functional annotation of not only RING-type ubiquitin E3 ligase genes, but any genes of interest.
6 REFERENCES


Hatfield, P. M., Gosink, M. M., Carpenter, T. B., and Vierstra, R. D. (1997). The


Physiol. 154, 479–82. doi:10.1104/pp.110.160846.


