

Identification of novel components and links in ubiquitin dependent protein degradation pathways of *Arabidopsis thaliana*

Inaugural-Dissertation
zur
Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln



vorgelegt von

Prabhavathi Talloji

aus Metpally, Indien

Köln, 2011

Berichterstatter:

Prof. Dr. George Coupland

Prof. Dr. Ulf-Ingo Flügge

Prüfungsvorsitzende:

Prof. Dr. Ute Höcker

Tag der mündlichen Prüfung: 06.12.2011

Die vorliegende Arbeit wurde am Max-Planck-Institut für Pflanzenzüchtungsforschung in Köln in der Abteilung Entwicklungsbiologie der Pflanzen (Direktor: Prof. Dr. George Coupland) und in den Max F. Perutz Laboratories in Wien in der Arbeitsgruppe von Prof. Dr. Andreas Bachmair angefertigt.

Max Planck Institute
for Plant Breeding Research



ACKNOWLEDGEMENT

It is my greatest pleasure to convey gratitude to each and every one whose support has made this thesis complete.

Many thanks to Prof. Andreas Bachmair, for your support and guidance to the end of this thesis work. The very good experience in the lab helped me to improve my scientific skills.

Big thanks to Prof. George Coupland, and Prof. Ulf-Ingo Flügge for being the referees of thesis. Special thanks to Prof. Ute Höcker for accepting to be the head of the committee. I am also thankful to Dr. Wim Soppe for being “Beisitzer” of the committee.

I would like to convey thanks to Fritz from Arndt von Haeseler group for helping with the analysis of the Solexa sequencing data. I also want to say thanks to Bruno for his support with the microarray data generation.

I always had great support from Kerstin Luxa, I enjoyed my time in the lab with you which I missed so much after moving to Wien. Thanks for being such a nice person and introducing me to future German national women soccer team captain “Tammi”. I acknowledge all other lab members and colleagues in the group.

Many thanks also to Sabina for introducing me to T-DNA library, where I met PRT7, and Valentina for joining me to the green houses at MPIZ for collecting thousands of samples.

Many thanks to Erna with official paper work, such a nice friendly person.

Dani you are one of the best scientist and humble person I know, I am thankful to you forever for your tremendous critical reading and comments on my thesis, which helped to take final shape of this book. My forever thanks for your wonderful and never ending moral support, for your magical words hang on...for redefining me with “Prabha chan”. I just cannot imagine without your support. Your selfless nature, kindness made me realize what a friend means and what a moral support can do. Thanks for standing by my side without being asked for it. The list just goes on.....

Mitzi you have been great company, always had great fun discussing general topics and science. What a hard working person you are, I am inspired with your determination. Thanks for making me feel comfortable at your place in all my visits to Cologne, and your guidance during last days of thesis preparation. I always counted on your support.

Michile and Misi my best neighbors in Cologne: Misi, it was always interesting to discuss science and life with you, thanks for the nice company and moral support which helped me a lot. Michile, many thanks for taking care of my post and the good times especially your support during big party time.

Johanna many thanks for your great company and the smile you brought on my face in the working time in the lab and for the name you gave me “Prabhala”.

Yami and Omar you have been great friends to me in Vienna, tons of thanks for your never ending support, especially for all nice surprises on birthdays. You are simply amazing. You always made sure that I am fine. Thanks for bringing smile on my face. Thanks for the “Prabhita” name you gave me. Thank you very very much.

Justy, the polish gulash, what a colourful company, thanks for the wonderful time and for entertaining me with your company. You always brought right balance to life. Amazing and nice time as always. Thanks for giving me your version of name “Prabhishka”.

Zahra, a wonderful soapy friend who is always totally booked, but still always found time to help me. I enjoyed your company very much had fruitful funny discussions, you have showed what a support means. Many many thanks for sharing last 10 grams, which brought required balance in right time.

Raji, Sravan, Mohit, Silpi, Madura, Shushan, Vova, and Sofy thank you all for your nice support made my time more joyful. Mamoona thanks a lot for the wonderful company, and nice time.

Roman and Simon, the two amazing and great persons, thanks for your great company in the lab, I had great fun and enjoyed my working time in the lab. Simon, thanks for sharing some good experience and nice ideas at the end of thesis. Thanks also to all others in Markus’ group.

Thank you so much Christine, Walter and Johannas you all made me so happy. For making me feel you are back home baby. Especially Christine for wonderful time, for sharing, listening, caring and awarding me “sunshine of Holohergasse “.

My friends back at home country, Ramaesh Pandre thanks a lot, you have been amazing support, wonderful advises, for the confidence and strength you gave me. Minni, Preshu, Suchi, Anantha, Vijju, Guptha, and Neelu you all have been great company and gave me the spirit.

My life time thanks to my parents, amma and nanna brothers, Suresh, Bhasker, Naveen, Pramodh, Karthik, and sister Kavitha, Srinivas bava, Roshni, Rohith, Vanitha vadina, Prathyusha and Rashika you all for being my confidence, energy and balance. I admire all your prayers, love and affection for being always near to me with all your encouragement and belief in me. With your support everything is possible.

CONTENT

Acknowledgement.....	1
Content	3
Zusammenfassung	6
Summary	8
1 Introduction	10
1.1 Components of the ubiquitin 26S proteasome pathway.....	11
1.1.1 Ubiquitin.....	11
1.1.2 Enzymatic steps of ubiquitination	12
1.1.3 Components of ubiquitination process	14
1.1.4 Types of E3 ligases and their functional importance	16
1.1.5 APC/C complex.....	19
1.1.6 The 26S proteasome: a site for protein breakdown.....	20
1.1.7 De-ubiquitinating enzymes (DUBs).....	20
1.2 The N-end rule pathway	21
1.2.1 Enzymatic and Non-enzymatic modifications of N-end rule pathway	23
1.2.2 E3 ligases of the N-end rule pathway and their specificity.....	24
1.2.3 An overview of functional importance of ubiquitin system components and their implication in cell death processes	24
1.3 Background to the thesis	27
1.4 Aim and strategy of the study	27
2 Results	31
2.1 Results part 1 – Search for <i>suppressor of cell death (sud2)</i> candidates	31
2.1.A1 Genetic screen to identify cell death responsible candidates	31
2.1.A2 The <i>suppressor of cell death (sud2)</i> rough mapping from previous work	31
2.1.A3 Generation of a large mapping population	32
2.1.A4 Phenotyping of the new recombinant mapping population	32
2.1.A5 Marker-based genotyping of the new recombinant mapping population.....	34
2.1.A6 Generation of bigger fine mapping population and molecular marker analysis ..	38
2.1.A7 Library construction of 350 kb genomic sub-region of chromosome III.....	40
2.1.A8 Solexa-based sequencing of fine mapped region of <i>sud2</i>	43
2.1.A9 Sequence alignment.....	43
2.1.A10 Identification of SNP candidates	45
2.1.A11 Validation of identified polymorphisms.....	48
2.1.B1 Differential gene expression of <i>sud2</i> and <i>RV86-5</i>	50

2.1.B2	Experimental validation of microarray-based identified candidates	55
2.2	Results part 2 – N-end rule pathway	58
2.2.A1	Generation of reporter lines expressing test substrates	58
2.2.A2	Seed scale-up and EMS-mutagenesis of pER-L-GUS expressing lines.....	60
2.2.A3	pER-L-GUS EMS-mutant screen by live tissue GUS assay	60
2.2.A4	Allelism test among the identified candidates.....	64
2.2.A5	Generation of mapping populations	66
2.2.A6	Phenotypic analysis of identified <i>prr8</i> mutants	66
2.2.A7	Experimental evidence to show that L-GUS is a proteasome substrate, stabilized in <i>prr8</i> mutants	67
2.2.B1	Isolation of UBR-domain proteins of the N-end rule pathway by T-DNA library screening.....	69
2.2.B2	Confirmation of isolated <i>prr7</i> mutant	71
2.2.B3	Analysis of the <i>prr7</i> phenotype.....	72
2.2.B4	Enzymatic analysis of the <i>prr7</i> mutant	73
2.2.B5	Characterization of <i>Arabidopsis BIG</i> , homolog of a mammalian N-end rule pathway E3 ligase.....	73
2.2.B6	Enzymatic analysis of the <i>big</i> mutant.....	74
2.2.C1	Deamidation components of the <i>Arabidopsis</i> N-end rule pathway	75
2.2.C2	Genotyping and phenotyping of <i>Arabidopsis ntan1-1</i> and <i>ntaq1-3</i>	76
2.2.C3	Analysis of enzymatic function of <i>Arabidopsis</i> Ntan and Ntaq	78
2.2.D	NO-mediated modification in N-end rule pathway in <i>Arabidopsis</i>	78
3	Discussion	80
3.1	Discussion part 1	80
3.1.1	The major consequence of ubiquitination inhibition in <i>Arabidopsis</i> is cell death ..	80
3.1.2	Suppressor of ubK48R expression, <i>sud2</i> , rescues lethal phenotype	81
3.1.3	Mapping - a way to hunt for <i>sud2</i> locus: results suggest <i>sud2</i> position on chromosome III	81
3.1.4	Does larger population help to overcome the limitation of low recombination?....	82
3.1.5	Fragment library construction: an alternative route to reach to the <i>sud2</i> locus	83
3.1.6	Next generation sequencing: an excellent tool for mapping process for the identification of genes of interest	83
3.1.7	Analysis method provides graphic quick view to identify candidates showing polymorphism.....	84
3.1.8	Microarray identified differentially expressed candidates	85
3.1.9	Lessons from expression comparison between un-induced and induced <i>RV86-5</i> ...	86
3.1.10	Experimental validation of identified candidates from microarray analysis.....	92
3.1.11	Biological importance of identified candidates.....	93

3.2 Discussion – part 2	97
3.2.1 Importance of the N-end rule pathway in development	97
3.2.2 Target specificities of N-end rule E3-ligases and their role in development	98
3.2.3 Transgene-based screen led to the identification of the novel E3-ligase PRT8.....	98
3.2.4 Do plant homologs of mammalian UBR-domain proteins exhibit E3-ligase function?.....	101
3.2.5 Do plants process tertiary residues via deamidation?	102
3.2.6 Does the N-end rule pathway play a role in NO signal perception?	103
3.2.7 Conclusion and outlook.....	104
4 Materials and methods	105
4.1 Material	105
4.1.1 Chemicals, kits, antibodies.....	105
4.1.2 Oligonucleotides, markers, enzymes.....	106
4.1.3 Bacterial strains and binary vectors.....	106
4.1.4 Plants	107
4.1.5 Buffers and solutions.....	108
4.1.6 Media.....	110
4.2 Methods	110
4.2.1 Transformations	110
4.2.2 Plant methods	111
4.2.3 Plant growth	113
4.2.4 Plant genetic methods.....	114
4.2.5 Model substrate generation and stability assays	116
4.2.7 Purification methods	117
4.2.8 Standard enzymatic reactions.....	118
4.2.9 Nucleic acids synthesis and quantification.....	120
4.2.10 Databases and Bioinformatics tools	120
4.2.11 Other plant-related methods	121
5 Abbreviations	123
6 Appendix	126
7 References	140
Eidesstattliche Erklärung.....	151
Curriculum Vitae.....	152

ZUSAMMENFASSUNG

Der Ubiquitin-26S-Proteasom-abhängige Proteinabbauweg und sein Teilbereich, der N-end-Rule-Weg, sind wichtige Ubiquitin-abhängige Vorgänge in Eukaryonten. Die meisten Substratproteine werden vorrangig durch das Proteosom ihrem Abbau zugeführt. Die Expression einer Ubiquitin-Variante mit Arg anstelle von Lys an Position 48 (ubK48R) in der *Arabidopsis*-Linie RV86-5 führt zum Zelltod. In dieser Arbeit wurden, um diesem Protein-Abbauweg nachgeschaltete Ereignisse zu verstehen, die ubK48R-exprimierende Linie RV86-5 und eine Suppressor-Linie des Ubiquitin-Varianten induzierten Zelltodes, *sud2* (suppressor of ubiquitin variant induced cell death), als Hilfsmittel genutzt. Feinkartierung mit Hilfe von 1239 rekombinanten Pflanzenlinien grenzte die Position des mutierten *SUD2*-Lokus auf die Region des Chromosoms III zwischen den Genen *At3g44400* und *At3g44900* ein. Durch niedrige Rekombinationsrate und repetitive Sequenzen verursachte Probleme wurden durch Herstellung einer sub-genomischen Bibliothek und anschließende Solexa-Sequenzierung dieser 350 kb großen Region überwunden. Die für einen Nukleotid-spezifischen Vergleich zu einer Referenz-Sequenz maßgeschneiderte Daten-Auswertung ermöglichte die Identifizierung von 15 Kandidaten für die *sud2*-Mutation, wovon fünf durch konventionelle Sequenzierung bestätigt werden konnten. Einer alternativen Strategie folgend, wurden mittels Microarray-Analyse von Transkriptmengen-Unterschieden zwischen RV86-5 und *sud2* zehn weitere Kandidaten-Gene für Zelltod-Suppressoren identifiziert, von denen die meisten eine unbekannte Funktion haben. Von Mutanten in neun der untersuchten Kandidaten-Gene waren acht in der Lage, den letalen RV86-5 Phänotyp zu supprimieren, was auf ihre Wichtigkeit für den Zelltod hinweist.

Ein Hauptinteresse der Ubiquitin-Forschung ist die Identifizierung von E3-Ligasen und ihrer Substrate. Der zweite Teil dieser Arbeit beschäftigte sich mit der Suche nach neuen pflanzlichen E3-Ligasen mit einer Funktion im N-end-Rule Weg. Kurzlebige Proteine mit der N-terminalen aliphatischen hydrophoben Aminosäure Leu werden von keiner der beiden bisher bekannten pflanzlichen N-end-Rule-E3-Ligasen, PRT6 und PRT1, erkannt. Mittels EMS-Mutagenese einer Pflanzenlinie, die L-GUS als Test-Protein exprimiert, gefolgt von einer GUS-Färbung an lebenden Pflanzen zum Nachweis der Stabilisierung des Testproteins, wurden die zwei Komplementationsgruppen PRT8 und PRT9 identifiziert, welche putative E3-Ligasen mit einer Rolle in der Destabilisierung von Proteinen mit amino-terminalem Leu repräsentieren könnten. Die *prt8*-Mutante zeigt eine verzögerte Entwicklung. Mit der

Herstellung einer Kartierungs-Population wurde die Grundlage zur Identifizierung dieses Lokus geschaffen.

Des Weiteren wurden in dieser Arbeit *Arabidopsis*-Mutanten der funktionell noch nicht charakterisierten UBR-Domäne-Proteine PRT7 und BIG analysiert, welche Homologie zu den Säugetier-N-end-Rule Komponenten UBR4 und UBR7 aufweisen. Eine *prt7* Mutante, die mit Hilfe eines T-DNA-Bibliothek-Screens isoliert wurde, zeigte verfrühte Blatt-Seneszenz. Im Gegensatz dazu wies die *big*-Mutante verzögerte Seneszenz und zudem keine enzymatische Affinität zu Test-Substraten mit basischem N-Terminus auf. Im Rahmen dieses Projektes wurden außerdem Mutanten für die zwei *Arabidopsis* Deamidasen NTAN und NTAQ isoliert. Diese sind entfernt mit Säugetier-Deamidasen verwandt. Die Mutanten-Linien wurden mit Reporter-Linien gekreuzt, welche N-GUS bzw. Q-GUS exprimieren, um daraus abzuleiten, ob diese Enzyme – wie in Säugetieren – Substrate für Arg-t-RNA-Protein-Transferasen zur Verfügung stellen. Diese hier generierten Linien bilden die Grundlage zur Erforschung unbekannter Funktionen von Komponenten des N-end-Rule Weges in *Arabidopsis*.

Die Bedeutung des NO-Signalweges in Pflanzen wird bereits lange untersucht, doch die molekularen Mechanismen desselben sind noch immer nicht gut verstanden. In dieser Forschungsarbeit wurde gezeigt, dass im N-end-rule Weg in *Arabidopsis* NO den Proteasom-vermittelten Abbau von Substraten mit N-terminalem Cys bewirkt und dass dieser Vorgang von Sauerstoff abhängig ist. Mit diesen Ergebnissen wurden starke Hinweise gefunden, dass der N-end-Rule Weg eine Rolle bei der NO-Signaltransduktion und -Rezeption spielen könnte. Dies erlaubt neue Einsichten in den pflanzlichen N-end-Rule Weg.

Zusammenfassend wurden in dieser Arbeit neue Methoden zur Überwindung des Problems niedriger Rekombination während des Mappings entwickelt, mögliche Bindeglieder zwischen Zelltod und Ubiquitin-abhängigem Proteinabbau identifiziert und neue putative E3-Ligasen des N-end-Rule Weges mit Hilfe einer neuartigen Methode des EMS-Mutanten-Screens, unterstützt durch GUS-Färbung an lebenden Pflanzen, entdeckt. Zudem hat diese Arbeit eine Verbindung zwischen NO und dem N-end-Rule Weg in *A. thaliana* aufgezeigt. Eine umfassende Sammlung von Mutanten des pflanzlichen N-end-Rule Weges wurde geschaffen, die eine Fülle an Möglichkeiten zur Identifizierung natürlicher Substrate der gefundenen Komponenten eröffnet.

SUMMARY

The canonical ubiquitin 26S proteasome dependent protein degradation pathway and its sub-branch N-end rule pathway are important ubiquitin dependent processes in eukaryotes. The majority of substrates are predominantly targeted for degradation by the proteasome. Expression of a ubiquitin variant with Arg instead of Lys at position 48 (ubK48R) in the *Arabidopsis* *RV86-5* line leads to cell death. In order to understand the downstream effects of this pathway, the ubK48R expressing line *RV86-5* and the suppressor line of ubiquitin variant induced cell death, *sud2*, were used as tools. Fine mapping with 1239 recombinants narrowed down the *sud2* mutant locus to the south arm of chromosome III, between loci At3g44400 and At3g44900. Problems caused by low recombination and repeated sequences were overcome by sub-genomic PCR-based amplification of a 350 kb region and subsequent Solexa sequencing of this region of interest. The data analysis tailored for nucleotide based comparison to reference sequence identified 15 candidates, 5 of which could be verified by conventional sequencing. In an alternative approach, microarray-based transcriptional expression differences between *RV86-5* and *sud2* identified 10 additional candidate suppressor genes, the majority of which are of unknown function. Among mutations in 9 of the tested candidates, 8 were able to prevent the lethal phenotype of *RV86-5*, indicating their involvement in the cell death process.

The main interest of the ubiquitin research field is to identify E3-ligases and their interacting substrates. The second part of this work involved the search for novel E3 ligases that modify a known test protein with an aliphatic hydrophobic amino-terminal residue, Leu, which is targeted by none of the known plant N-end rule E3 ligases, PRT6 and PRT1. EMS mutagenesis on a plant line expressing a test protein with L-GUS followed by live tissue GUS staining, to screen for transgene stabilization, identified the 2 complementation groups PRT8 and PRT9, representing candidates for putative E3-ligases involved in destabilization of test proteins with amino-terminal Leu. The *prt8* mutant shows delayed development. With the creation of a mapping population, the basis for the identification of locus was laid in this work.

Arabidopsis mutants in the functionally unknown UBR domain proteins BIG and PRT7, which share homology with the mammalian N-end rule pathway components UBR4 and UBR7, were analyzed. A mutant in PRT7, isolated by T-DNA library screening, showed premature leaf senescence. In contrast, the *big* mutant showed delayed senescence and in addition no enzymatic affinity to test substrates with a basic N-terminus. Mutants were

isolated in two putative *Arabidopsis* deamidases, NTAN and NTAQ that are distantly related to mammalian deamidases. These were crossed into reporter lines expressing N-GUS and Q-GUS test proteins to deduce whether these enzymes provide substrates to Arg-t-RNA protein transferase as in mammals. These created mutants have laid the basis to analyse unknown functions of N-end rule pathway components in *Arabidopsis*.

The importance of NO in signaling in plants has been long studied, but its molecular mechanism is still not well understood. In this work, it was found that in the *Arabidopsis* N-end rule pathway, NO targets test substrates with N-terminal Cys for degradation in a proteasome dependent manner and that this process is dependent on O₂. With these results, strong evidence was obtained that the N-end rule pathway has a role in NO signaling and sensing. This finding has brought new insights into the plant N-end rule pathway.

Taken together, the research work of this Thesis has developed new methods to overcome the low recombination problem during the mapping process, identified candidates that could potentially link the cell death processes to the ubiquitin dependent degradation pathway and identified putative E3-ligases of the N-end rule pathway by a novel way of EMS mutant screening supported by live tissue GUS assay. This research work found a connection between NO and the N-end rule pathway in *A. thaliana*. A complete set of mutants in all known plant N-end rule pathway components has been created, opening a window of possibility to further find natural substrates of this pathway.

1 INTRODUCTION

All phases of plant life are tightly regulated by the protein levels present at that given stage. In eukaryotes, the ubiquitin 26S proteasome pathway (UPS) and its related N-end rule pathway (NERP) are the main protein degradation pathways and crucial to maintain and establish a characteristic pool by removing unnecessary proteins in a given developmental phase of the eukaryotic life. Cell death, paradoxically an integral part of growth and development, is tightly linked to and regulated by the protein degradation machinery.

In plants, cell death processes occur at different stages of life and can be divided into three major categories. Firstly, Programmed cell death (PCD) takes place during developmental processes (at the organ level), for example during embryogenesis, cell differentiation and organ development. Secondly, PCD is executed during defense against biotic or abiotic stress. This type of PCD is also known as hypersensitive response (HR). A third type of PCD is senescence, which occurs at the whole organ or plant level and is a slower form of cell death. To what degree all these types of cell death share common components is not known. But one common phenomenon that can be noticed among all cell death processes is protein degradation.

Many of the programmed cell death substrates are suspected to be targeted for degradation via ubiquitin dependent pathways (UPS and NERP). In animals it was also identified that nitric oxide (NO) involves in S-nitrosylation, which promotes its recognition by the E3-ligases, resulting either in degradation or translocation (MANNICK 2007). This process is not well studied in plants. Unlike in animals, in plants mechanisms of interlinks between ubiquitin dependent protein degradation and cell death are not well developed. Nevertheless, in plants a direct connection between ubiquitin system and cell death became evident from the research findings of Andreas Bachmair's Lab (2006) (SCHLOGELHOFER *et al.* 2006). So far some positive or negative regulatory components of hormonal signaling pathways and plant pathogen interaction that influence cell death have been identified as targets of ubiquitin dependent protein degradation systems (UPS/NERP). Genetic and biochemical studies using mutants in the ubiquitin proteasome dependent protein degradation pathways and other hormonal signaling pathways directly show the importance of components of the ubiquitin proteasome pathway in regulation of various processes, but information about connections between ubiquitin system and cell death in plants is still missing. Therefore it is very interesting to understand the biochemical and molecular basis behind the interactions between

PCD and ubiquitin dependent protein degradation pathways (UPS/NERP), and identification of key players would help towards this end and this formed the central aim of this thesis work.

1.1 Components of the ubiquitin 26S proteasome pathway

1.1.1 Ubiquitin

Ubiquitin is encoded by a multi gene family resulting in production of ubiquitin as protein fusion; functional monoubiquitin is released by ubiquitin C-terminal hydrolases (CALLIS *et al.* 1995; JENTSCH 1992). In *Arabidopsis thaliana*, a model organism for plants, there are 14 different ubiquitin genes present (*AtUBQ1-14*). Ubiquitin (ub) is a structurally most conserved small protein with 76 amino acids. In plants it differs by two and three residues from the yeast and animal protein respectively (BURKE *et al.* 1988; CALLIS *et al.* 1995).

Ubiquitin has a very compact globular structure with extensive hydrogen bonds making it highly stable and it refolds spontaneously if unfolded. Ubiquitin has a flexible protruding carboxyl terminus with a glycine at the end facilitating covalent interaction with E1s, E2s, and some E3s in a conjugation cascade. It finally ends up forming an isopeptide bond with a lysyl ϵ -amino group on the targeted substrate with rare exceptions where ubiquitin binds to a cysteinylsulphydryl group of the substrate (CADWELL and COSCOY 2005). Ubiquitin contains 7 lysine residues, which are positioned at 6, 11, 27, 29, 33, 48 and 63 (Fig 1). These seven lysyl residues can be used for marking the substrates with ubiquitin-ubiquitin linkages (PENG *et al.* 2003) and the fate of substrates mainly depends on the type of lysyl bond used and on the length of the ubiquitin chain on the substrate. Substrates marked with Lys48 poly-ubiquitin chains are predominantly targeted for degradation by the 26S proteasome. A chain consisting minimally of four ubiquitins on the substrate is required to be recognized by the 26S proteasome (CHAU *et al.* 1989; FINLEY *et al.* 1994; THROWER *et al.* 2000). There are some exceptions noticed by Kirkpatrick and coworkers, that an endogenous substrate, yeast cyclinB, could be ubiquitylated *in vitro* with ubiquitin lysine 11 and 63 linkages which served as proteolytic signals (KIRKPATRICK *et al.* 2006) and Hofmann and Pickart showed that attachment of a lysine 63-linked poly-ubiquitin chain to lysine 48 of the ubiquitin protein of the ubiquitin dihydroxyfolate reductase (UbdHFR) fusion protein leads to degradation of this fusion substrate *in vitro* by purified proteasome (HOFMANN and PICKART 2001), suggesting that other ubiquitin-ubiquitin linkages can serve as proteolytic signals. Ubiquitin chains formed on substrates by means of other than lys48 poly-ubiquitin are involved in other processes than degradation by proteasome. For example modification by mono-ubiquitination

can drive substrates to the lysosome/vacuole for turnover (HICKE 2001) or can also affect the transcription machinery (BACH and OSTENDORFF 2003). There is some evidence that ubiquitin chains formed via lysine 63 are involved in processes such as DNA repair, protein activation (SCHNELL and HICKE 2003), ribosomal regulation and endolysosomal degradation (DUNCAN *et al.* 2006; SPENCE *et al.* 2000). It is also noticed that lysine 6 linked chains take part in regulation of DNA replication and repair (MORRIS and SOLOMON 2004; NISHIKAWA *et al.* 2004).

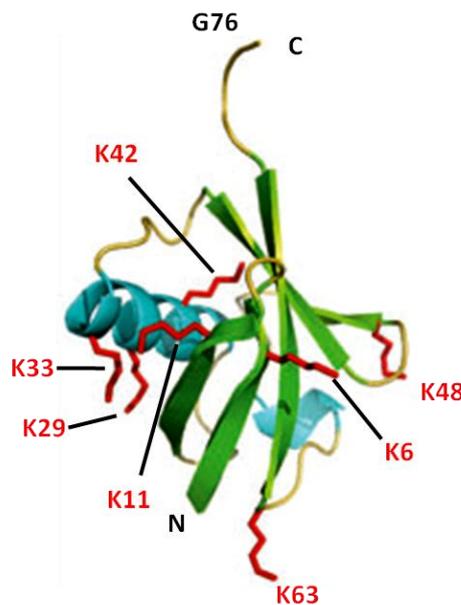


Figure 1 Three dimensional ribbon model of plant Ubiquitin. Figure taken from (HUA and VIERSTRA 2011) modified. Three dimensional ribbon model of plant ub (VIJAY-KUMAR *et al.* 1987) the side chains from the seven lysines in ub that can be used for poly-ub chain formation are shown in red. The β strands are in green, the α helices are in cyan, and the C-terminal Gly76 used to ligate ub to other proteins is indicated. N, N-terminus; C, C-terminus.

Taken together it is clear that ubiquitin with its seven lysine residues marks the substrates and drives them to various destinations and takes part in wide variety of biological processes.

1.1.2 Enzymatic steps of ubiquitination

The ubiquitin 26S proteasome system (UPS) and its sub pathway N-end rule pathway (NERP) are main protein degradation pathways. The functional part of pathways involves two important steps. The first one is to recognize substrates and tag them with single or poly-ubiquitin; the second step is to drive the ubiquitinated substrates to different destinations depending upon type and length of the ubiquitin chain. The majority of substrates are

designated for degradation by the 26S proteasome. Recognition of substrate and ubiquitin attachment process involves three enzymes known as activating enzyme (E1), conjugating enzyme (E2) and ligase (E3) (Fig2).

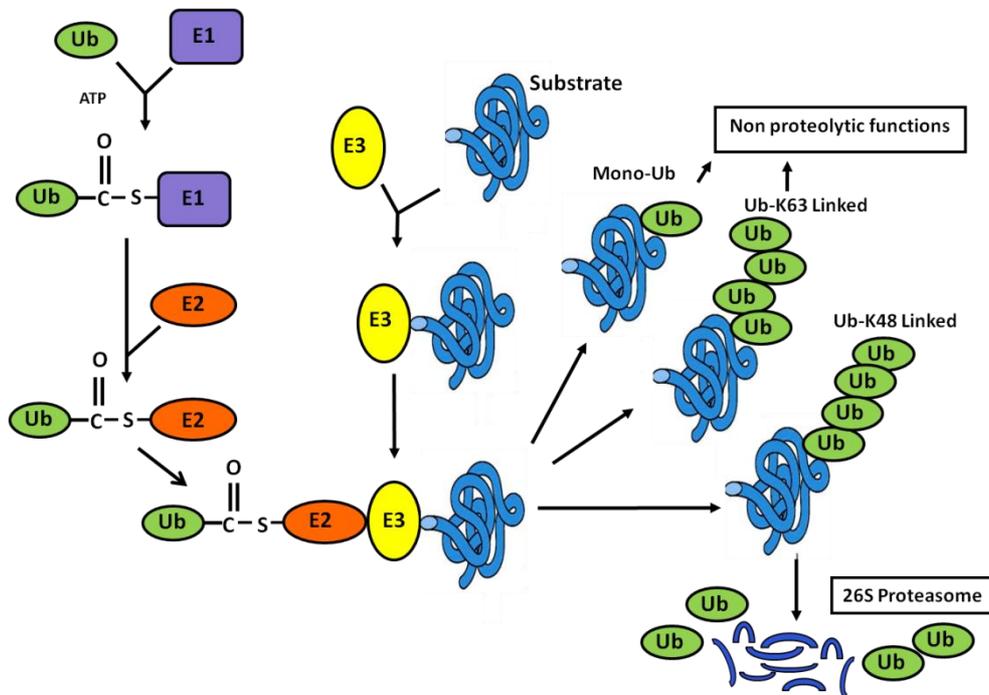


Figure 2 Simplified overview of steps involved in the ubiquitination process. E1-ubiquitin activating enzyme, E2-ubiquitin conjugating enzyme and E3-ubiquitin ligase. ub=Ubiquitin

These pathways begin with activation of ubiquitin by E1 enzyme; by utilizing ATP a thioester bond is formed between E1 and ubiquitin, by linking the C-terminal Gly-76 residue of ubiquitin on to a conserved cysteine residue within E1 (HATFIELD *et al.* 1990; HATFIELD and VIERSTRA 1992). This activated form of ubiquitin is transferred from E1 to a specific cysteine residue of E2 by transthioesterification. In the last step E3 ligase mediates the attachment of ubiquitin to substrates through an isopeptide bond between C-terminal Gly-76 of ubiquitin and a free lysine ϵ -amino group in the substrate. Poly-ubiquitination of substrates is a prerequisite for degradation via 26S proteasome (DOHERTY *et al.* 2002; WILKINSON 2000). In some cases, ubiquitin is transferred onto the E3 ligase before it is linked to the substrate and in some cases ubiquitin (ub) is transferred from E2 to the substrate, but in both the cases E3 are specifying which substrate to be ubiquitylated. It is also reported that in some cases E4, an additional factor is required for substrates poly-ubiquitination (KOEGL *et al.* 1999). The majority of the substrates that are poly-ubiquitinated via Lys 48 linked ub chains are destined for subsequent destruction by 26S proteasome. Because of the crucial role played by ub-Lys

48 linked chains in substrate degradation, one of the major topics of this Thesis research is to gain insight into functional importance of ub-Lys 48 and its downstream signaling in *Arabidopsis*.

1.1.3 Components of ubiquitination process

E1s – ub activating enzymes

E1s are the enzymes that catalyze the first reaction of the ubiquitination pathway. Two E1s reported in the genome of *Arabidopsis*, ubiquitin activating enzyme 1 (UBA1) and ubiquitin activating enzyme 2 (UBA2). These E1s have conserved cysteine that facilitates binding of ub and a nucleotide binding motif which can interact with ATP or the AMP-ub (HATFIELD *et al.* 1997). The main function of these enzymes is to activate ubiquitin. They are catalytically very active ensuring the levels of activated ub required by downstream activities. Localization studies revealed that E1s are present in most of the tissues and one is suspected to be nuclearly localized (HATFIELD *et al.* 1997). Mutational studies revealed that UBA1 accounts for more physiological functions in comparison to UBA2 (GORITSCHNIG *et al.* 2007). As these enzymes function upstream of the substrate recognition step, they have no impact on substrate specificity (PICKART 2001), but only transfer activated ubiquitin to E2s.

E2s –ub conjugating enzymes

E2s function downstream to E1s, and have a conserved UBC domain consisting of 150 amino acids which, serve as core domain. This ubiquitin conjugation domain (UBC) consists of catalytically active cysteine residue; on to it ubiquitin is transferred from E1 by a transthiol esterification. In plants, a large family of E2s exists. In *Arabidopsis* so far 37 E2s are reported (KRAFT *et al.* 2005) and are clustered into 12 subfamilies (BACHMAIR *et al.* 2001; VIERSTRA 1996). E2s are very heterogenous varying in size from 14 to over 100 kDa, show substantial variation in the amino acid sequence within the core domain but are still able to fold into a similar three dimensional structure (COOK *et al.* 1992; COOK *et al.* 1993). Some E2s are restricted in having only a core domain, whereas certain E2s have extended N- and C-termini (JENTSCH 1992; MERCHANT and BOGORAD 1986), which probably direct them for correct localization or specify interaction with suitable E3s.

E2 isoforms from yeast and animals display specific functions ranging from cell cycle regulation; DNA repair to ER translocated protein degradation (HERSHKO and CIECHANOVER 1998; PICKART 2001). Functional importance of plant E2s is still not clearly known because

of lack of availability of mutant information. For UBC24, known as PHO2, it has been explained in the literature that mutation in this gene influences phosphate signaling (BARI *et al.* 2006). Based on the orthology to yeast UBC6, the *Arabidopsis thaliana* UBC6 E2 subfamily is suspected to have a role in targeting ER retrotranslocated proteins for degradation (KOSTOVA and WOLF 2003). *Arabidopsis* E2s belonging to the UBC8 family interact with a wide variety of E3s *in vitro*, supporting their wide expression. Based on their expression it suggests probably the UBC8 family is the most functionally interacting E2s in *Arabidopsis*. (BATES and VIERSTRA 1999; GIROD *et al.* 1993; HARDTKE *et al.* 2002; SEO *et al.* 2003; XIE *et al.* 2002).

In *Arabidopsis*, like other eukaryotes, a family of Ubiquitin-conjugating E2 enzyme variant (UEVs) is present. There are eight genes found to encode putative UEVs in *Arabidopsis* (KRAFT *et al.* 2005). These UEVs have the conserved E2s catalytic core domain but lack the catalytic cysteine (BROOMFIELD *et al.* 1998; SANCHO *et al.* 1998) hence they lack the ability to participate in direct conjugation of ubiquitin. In order to be functionally active they require additional E2s as in the case of yeast UEV Mms2, which needs to interact with E2 UBC13 to make ub-63 linked poly-ubiquitin chains (HOFMANN and PICKART 1999; VANDEMARK *et al.* 2001). Another UEV, the COP10 of *Arabidopsis*, seems to interact with a number of E2s and helps formation of ub Lys-48 and Lys -63 linked poly-ubiquitin chains *in vitro* (YANAGAWA *et al.* 2004). It is also reported to interact with Cullin4-based E3 ligase (CHEN *et al.* 2005). With the known information it's clear that E2s are intermediate players between E1s and E3s, transfer ubiquitin either onto E3s or onto substrates selected by E3s. In either case they have no impact on substrate specificity in general.

E3s –ub ligating enzymes

E3 enzymes play a very significant role in spatial and temporal selection and ubiquitination of substrates, which is a crucial step in UPS and for various biological functions. In the *Arabidopsis* genome, around 1600 loci encode putative components of UPS, which account for 6% of the total proteome. Around 1300 genes encode E3 ligases (SMALLE and VIERSTRA 2004). Considering the E3s position and their number it is very clear that they are key components of the UPS system. E3s transfer ubiquitin from E2s to a free lysine ϵ -amino group in the substrate. Depending on the presence of domains and subunits, E3s either transfer ubiquitin directly from E2s to the substrates or E3s first binds ubiquitin on them before being transferred to substrates. Although different researchers reported different ways of E3s classification, in a simplified way E3 ligases can be classified into two major types,

HECT E3s (Homology to E6AP C-Terminus) and RING E3s/ U-box (Really Interesting New Gene). There are few E3s specifically functioning in the N-end rule pathway. These E3s and other components of the N-end rule pathway are one of the major topics of this Thesis research work hence they are discussed in more detail in subsequent parts of this Thesis and other E3 classes are explained below.

1.1.4 Types of E3 ligases and their functional importance

HECT-ligases (single subunit)

HECT E3s are single subunit polypeptides that range from 100-400 kDa, with a C-terminal signature domain called HECT domain formed by 350 amino acids that includes a conserved cysteinylsulfhydryl residue. In HECT E3s, ubiquitin is first transferred onto the conserved cysteinylsulfhydryl group on the HECT domain, before being transferred to a lysine residue on the substrate (PICKART 2001; SCHEFFNER *et al.* 1995). In addition to this domain, they are also known to have other protein-protein interaction sites such as RING-finger domain, coiled-coil or SH3 domain. These interacting sites are suspected to be involved in substrate recognition and or in protein localization. Many HECT E3s are present in animals, up to 50 putative HECTs reported in human (SCHWARZ *et al.* 1998). Unlike animals, plants are reported to contain a smaller number of HECTS.

Arabidopsis contains a small family of HECT proteins, consisting of 7 members (*UPL1* to *UPL7*) (DOWNES *et al.* 2003). *UPL1* and *UPL2* were identified by Bates et al (BATES and VIERSTRA 1999), whereas *UPL3* to 7 were identified by Downes et al (DOWNES *et al.* 2003). *UPL3* mutation leads to a similar phenotype to *spy5*, a constitutive GA mutant, which shows supernumerary trichome branching, which may result from disruption of gibberellic acid (GA) mediated trichome development (PERAZZA *et al.* 1998). *UPL3* mutation leads to hypersensitivity to GA, which leads to increased hypocotyl elongation upon GA treatment, but other GA related responses such as flowering and germination, are not affected. The direct substrates targeted by *UPL3* are not known. Miao et al. showed recently that *UPL5* mediates degradation of WRKY53, a transcription factor involved in senescence. *UPL5* is a negative regulator of senescence (MIAO and ZENTGRAF 2010). Although the function of the remaining *UPLs* is not known, studies of *UPL3* and *UPL5* clearly indicate that HECT E3s are important for plant life.

RING E3 Ligases

The *Arabidopsis* genome encodes a large family of RING E3 ligases; they are further sub-classified into single component RING/U-box and multi component complexes such as Cullin Ring Ligases (CRLs) and Anaphase Promoting Complex/Cyclosome (APC/C). RING E3s have specific domains that serve as docking site to interact non-covalently with E2s charged with ubiquitin. In addition they have a direct substrate binding site or adaptor binding site as in case of CLR (SEOL *et al.* 1999). Unlike HECT E3s, ubiquitin is not transferred onto these E3s, they tether E2s charged with ubiquitin and substrates facilitating close enough proximity for transfer of ubiquitin from E2s to respective substrates.

U-box E3s RING finger derived ligases

U-box E3 ligases contain a conserved 70 amino acid U-box domain that is known to be structurally similar to the RING domain E3s. Initially, the U-box was identified in yeast UFD2 (Ubiquitin Fusion Degradation2) protein (KOEGL *et al.* 1999). Unlike the RING domain, the U-box lacks the scaffold stabilizing Zinc-chelating cysteine and histidine residues but is still able to adopt a RING structure by depending on intramolecular interactions of salt bridges and hydrogen bonds to stabilize the scaffold. It is able to function as E3 ligase and promotes substrate degradation (ARAVIND and KOONIN 2000; HATAKEYAMA *et al.* 2001; JIANG *et al.* 2001). Plants do exhibit a significantly higher number of U-box genes in comparison to yeast and humans. The *Arabidopsis* genome contains 64 U-box genes in comparison to 21 and 2 U-box genes in human and yeast, respectively (AZEVEDO *et al.* 2001; KOSAREV *et al.* 2002; WIBORG *et al.* 2008). Plant U-box (PUBS) E3s are sub-classified depending on other domains they contain, such as Armadillo repeats (ARM), and WD40 repeats (ANDERSEN *et al.* 2004; KOSAREV *et al.* 2002; MUDGIL *et al.* 2004; SAMUEL *et al.* 2006). Genetic analysis and mutant studies revealed that U-box E3s are involved in several biological processes. ARC1 is a positive regulator of the self-incompatibility (STONE *et al.* 2003), *AtPUB9* plays a role in abscisic acid (ABA) hormone response (SAMUEL *et al.* 2008), AtCHIP contains a tetratricopeptide repeat (TPR) and is thought to be implicated in abiotic stress response (DAI *et al.* 2003; QIAN *et al.* 2006; YAN *et al.* 2003). A recent work has provided evidence that AtCHIP plays a role in protein degradation in the chloroplast (SHEN *et al.* 2007a; SHEN *et al.* 2007b). More interestingly AvrPtoB is a U-box E3 ligase that is transferred by the plant pathogen *Pseudomonas syringae* pv. *tomato* DC3000 to plant cells, suppressing plant immune response by evading basal defense responses. It helps to inhibit cell death and thereby ensures bacterial virulence (ABRAMOVITCH *et al.* 2006; ABRAMOVITCH and

MARTIN 2005; DE TORRES *et al.* 2006; JANJUSEVIC *et al.* 2006). The rice plant gene *Spotted leaf11* (*spl11*) encodes a U-box E3 ligase shown to have a role in PCD (ZENG *et al.* 2004). Taken together, it is clear that PUBS do function as E3 ligases and influence a plethora of plant developmental processes.

Cullin RING E3 ligases (CRLs)

Arabidopsis genome level search based on sequence similarity identified 11 predicted CULLIN-related genes. Only five are putative functional genes (CUL1, CUL2, CUL3a, CUL3b, and CUL4). All multi-subunit CRLs can be classified into 3 major subclasses CUL1, CUL3 and CUL4 (GINGERICH *et al.* 2005; RISSEEUW *et al.* 2003; SHEN *et al.* 2002) (Fig 3).

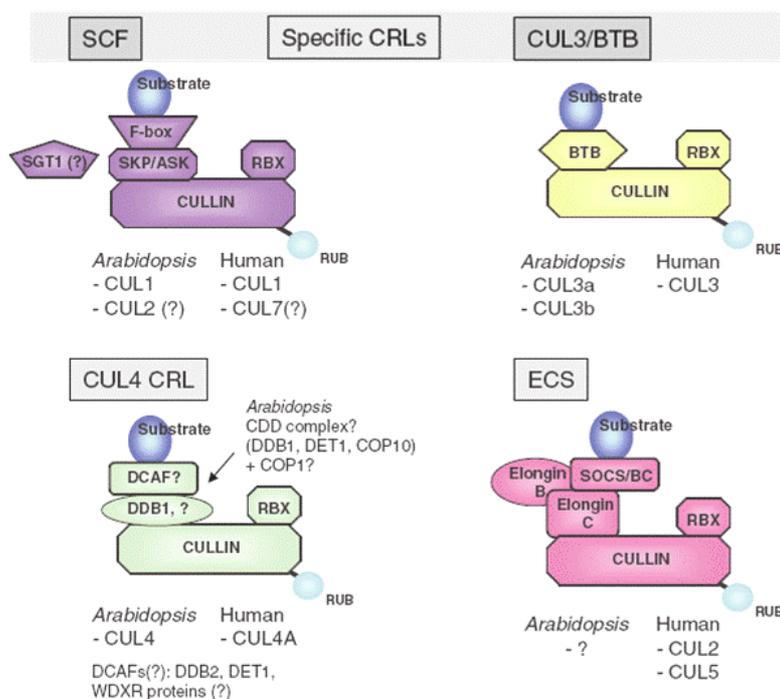


Figure 3 Cullin based ubiquitin ligases. Figure taken from (DREHER and CALLIS 2007). This picture depicts an overview of CRLs; details of each type are explained under topic CRLs.

In all the CRL subclasses, the CULLIN subunit functions as a scaffold protein characterized by a conserved 150 amino acid CULLIN domain, facilitating binding of RBX1 at the C-terminus, to which E2 can bind. Substrate recruiting adaptors bind to the cullin at the N-terminus. Hence all CRLs have the substrate binding motif and E2 binding motif on different parts of the single cullin subunit.

The CUL1 complex consists of four components: *Arabidopsis* S-phase Kinase associated Protein1 (ASK1), CUL1, substrate recruiting F-box protein and RING-Box 1 (RBX1) (Fig 3). In *Arabidopsis*, there are above 700 F-box proteins predicted (SMALLE and VIERSTRA 2004). CUL2 is closely related to CUL1 and builds SCF complexes (GRAY *et al.* 1999). Mammalian and yeast SCF structure has been resolved (HUIBREGTSE *et al.* 1995). Recent identification of three dimensional structure models of two *Arabidopsis* SCF subunits revealed mechanistic details (SHEARD *et al.* 2010; TAN *et al.* 2007).

The second subclass, E3 CUL3 complex, contain BTB/POZ (Bric a brac, Tramtrack and Broad complex/Pox virus and Zinc finger) domain proteins which function as substrate specific adaptors, CUL3 and the RBX1 protein(PINTARD *et al.* 2004)(Fig 3). In *Arabidopsis* 80 putative BTB proteins were found. Some of them also contain additional domains such as MATH and ankyrindomains known to be involved in protein-protein interaction (MICHAELY and BENNETT 1992; XU *et al.* 2003).

The third subclass E3, CUL4 complex contains DDB1, CUL4 and RBX1. DDB1 (UV-Damaged DNA –Binding Protein 1) functions as a substrate recruiting subunit either alone or in combination with De-Etiolated-1 (DET1) and Constitutively Photomorphogenic-1 (COP1) proteins (MCCALL *et al.* 2005) (Fig 3).

In humans, one more type of CRL has been reported that interacts with the adaptor proteins elongin B/C and recruits substrates to human CUL2 and CUL5 (KAMURA *et al.* 2004) (Fig 3).Neither this CUL2 nor other components of this complex show homology to components in *Arabidopsis*, except elongin C adaptor protein, which has one distant relative in *Arabidopsis* (RISSEEUW *et al.* 2003; SHEN *et al.* 2002).

1.1.5 APC/C complex

The anaphase-promoting complex/cyclosome (APC/C) protein complex has E3 ligase function and consists of 11 subunits (GIEFFERS *et al.* 2001). Two subunits of the APC complex, APC2 and APC11 show homology to subunits in SCF, cullin and a RING protein respectively. These two subunits together form the minimal ub-ligase complex of the APC (TANG *et al.* 2001).The *APC3/HOBBIT* (HBT)gene mutation leads to strong defects in meristem organization, giving direct indication that HBT is crucial for cell division and differentiation (BLILOU *et al.* 2002). In plants like in animals, targets of APC are cyclins; in many plant cyclins a conserved D-box is identified. Hence the APC appears to target them to

regulate cell division and differentiation in early stages of development (BLILOU *et al.* 2002; RENAUDIN *et al.* 1996).

1.1.6 The 26S proteasome: a site for protein breakdown

The 26S proteasome is present in the cytoplasm and nucleus of all eukaryotes (BOOK *et al.* 2009; YANG *et al.* 2004). It is a 2.5 MDa ATP dependent protease complex, consisting of 20S core particle (CP) and 19S regulatory particles (RP). The 20S core is made up of 4 rings of 7 different β and 7 different α subunits and forms a barrel shape (FU *et al.* 1999; GROLL *et al.* 1997). The inner β -subunits are the main substrate destruction sites having chymotryptic, tryptic and caspase like proteolytic activity. The 19S regulatory part is present at both ends of core particle. The 19S complex has 11 subunits and is divided into lid and base. The 19S lid part recognizes poly-ubiquitylated proteins, whereas the base consists of 6 RP triple A (AAA⁺) ATPases (RPTs 1-6) that help in unfolding of substrates and feeding them into the 20S core subunit for further degradation. The RP non-ATPases (RPNs) 10 and 13 function as ubiquitin receptors and RPN11 is involved in deubiquitination and release of attached ubiquitins from substrate (FU *et al.* 1998; HUSNJAK *et al.* 2008). Some RPs show specific hormone signaling function, e.g. RPN10 and RPN12a affect ABA and cytokinin signaling, respectively (SMALLE *et al.* 2002; SMALLE *et al.* 2003). It is also known that DUBs, E3s and a range of accessory proteins such as radiation sensitive 23 (Rad23), DNA-damage inducible 1 (Ddi1), the Hsp70 chaperone and proteasome assembly factors interact at substoichiometric level with the 26S proteasome (SCHMIDT *et al.* 2005).

1.1.7 De-ubiquitinating enzymes (DUBs)

De-ubiquitination is mainly involved in releasing poly-ubiquitin chains from substrates and ensures the existence of pool of reusable ubiquitin moieties. De-ubiquitinating enzymes (DUBs) are also known to act in proofreading or as processing enzymes of ubiquitin precursors. Like other eukaryotes, plants contain several DUBs (WEISSMAN 2001; WING 2003; YAN *et al.* 2000). The *Arabidopsis* genome contains nearly 30 genes that encode putative De-ubiquitinating enzymes (DUBs). These DUBs are subdivided into two main classes; the ubiquitin carboxyl-terminal hydrolase (UCH) class that includes two enzymes, and the ubiquitin specific processing protease (UBP) class, which includes 27 members (JOHNSTON *et al.* 1999). *AtUBP14* is involved in recycling free ub units (DOELLING *et al.* 2001). RPN11 is one subunit of the 26S proteasome lid particle, also functions as DUB. All

functional DUBs recognize ubiquitin and detach any amino acid or peptide bound to the C-terminal glycine (DOELLING *et al.* 2001; YAN *et al.* 2000).

In a recent work it was showed that UBP19 is involved in growth regulation in normal and malignant cells further providing evidence for the role of deubiquitinating enzymes in biological processes (LU *et al.* 2011). In *Arabidopsis*, AtUBP12 and AtUBP13 function as negative regulators of plant immunity (EWAN *et al.* 2011).

1.2 The N-end rule pathway

The N-end rule pathway (NERP) is a ubiquitin dependent selective protein degradation pathway present in all eukaryotes. The N-terminal sequence of a protein substrate impacts on its stability in the cell; this phenomenon was termed the N-end rule. Initially this pathway was discovered in the 1980s in the laboratory of Alexander Varshavsky (BACHMAIR *et al.* 1986; BACHMAIR and VARSHAVSKY 1989). A set of experiments proved that when certain amino acids such as Met, Thr, Ser, Gly and Val are present at the N-terminus, the proteins are relatively stable. However, when amino acids such as Lys, Arg, His, Phe, Tyr, Trp, Ile, Leu, Asp, Asn, Glu, Gln and Cys are present, the proteins are either directly degraded or enzymatic modified which in turn channels them for degradation via the ubiquitin dependent N-end rule pathway.

In this pathway, substrate recognition involves the same enzymatic cascade as in the case of UPS (E1, E2, and E3) (Fig 4). NERP E3s are central to this pathway and show specificity to substrates (Fig 4). Different E3s shows preference to bind to different proteins mainly depending on the type of N-terminal residue (MOGK *et al.* 2007; VARSHAVSKY *et al.* 2000).

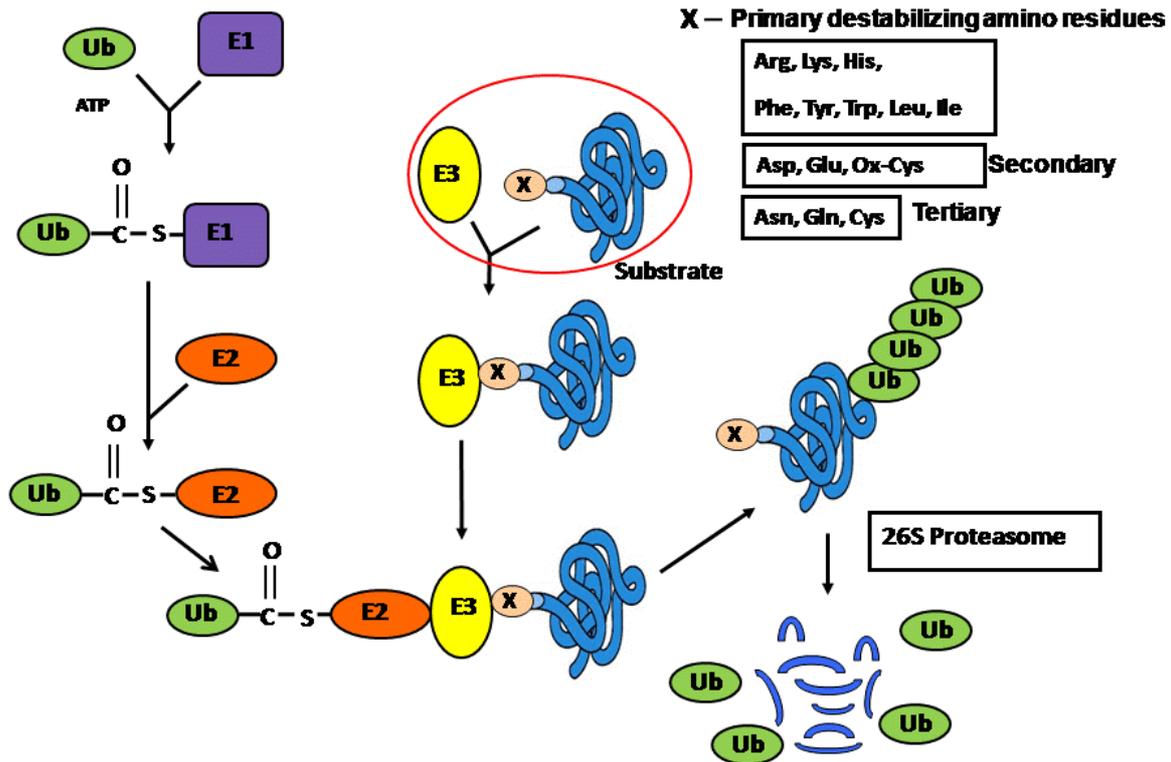


Figure 4 Overview of steps involved in the ubiquitination process of the N-end rule pathway. E1-ubiquitin activating enzyme, E2-ubiquitin conjugating enzyme and E3-NERP ubiquitin ligating enzyme. X-denotes primary destabilizing residues.

These destabilizing amino acids can be divided into three major classes, primary, secondary and tertiary. Only substrates with primary residues at the N-terminus are directly recognized by E3 ligases and further processed for degradation by the 26S proteasome (TURNER *et al.* 2000; Varshavsky, 1997 #132). The primary residues fall into two categories, basic (type 1 Arg, His and Lys) and bulky hydrophobic residues (type 2 Phe, Tyr, Trp, Leu and Ile). The secondary destabilizing residues are Asp, Glu and ox-Cys (oxidized-Cys). The tertiary residues encompass Asn, Gln and Cys. However, secondary and tertiary residues need to undergo arginylation and deamidation, respectively, in order to become primary substrates. Whereas the tertiary type of Cys residue needs a non-enzymatic NO and O₂-mediated modification in order to be recognized by E3s and as a consequence undergoes degradation. A very recent advancement in the understanding of this pathway is the discovery N-terminal acetylation process that generates substrates for N-end rule specific E3 ligase in yeast (FROTTIN *et al.* 2006; HWANG *et al.* 2010).

1.2.1 Enzymatic and Non-enzymatic modifications of N-end rule pathway

Deamidation and Arginylation processes

In mammals, tertiary residues Asn and Gln are converted into secondary residues Asp, Glu respectively through enzymatic modification by N-terminal amidases NTAN1 and NTAQ1, respectively (Fig 5) (GRIGORYEV *et al.* 1996; WANG *et al.* 2009). In yeast there is only the single amidase NTA1 that modifies tertiary residues into secondary ones. In mammals, N-terminal Cys undergoes oxidation in the presence of NO and oxygen and turns into ox-Cys (Fig 5) (GRIGORYEV *et al.* 1996; WANG *et al.* 2009). *Arabidopsis* contains homologs of mammalian deamidases but the functional importance is not known. NO/O₂ mediated Cys oxidation is also not known in plants. These unsolved points are addressed in this Thesis work. It is very interesting to know if nitric oxide (NO) signaling is mediated via N-end rule pathway in plants.

Arginylation is a process where Arg is attached to substrates which possess at their N-terminus secondary (Asp, Glu and ox-Cys) destabilizing residues (Fig 5). This reaction is carried out by an arginyl-tRNA-protein transferase (R-transferase). As a consequence of arginylation, protein degradation takes place (BALZI *et al.* 1990; CIECHANOVER *et al.* 1988; GONDA *et al.* 1989). The arginylation function was also reported in plants (YOSHIDA *et al.* 2002).

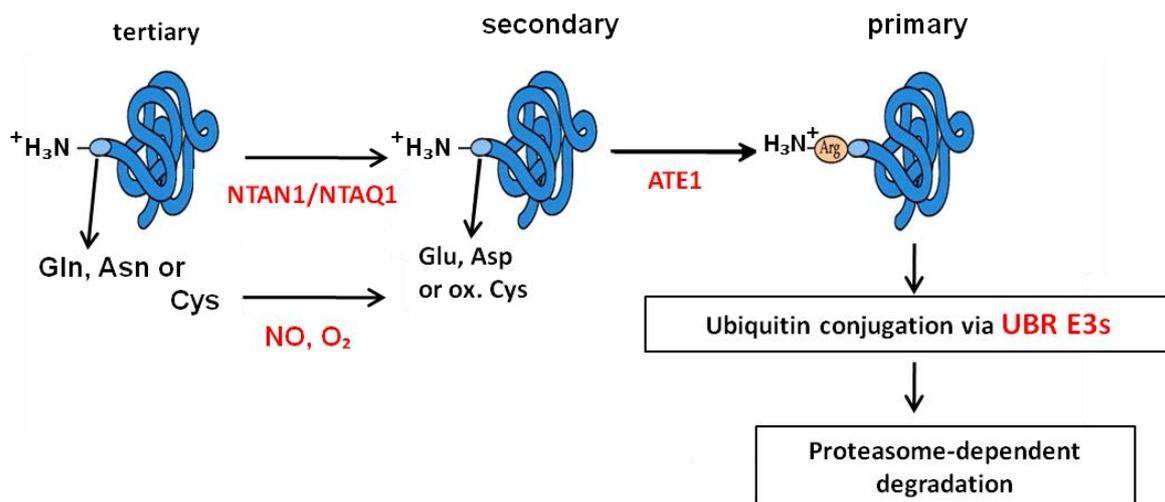


Figure 5 An overview of enzymatic and non-enzymatic steps involved in mammalian N-end rule pathway.

1.2.2 E3 ligases of the N-end rule pathway and their specificity

The first E3 ligase identified in the NERP is the UBR1 from yeast, which recognizes primary destabilizing signals (type 1 and type 2) of the substrate (HERSHKO *et al.* 1986). Tasaki and colleagues have identified a family of E3 ligases (UBR1-UBR7) that function as mammalian E3 ligases (TASAKI *et al.* 2005). Mutational studies in *Arabidopsis* revealed interesting insights into this pathway in plants. The E3 ligase *PROTEOLYSIS1* (*PRT1*) recognizes only aromatic residues and functionally replaces UBR1 of yeast (POTUSCHAK *et al.* 1998; STARY *et al.* 2003). A more recent finding in plants identified a second E3 ligase, *PROTEOLYSIS6* (*PRT6*) that targets primary basic residues such as Arg (GARZON *et al.* 2007). These two known E3s clearly show differences in their domain structures as well as in substrate specificity. Neither of these plant-specific E3s, *PRT1* and *PRT6* recognizes a Leu residue, which is considered as primary destabilizing residue. A report from Tasaki and colleagues (TASAKI *et al.* 2005) lists two plant proteins, At3g02260 and At4g23860, with a Zinc finger-like domain called UBR-box, which is a conserved domain in all known mammalian UBR proteins. Hence these two are suspected to be potential E3 ligases. Based on known facts these two proteins are considered for analysis in this Thesis work to see if any of these two could function as plant E3 ligase, e.g. binding aliphatic hydrophobic destabilizing residues such as Leu or any other primary destabilizing residues.

1.2.3 An overview of functional importance of ubiquitin system components and their implication in cell death processes

Firstly in the ubiquitin system, ubiquitin's seven Lys residues (Lys6, Lys11, Lys27, Lys29, Lys33, Lys48, and Lys63) proved to influence substrate fate by forming different chain length and linkage types. Such post-translational modification of substrates influences cellular localization, protein-protein interaction, alters recognition and promotes degradation via 26S proteasome (FINLEY 2009; GLICKMAN and CIECHANOVER 2002; HICKE and DUNN 2003; IKEDA and DIKIC 2008; MUKHOPADHYAY and RIEZMAN 2007; PICKART and FUSHMAN 2004). A well known consequence of ubiquitination is breakdown of substrates at the 26S proteasome, which accounts for degradation of most cytosolic, nuclear, endoplasmic reticulum lumen/membrane proteins and mitochondrial proteins (FINLEY 2009; GLICKMAN and CIECHANOVER 2002). The majority of the proteasome substrates are marked with poly-ubiquitin chains formed using Lys 48. Among ubiquitin's seven Lys residues, Lys48 is the only lysine, whose replacement with arginine is lethal, emphasizing the essential and unique role for Lys48-linked chains (FINLEY *et al.* 1994). Importance of Lys 48-linked chains is

studied in plants by expressing a ubiquitin variant with arginine at Lys48 position, results of this study revealed a possible role of ubiquitination in cell death processes in plants (SCHLOGELHOFER *et al.* 2006). The current thesis work is a step forward in deciphering molecular candidates involved in this process.

E3 ligases of the ubiquitin proteasome pathway are central to the substrate recognition and degradation. Some of the E3s are known to act as positive and negative regulators of cell death process in animals and in plants. Mdm2, E3 ligase targets tumor suppressor gene products such as p⁵³(HONDA *et al.* 1997) and retinoblastoma proteins (pRB)(MIWA *et al.* 2006; UCHIDA *et al.* 2005), for ubiquitin-mediated degradation. SCF^{skp2} targets other tumor suppressor gene products such as p130, Tob1, p27^{kip1}, p57^{kip2}, and p21^{cip1}(CARRANO *et al.* 1999; GANOTH *et al.* 2001; HIRAMATSU *et al.* 2006; KAMURA *et al.* 2003; NAKAYAMA *et al.* 2000; TEDESCO *et al.* 2002). These two act like oncogene products. However, SCF^{Fbw7} is involved in degradation of several oncogene products such as cyclin E, Notch, c-Myc, c-Jun, and c-Myb(KANEI-ISHII *et al.* 2008; KITAGAWA *et al.* 2009; KOEPP *et al.* 2001; WEI *et al.* 2005; WELCKER *et al.* 2004; WU *et al.* 2001; YADA *et al.* 2004). Fbw7 is often either deleted or mutated in human cancers and acts like a tumor suppressor. Mutations in oncogene products or suppressors deregulate cell death. Caspase modified DIAP1, a member of an IAP1 family of PCD inhibitors, in *Drosophila* is channeled into the N-end rule pathway, a process which is important for the correct regulation of apoptosis (DITZEL *et al.* 2003).

The ubiquitin proteasome dependent protein degradation pathway components are vital for cell death regulation in plants. At-PUB44 is also known as Senescence associated E3 ubiquitin ligase 1 (SAUL1) has been reported to prevent premature senescence (RAAB *et al.* 2009). A very recent study revealed *Arabidopsis* UPL5, a HECT E3 ligase, is a negative regulator of senescence (MIAO and ZENTGRAF 2010). SPL11, a rice U-box E3 ligase, functions as negative regulator of HR-associated leaf lesion formation and pathogen defense (ZENG *et al.* 2004). The AvrPtoB bacterial effector is a U-box type E3 ligase that functions in plant hosts and inhibits cell death and basal defense responses and ensures bacterial virulence (ABRAMOVITCH *et al.* 2006; JANJUSEVIC *et al.* 2006). *Arabidopsis* ATE1 is an arginyl-t-RNA: proteinarginyltransferase, which has been shown to regulate senescence by converting secondary residues into primary ones which are direct substrates of E3 ligase PRT6 (GARZON *et al.* 2007; YOSHIDA *et al.* 2002). The *ORE9/MAX2* gene encodes an F-Box protein with LRR domain, which forms a complex with SCF Cullin. It was identified in a genetic screen intended to identify delayed senescence (for *ORESARA* which means “long live” in Korean),

and *MAX2* is identified in a screen for plants with enhanced lateral branching (WARD and LEYSER 2004; WOO *et al.* 2001). A study by Trujillo *et al.* (2008) showed that *AtPUB22*, *23*, and *24* are involved in oxidative burst and cell death, acting as negative regulators of abiotic stress and plant defense responses (CHO *et al.* 2008; TRUJILLO *et al.* 2008).

All the above listed examples emphasize the role of the ubiquitin proteasome dependent protein degradation components in cell death processes, either as positive or negative regulators. It is also evident that the N-end rule mediated protein degradation pathway is essential to maintain proper cell death.

1.3 Background to the thesis

In previous work it was shown that in higher plants expression of a ubiquitin variant (ubiquitin Lys48 to Arg called as ubK48R) leads to cell death phenotypes (Bachmair et al 1990, Peter Schlogelhofer 2006). This provided first information that ubiquitin proteasome dependent degradation is linked to cell death processes. In *Arabidopsis* a progenitor line carrying ubR48 was EMS mutagenized to isolate survivors of ubK48R lethality. This screen identified 5 complementation groups. One promising line among them was *sud2* (suppressor of ubiquitin UbK48R-induced cell death) considered for mapping to identify responsible candidate(s) leading to cell death in ubK48R background. The rough mapping suggested the position of the mutation on chromosome 3 between markers MUO22 and CIW4. This formed the basis for one of the research topics of this Thesis work.

The basis for the second topic comes from N-end rule pathway research; previously it was shown that PRT6 and PRT1 are E3 ligases and target model substrates with basic and aromatic residues at the N-terminus, respectively (GARZON *et al.* 2007; POTUSCHAK *et al.* 1998). A Leu residue, which belongs to the group of primary residues, is a destabilizing residue in *Arabidopsis*. It is not stabilized by either of the known E3s, giving a hint at for the existence for at least one or more unknown E3 ligases in *Arabidopsis*.

1.4 Aim and strategy of the study

The main aim of this Thesis is to deduce the molecular links between ubiquitin dependent protein degradation pathways (UPS/NERP) and programmed cell death by using the model plant *Arabidopsis*. To this end, two independent projects were undertaken to decipher molecular mechanisms interlinking ubiquitin dependent protein degradation and cell death processes. The first project is aimed to identify the molecular candidates leading to cell death upon expression of ubiquitin variant (ubK48R) by using suppressor of cell death phenotype (*sud2*) *Arabidopsis* (Fig 6). To unravel this two different approaches, one at the genome level and the other one at the transcriptome level were designed.

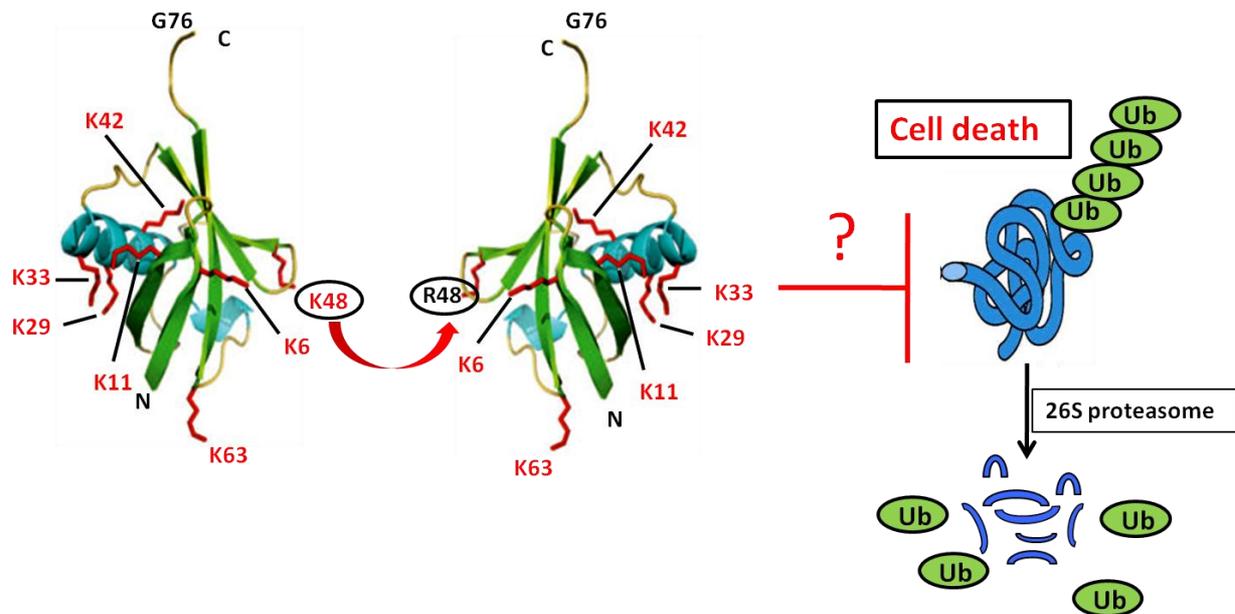


Figure 6 Ubiquitin variant inhibits poly-ubiquitination. A ubiquitin inhibition leads to downstream protein degradation deregulation via 26S proteasome. Ubiquitin structure is taken from (HUA and VIERSTRA 2011) and modified. Three dimensional ribbon model of plant ub (VIJAY-KUMAR *et al.* 1987).

The first one involves fine mapping of *sud2* candidate and identification of the molecular candidate(s) responsible for rescuing the cell death phenotype imposed by ubR48 in *Arabidopsis*. A forward genetics approach using EMS mutagenesis followed by map based identification of a candidate region responsible for specific phenotype is already well established method, several lines of evidence proved the importance of this method to plant research by allowing remarkable achievements in identification of key players in several biochemical pathways. A well-designed platform of high-density polymorphic markers is a backbone to this method in addition to very advanced next-generation sequencing methods. To achieve the first step in fine mapping of suppressor of ubR48 (*sud2*), a large mapping population was generated for genotyping and phenotyping. As there was very low recombination in the candidate region, an extra step aimed to generate a sub-genomic library of PCR fragments from the region of interest. Purified fragments subject to next generation sequencing and reads were aligned to wild type Columbia sequence to find the probable SNPs. On the other hand, to analyze the transcriptome, a microarray based transcriptome profiling was chosen. In both cases the final goal of this work was to find candidate genes and to confirm their involvement in cell death processes. These identified candidate genes would help to draw biologically meaningful connections between ubiquitin dependent protein degradation and cell death processes in plants.

The second project aimed at identifying various components of the N-end rule pathway and studying their role in cell death process in *Arabidopsis*.

The N-end rule pathway's known mutants suggest that this pathway regulates senescence, germination and apoptosis. This pathway is well characterized in yeast and mammalian systems in comparison to plants. PRT1 and PRT6 are two plant E3s responsible for targeting primary residues. However, though Leu is an aliphatic hydrophobic primary residue, it is not recognized by the known E3s, and this suggested the existence of one or more E3s in N-end rule pathway (Fig 7). The generation of reporter lines expressing chemically inducible ubiquitin fusion model substrates with R, L, D, M, N, Q, D, E, F, and C residues at the N-terminus of the test protein part of the constructs was planned. In a second step, EMS mutagenesis of a L-GUS expressing line leading to the identification of mutant lines, that stabilize L-GUS was envisioned. To achieve this, a live tissue GUS assay also needed to be developed.

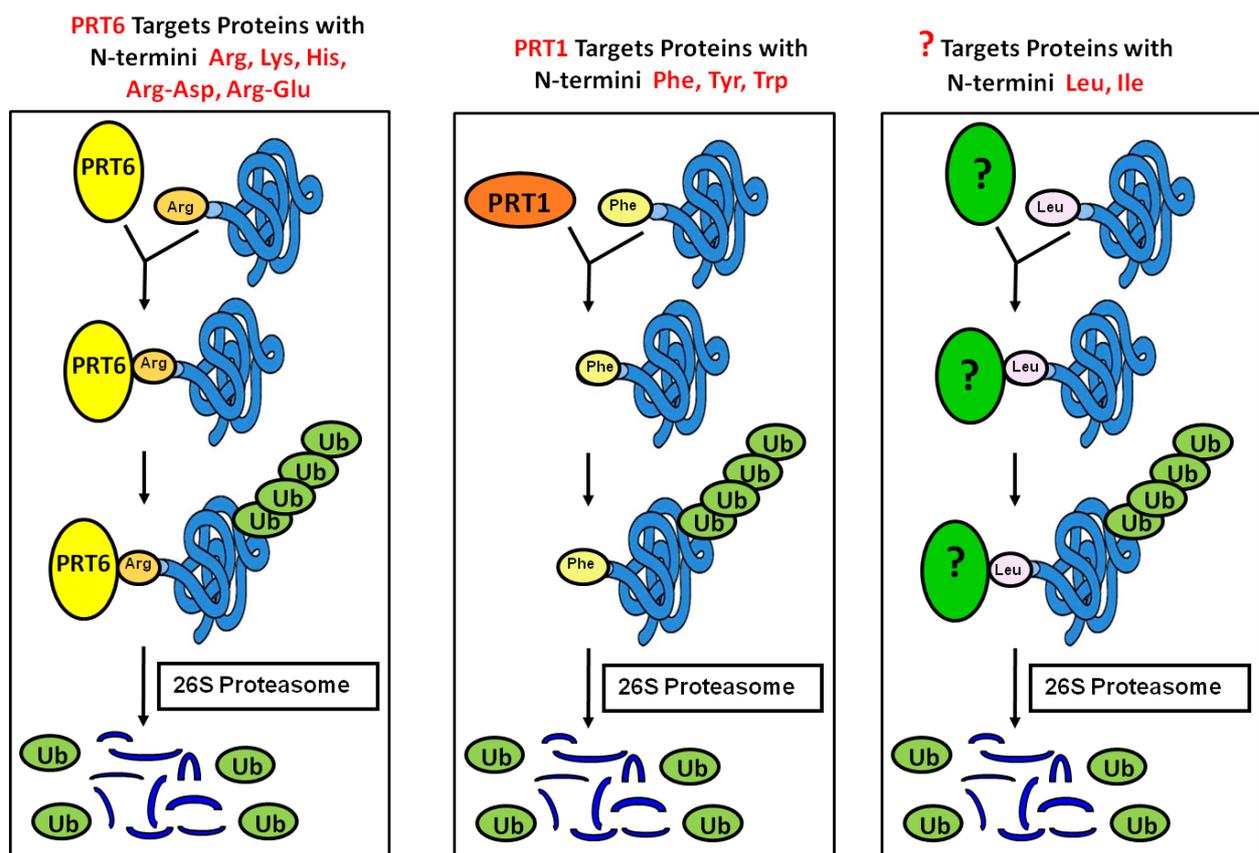


Figure 7 The N-end rule E3-Ligases of *Arabidopsis*. PRT6, PRT1 dependent ubiquitin mediated degradation via 26S proteasome of model substrate with basic and aromatic residues at N-termini. An unknown ligase degrades model substrate with aliphatic hydrophobic residue such as Leu at N-termini.

A literature based study suggests existence of two proteins, BIG (At3g02260) and PRT7 (At4g23860) that might function as E3s as they possess a UBR-box which is a common domain in known N-end rule E3s. Deamidation, a process where tertiary residues are converted into secondary residues is well studied in mammals. Molecular information about this mechanism is lacking in plants. Furthermore other findings revealed that distinct counterparts of mammalian deamidases do exist in plants, Ntan and Ntaq, but functional importance has not been deciphered yet (Fig 8). To understand if these plant distinct homologs function in a similar fashion to mammalian proteins and have any role in plant N-end rule pathway, T-DNA mutant lines and Tilling lines were considered for examination.

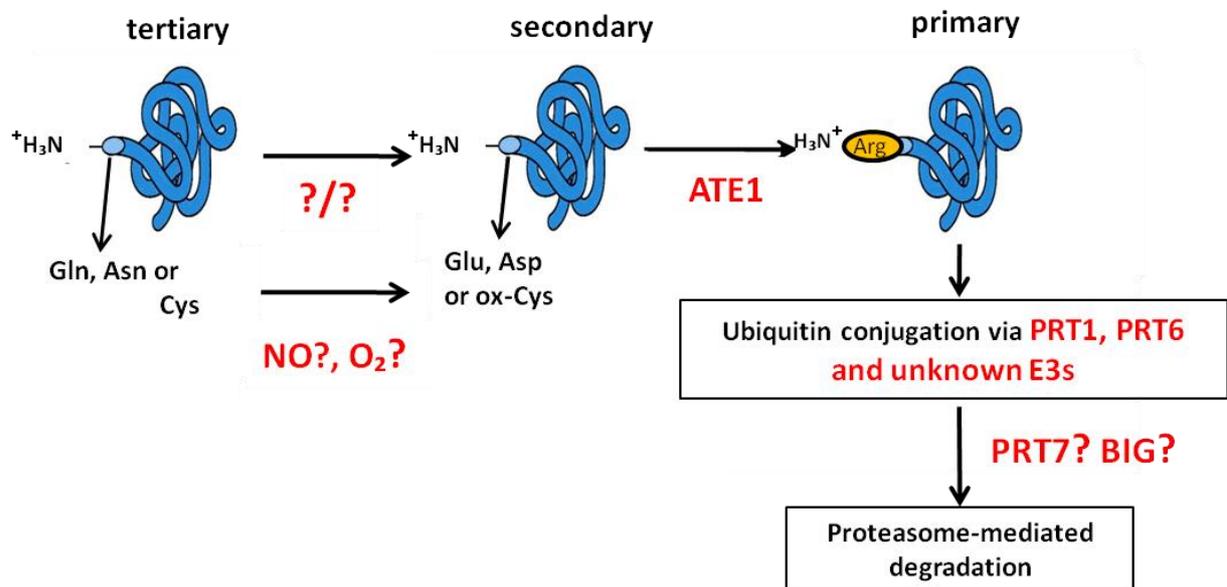


Figure 8 Unknown aspects of the N-end rule pathway in *Arabidopsis*.

It is also shown in other eukaryotes that NO mediated degradation functions via the N-end rule pathway, misregulation of this process leads to developmental defects. It is an open question to investigate, if NO has any role in converting tertiary Cys residue into ox-Cys, when present at the N-terminus of the test protein and further promotes its degradation through the N-end rule pathway in *Arabidopsis*. And also to see if this unknown branch has any role in plant physiological processes (Fig 8). These unknown areas of N-end rule pathway in plants formed another critical research subject of the current Thesis work.

2 RESULTS

2.1 Results part 1

Search for *suppressor of cell death (sud2)* candidates

2.1.A1 Genetic screen to identify cell death responsible candidates

In plants ubiquitin mediated protein degradation and cell death connections are not well known. To understand links between ubiquitin and cell death, a ubiquitin variant was generated by modifying Lys 48 to Arg (ubK48R). Expression of ubK48R leads to inhibition of poly-ubiquitin chains formed by Lys 48 and leads to cell death phenotype in *Arabidopsis*. To identify candidate(s) responsible for the lethal phenotype, a genetic screen was started. As a part of this process, suppressors of ubK48R mutants were generated. One of the suppressor line that was isolated in an EMS screen was named after its phenotype, the rescue of cell death caused by expression of ubK48R, *suppressor of ubiquitin variant (ub K48R) induced cell death (sud2)* (SCHLOGELHOFER *et al.* 2006). A classical mapping using suppressor of *ubiquitin variant induced cell death2 (sud2)*, would uncover a molecular candidate involved in phenotypic change in the progenitor line. This finding will provide connections between cell death and ubiquitin dependent protein degradation process.

2.1.A2 The *suppressor of cell death (sud2)* rough mapping from previous work

Initial rough mapping of *sud2* was performed by using 29 molecular markers on 26 mapping individuals. These markers had 20 cM distance and were spread over 5 chromosomes. The results of this analysis suggested the position of the *sud2* mutation on chromosome 3, to the south arm between the markers CIW4 (18.9 cM) and MNZ14 (10, 7cM). Although two more markers were analyzed from the south end of the chromosome 3, this could not further decrease the candidate region. To increase the recombination events and to generate a new mapping population, 3 lines from the existing rough mapping population (Mp# 5-1-50, Mp# 6-1-25 and Mp# 6-1-37) were crossed to two 86*Ler* lines (# 28-1 and #48-1; RV#86-5 introgressed into *Ler*) forming new lines to continue further analysis.

In this thesis these newly generated recombinants (derived from crossing) were used as starting material to continue the mapping process of suppressor of cell death mutant to

identify candidates responsible for causing cell death phenotype upon expression of ubiquitin variant (ubK48R).

2.1.A3 Generation of a large mapping population

The large mapping population was required to fine map the candidate suppressor of cell death in *sud2* genetic background. To attain this, new recombinants (F1) were used for the generation of large mapping population by selfing. From 37 F1 selfed plant lines, 192 F2 lines were generated. The details of these lines are shown in table 1. These lines were screened by phenotypic and genotypic analysis. The genotypic analysis was performed by analyzing the PCR-based markers.

Table 1: Details about the origin of large mapping population.

Cross	No. of F1 Selfed plants	No. of F2s lines generated
Mp#5-1-50 x 86Ler#28-1-2	9	54
Mp#6-1-25 x 86Ler#48-1-1	9	48
86Ler#48-1-1 x Mp#6-1-37-1	10	36
86Ler#48-1-1 x Mp#6-1-37-2	9	54
Total no. of plants	37	192

2.1.A4 Phenotyping of the new recombinant mapping population

The altered phenotype was examined by using a survival test on a selection medium containing of Hygromycin (Hyg 25µg/l), Dexamethasone (Dex 0.7µM), Methotrexate (Met 50 µg/l) (HDM). The selection media was used because the ubK48R expression was kept under the control of chemically inducible promoter. Hyg was for selection of the transgene, Dex for induction of the transgene, and Met was to test transgene expression in the induced plant line. First a control survival test was performed. To this end first the seeds of *sud2*, RV#86-5 and 86Ler were germinated on selection medium. The survival status of these lines was scored after two to three weeks. The survival of *sud2* clearly indicates that the *sud2* line can withstand lethal effect of ubK48R as it might be carrying an EMS mutation in the cell death responsible candidate gene. Transgenic *sud1* line is also an EMS suppressor but shows

low survival in comparison to *sud2*. This might have resulted from less direct influence on survival. *RV86-5* and *86Ler* are transgene (ubK48R) expressing lines but have no an EMS mutation. Thus these two lines underwent death because of lethal effect of transgene (Fig 9). This test served as basis for further phenotyping of the mapping population.

The newly generated 192 F2 mapping lines were germinated on HDM (Hyg, Dex and Mtx), plates to score the survival and for phenotyping. Phenotyping results were scored after two to three weeks. If all seedlings survived on selection medium that line was scored as homozygous (for suppressor mutation). In case all seedlings died that line was considered as wild type for the suppressor locus (i.e. only containing either the *RV86-5* or *86Ler* locus). In an intermediate scenario, where only 1/4 (a fraction) half of the seedlings survived, such a line was scored as heterozygous. The segregation scores of 192 individuals are summarized in the table 2. The results are clearly indicating that these individuals showed monogenic segregation for the *sud2* mutation.

Survival test on selection medium (HDM)



Figure 9 Survival test of two-week old seedlings of different genetic background lines carrying an inducible ubK48R transgene. *sud2* is the *RV86-5* progenitor line with EMS induced mutation in Columbia background, *RV86-5* is a progenitor line without EMS mutation in Columbia background, *86Ler* is a Landsberg *erecta* background transgenic line with introgressed ubK48R from *RV86-5* line, has no EMS mutation. *sud1* is another suppressor *RV86-5* progenitor line generated by EMS mutagenesis in Columbia background.

Table 2: Summarized result of phenotyping of mapping population. Results of survival test are obtained from three-week old seedlings.

Phenotype Segregation in F2 population	Observed number of segregates	Expected number of segregates
Homozygous (Suppressor)	40	48
Heterozygous	102	96
Wild type (<i>RV86-5</i> or <i>86Ler</i>)	50	48
Total	192	192

2.1.A5 Marker-based genotyping of the new recombinant mapping population

To identify the mutation on chromosome III, marker based genotyping was chosen. To accomplish this, recombinant lines of the mapping population were germinated on Hyg plates (for selection of the transgene *ubK48R*). From two-week old seedlings, DNA was extracted by using a 96-well biosprint automated machine. This genomic DNA was used as template for PCR-based amplification of markers. Since recombination was between *Col* and *Ler* genetic backgrounds, molecular markers showing polymorphism between these two ecotypes were selected for analysis. Before being used as markers on mapping individuals, they were first examined for their reliability as polymorphic markers between these two ecotypes. To this end, markers were first amplified using genomic DNA of these two ecotypes. These PCR based amplified marker products were separated on high percentage agarose gel to analyze product size differences. The markers that showed prominent polymorphism between the two ecotypes were further used for analysis of the mapping population. The single nucleotide polymorphism (SNP) markers that were chosen for analysis are listed in table 3. These four markers covered a genomic region of chromosome 3, south end, between 11.4 Mb and 18.9 Mb. The polymorphism was analyzed depending on the recombination event of the segregation population at that specific locus for that given marker. These results narrowed down the mapping region to 7.5 Mb. Figure 10 shows an example of SNP marker T6H20 analysis as homozygous for *Col/Ler* or heterozygous status of the particular recombinant lines (number) from the mapping population. According to the analysis, recombinant line 176

shows homozygosity for Col-0, 177 is heterozygous and line 178 is homozygous for *Ler*. Results for 192 recombinants and four markers are shown in table 4.

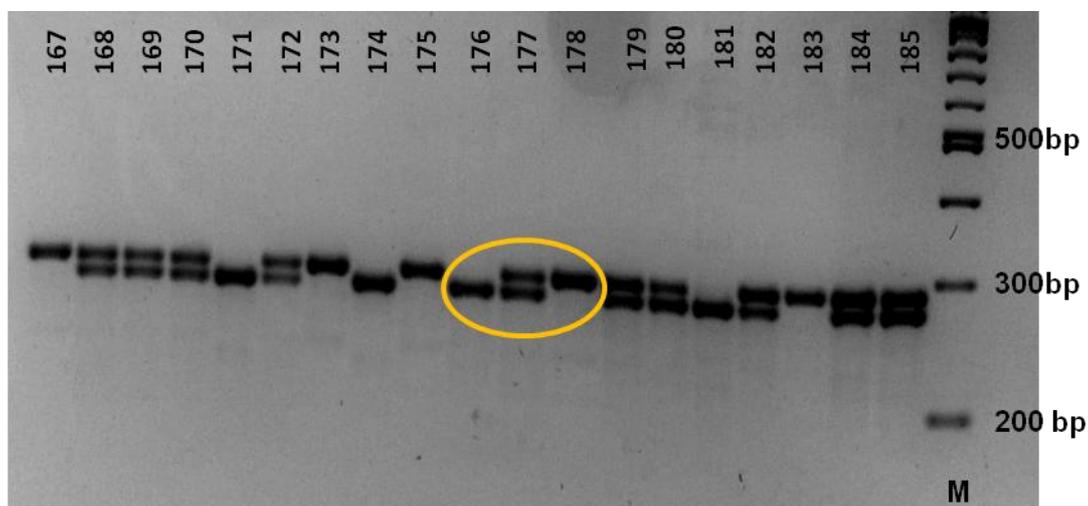


Figure 10 SNP marker T6H20 polymorphism analyses on recombinants, a sub-pool of mapping population. The SNP marker T6H20 PCR product of recombinants from lines 167 to 185 analyzed on high percentage agarose gel. The T6H20 marker product size for Col is 273 bp and for *Ler* is 293 bp. The number 176 recombinant line is homozygous for Col, whereas number 177 recombinant line is heterozygous and number 178 recombinant is homozygous for *Ler*.

Table 3: SNP markers between Col and *Ler* applied on new mapping population. Gene marker is the name of the marker tested; Map position is the position of the marker on chromosome III.

Gene marker	Map position (Mb)	PCR primers for amplification (5' to 3')	
		Forward primer	Reverse primer
MUO22	11.4	ATT GAT CAT ATC GCC CAA CAC	ACA TTG CAG CAG GAT AGG TTG
T32N15	16.36	ATC TGA AAA TCC TTG CGT GAG	TTG TGA CGA ATA GTG AAA GGA GAG
T6H20	17.2	CGG CTG AAA CTT GGA AGG GAC	AGG AAG AAC GTG TGA TTG TG
CIW4	18.9	GTT CAT TAA ACT TGC GTG TGT	TAC GGT CAG ATT GAG TGA TTC

Table 4 Phenotyping and SNP markers analyzed on 192 recombinants. For the colour code and details see foot note

Number	Phenotyping	Genotyping				Number	Phenotyping	Genotyping			
		11.4Mbp MUO22	16 Mbp T6H20	16.2 Mbp T32N15	18.9 Mbp CIW4			11.4Mbp MUO22	16 Mbp T6H20	16.2 Mbp T32N15	18.9 Mbp CIW4
1	CL	LL	CL	CL	51	CC	CL	CL	CL	CL	
2	CL	LL	CL	CL	52	LL	LL		LL	LL	
3	CC	LL	CC	CC	53	CL	CL	CC	CC	CC	
4	CL	LL	CL	CL	54	CL	CL	CL	CL	CL	
5	CC	LL	CC	CC	55	CL	CL	CL	CL	CL	
6	CL	LL	CL	CL	56	CL	CC		CC	CL	
7	CL	LL	CL	CL	57	CL	CL	CL	CL	CC	
8	CL	CL		CL	58	CL	CL	CL	CL	CL	
9	LL	LL	LL	LL	59	CL	CL	CL	CL	CL	
10	LL	LL	CL	CL	60	LL	LL	LL	LL	CL	
11	CL	LL	CL	LL	61	CL	LL	CC	CC	CC	
12	LL		LL	LL	62	CL	LL	CL	CL	CL	
13	CL	CL	CL	CL	63	CC					
14	CL	CL	LL	LL	64	CL	LL	CL	CL	CL	
15	CL				65	CL	LL	CL	CL	CL	
16	LL	LL	LL	LL	66	CL	LL	CL	CL	CL	
17	CL	CL	CL	CL	67	CC	LL	CC	CC		
18	LL	LL	LL	LL	68	LL	LL	LL	LL	LL	
19	CL	CC	CL	CC	69	CL	LL	CL	CL	CL	
20	CL	CL	CL	CL	70	LL	LL	CL	CL	CL	
21	CC	CL	CL	CL	71	CC	LL	CC	CC	CC	
22	LL	LL	LL	LL	72	CC	LL	CL	CL	CL	
23	CC	LL	LL	LL	73		LL	LL	CL	CL	
24	CL	CL	CL	CL	74	CC	LL	CC	CC	CC	
25	CL	CC	CC	CC	75	CL	LL	CL	CL	CL	
26	CC	CC	CC	CC	76	CC	LL	CC	CC	CC	
27	CL	CL	LL	CL	77	LL	LL	LL	LL	LL	
28	CL	CL	CL	CL	78	CC	LL	CC	CC	CC	
29	CC	LL	CL	CL	79	CL	LL	CL	CL	CL	
30	CL	CL	CL	CL	80	CL	LL	CL	CL	CL	
31	CL	LL	CL	CL	81	LL	LL	CL	CL	LL	
32	LL	LL	LL	LL	82	LL	LL	LL	LL	LL	
33	CL	LL	CL	CL	83	CC	LL	CC	CC	CC	
34	CC	LL	CC	CC	84	LL	LL	LL	LL	CL	
35	CC	LL	CC	CC	85	CL	LL	CC	CC	CC	
36	CL		CL	CL	86	CL	LL	CC	CC	CC	
37	CL	CL	CL	CL	87	LL	LL	CL	CL	LL	
38	CL	LL	CL	CL	88	LL	LL	CL	CL	CL	
39	LL	LL	LL	LL	89	CL	LL	CL	CL	CL	
40	CL	CL	CL	CL	90	CL	LL	CC	CC	CC	
41	CL		CC	CC	91	CL	LL	CL	CL	CL	
42	CL	CL	CL	CL	92	LL	LL			LL	
43	LL	LL	CL	CL	93	LL	LL	CL	CL	CL	
44	LL	LL	LL	LL	94	CC	LL	CC	CC	CC	
45	CL	CL	CL	CL	95	CC	LL	CL	CL	CL	
46	CL	CL	CL	CL	96	LL	LL	LL	LL	LL	
47	CL	CC	CL	CL	97	CC			CL		
48	CL	CL	CL	CL	98	LL			LL		
49	CL	CL	CL	CL	99	LL			LL		
50	CL	CL	CL	CL	100	LL	LL		CC	CL	

Table continues on following page

101	CC	LL		CC	CC	147	LL	LL	LL	LL	LL
102	CL	LL		CL	CL	148	LL	LL	LL	LL	LL
103	CC	LL	CL	CL	CL	149	LL	LL	LL	LL	LL
104	CL	LL			CL	150	LL	LL	LL	LL	LL
105	LL	LL	LL	LL	LL	151	CL	LL	CL	CL	CL
106	CC	LL	CL	CL	CL	152	CL	CL	CC	CC	CC
107	LL	LL	LL	LL	LL	153	CL	CL	CL	CL	CL
108	CL	LL	LL	LL	LL	154	LL	LL	LL	LL	LL
109	CC	CC	CC	CC	CC	155	CL	CL	CL	CL	CL
110	LL	LL	LL	LL	LL	156	CL	LL	LL	LL	CL
111	CL	CL	CL	CL	CL	157	CC	LL	CC	CC	CC
112	CC	CC	CC	CC	CC	158	CC	LL	CC	CC	CC
113	CL	CL	CL	CL		159	CL	LL	CL	CL	CL
114	CL	LL	LL	LL	CC	160	CC	LL	CC	CC	CC
115	LL	LL	LL	LL	LL	161	LL	LL	LL	LL	LL
116	LL	LL	LL	LL	LL	162	LL	LL	LL	LL	LL
117	CL	CL	CL	CL	CL	163	CL	CL	CL	CL	CC
118	CL	CL	CL	CL	CL	164	LL	LL	LL	LL	LL
119	CC	CC	CL	CC	CL	165	CL	LL	CL	CL	CL
120	CL	CL	LL	LL	LL	166	CL	LL	CL	CL	CL
121	CL	CL		CL		167	LL	LL	LL	LL	LL
122	LL	LL	LL	LL	LL	168	CL	LL	CL	CL	CL
123	CL	CL	CL	CL	CL	169	CL	LL	CL	CL	CL
124	CL	CL	CL	CL	CL	170	CL	LL	CL	CL	CL
125	LL	LL	LL	LL	LL	171	CL	LL	CC	CL	CC
126	CL	CL	CL	CL	CL	172	CL	LL	CL	CL	CL
127	CL	LL	LL	LL	LL	173	LL	LL	LL	LL	LL
128	CC	CC	CC	CC		174	CC	CC	CC	CC	CC
129	CL	CC	CL	CL	CC	175	LL	LL	LL	LL	LL
130	CL	CL	LL	CL	LL	176	CL	CC	CC	CC	CL
131	CC	CC	CC	CC	CC	177	CL	LL	CL	LL	CL
132	CL	CL	CL	CL	CL	178	LL	LL	LL	LL	LL
133	CC		CC	CC	CC	179	CL	CL	CL	CL	LL
134	CL	LL	CL	CL	CL	180	CL	CL	CL	CL	CL
135	CL	CL	CL	CL	CL	181	CC	CC	CC	CC	
136	LL	CL	CL	CL	CL	182	CL	CL	CL	CL	CL
137	LL	LL		LL	LL	183	LL	LL	LL	LL	LL
138	CC	CC	CC	CC	CC	184	CL	CL	CL	CL	CL
139	CC	LL	CC	CC	CC	185	CL	CL	CL	CL	CL
140	LL	LL	CL	CL	CL	186	CL	CL	CC	CC	CC
141	CL	LL	CC	CC	CC	187	CC	LL	CC	CC	CC
142	CL	LL	CL	CL	CL	188	LL	LL	LL	LL	LL
143	CC	LL	CC	CC	CC	189	CL	LL	CL	CL	CL
144	CL	LL	CL	CC	CL	190	CC	CC	CC	CC	
145	CL	LL		LL	LL	191	CC	CL	CL	CL	CL
146	CL	CL	CL	CC	CC	192	CL	CL	CL	CL	LL

Colour code: Green with CC indicates homozygous for Columbia for that given marker locus for that given recombinant line, likewise red with LL indicates Landsberg *erecta* and yellow with CL indicates heterozygosity. Lines without any colour code indicate marker product is missing for that recombinant.

2.1.A6 Generation of bigger fine mapping population and molecular marker analysis

As the analyzed 192 recombinant mapping individuals could not result in further narrowing down of the *sud2* mutant locus, in a next step a larger mapping population of 1239 individuals was generated in the similar fashion as the previous 192 lines were generated (from F1 selfed recombinants). Table 4 shows detailed information about the origin of the larger mapping population. DNA was extracted from this mapping population for each individual using an automated extraction method. This DNA served as template for PCR based amplification of markers. The dCAPS markers were designed by using public databases TAIR/TIGR, the SNP markers between Col and *Ler* were converted to dCAPS markers in order to be able to detect polymorphism between Col and *Ler* ecotypes. SNP marker products were directly analyzed on high percentage agarose gel by electrophoresis. The dCAPS markers were first digested with the respective restriction enzyme. These digested products were used for detection of polymorphism on agarose gel. Analysis of several SNP and dCAPS markers on this mapping population suggested a position of the mutant locus between markers T32N15 and *Acl*. The marker analysis is shown in Figure 11. This has reduced candidate area from 7.5 Mb to a 350 Kb region on chromosome III. The *sud2* phenotype mainly co-segregated with the T32N15 marker. Table 6 shows list of markers that were analyzed. The results suggested there was low recombination in this mapped region, it hampered the further mapping process to the candidate level.

	16.23 Mbp	16.26 Mbp	16.3 Mbp	16.3 Mbp	16.35 Mbp	16.36 Mbp	16.4 Mbp	16.5 Mbp	16.66 Mbp	16.75 Mbp	16.8 Mbp	17.2 Mbp
Number	Psi	Xmn	BsaH	Bst	Nde	T32N15	<i>Acl</i>	Sty I	Apa	A1w	BsrD	T6H20
197	CC	CC	CC	CC	CC	CL	CL	CC	CC	CC	CC	CC
267	CL	CL	CL	CL	CL	CL	CL	CL	LL	LL	LL	LL
300	CC	CC	CC	CC	CC	CC	CL	CC	CL	CL	CL	CL
304	CL	CL	CL	CL	CL	CC	CL	CL	CL	CL		
391	CL	CL	CL	CL	CL	CL	CC	CC	CC	CC	CC	CC
516	LL	LL	LL	LL	LL	LL	LL	LL	CL	CL	CL	CL
577	CL	CL	CL	CL	CL	CC	CL	CL	LL	LL		LL
609	CL	CL	CL	CL	CL	CC	CL	CL	CL	CL		CL
686	CL	CL	CL	CL	CL	CL	CC	CC	CC	CC	CL	CC
737	CL	CL	CL	CL	CL	CC	CL	CL	CL	CL		
1179	CL	CL	CL	CL	CL	CC	CL	CL	CL		CL	

Figure 11 Fine mapped region of *sud2* mutant. The fine mapped region of expected position of mutant locus is out lined by orange box. Analyzed markers and their position on chromosome III. Number represents the recombinant line. The colour code Green with CC indicates homozygous for Columbia for that given marker locus for that given recombinant line, likewise red with LL indicates Landsberg *erecta* and yellow with CL indicates heterozygous. Lines without any colour code indicate marker product is missing for that recombinant.

Table 5: The origin of the 1239 individuals of the fine mapping population

Cross	No. of F1 selfed plants	No.F2 generated	No. of F2 upgraded
Mp#5-1-50 x 86Ler#28-1-2	9	54	54+270
Mp#6-1-25 x 86Ler#48-1-1	9	48	48+240
86Ler#48-1-1 x Mp#6-1-37-1	10	36	36+267
86Ler#48-1-1 x Mp#6-1-37-2	9	54	54+270
Total no. of plants	37	192	1239

Table 6: The markers analyzed on mapping population of the 1239 individuals.

Gene Marker	Marker position on chromosome 3	PCR primers for amplification	Enzyme
Psi	16.23	TCG GGA GCA TTA TAC AGT TCA GTT AAA GTA TGC ATA TTT ATT GTG TCT CCT TA	Psi I
Xmn	16.23	ATA ATC TCA TTT AGC CCA CTC GAT TCT TGA TGG TTG CTT CAT CCT GAA GAT ATT	Xmn I
BsaH	16.3	TAC TTT CCG ATG AGA TTA AGA GTT GTT CTC AAA ATA TTG AAT TGC CGA TGG	Bsa HI
Bst	16.3	GGA CTG TGA GGA TAG TAT CAT TCA TTC TAG AAG CTG AAG CGC GAA ATG ATC	Bsp 1191
Nde	16.36	GAC AAC TGT TAT ATT TGG TGC CTT GAT TTA CAA CAG TTA GGC GAA ATC ATA	Nde I
T32N15*	16.36	ATC TGA AAA TCC TTG CGT GAG TTG TGA CGA ATA GTG AAA GGA GAG	
Acl	16.4	TAT CCG TCC GAT GAT CAA TCT CCT GCT AAC GAA AAC AGA GCC GAA AAA AAC	Acl I
LASSty	16.5	TTC CCC ATT TGG GCT CTT GGA GAA CAA TCA TTG AAA ATA AAA CAT GGT CCA A	Sty I
ApaL	16.6	ATT TGC TTG CAT CGG AGT ATG AGA GTG C GTC AAA AAC GTA ACC AAC TTC CCT T	Apa LI
Alw	16.7	CAC AAA AAA CAA TGA AAA TCA GAG ACA TGT CTT ATG ATG AGA CAT GAT T	Alw NI

BsrD	16.86	CTG TAT TTC TCT TCA AGA TCA AGC AAT AAC TGC AGA GTT GAG TGA GAA ACT T	Bsr DI
T6H20*	17.2	CGG CTG AAA CTT GGA AGG GAC AGG AAG AAC GTG TGA TTG TG	

* Sign markers are simple sequence length polymorphism (SSLP) markers, The PCR product wa directly analyzed for the polymorphism. Markers without any special sign are dCAPS, the PCR product was digested with respective enzyme before being analyzed to score polymorphism.

2.1.A7 Library construction of 350 kb genomic sub-region of chromosome III

This analysis of several markers on 1239 recombinant plant lines suggested low recombination events in this fine mapped region of interest. In addition this region showed nearly 35% of sequence repeats. These two reasons did not allow further marker-based analysis to delineate the *sud2* mutant locus on chromosome III. To overcome these limitations, an alternative approach was designed to generate a library of PCR fragments of the region of interest by using *sud2* DNA as a template for amplification. For schematic over view of the strategy see Figure 12.

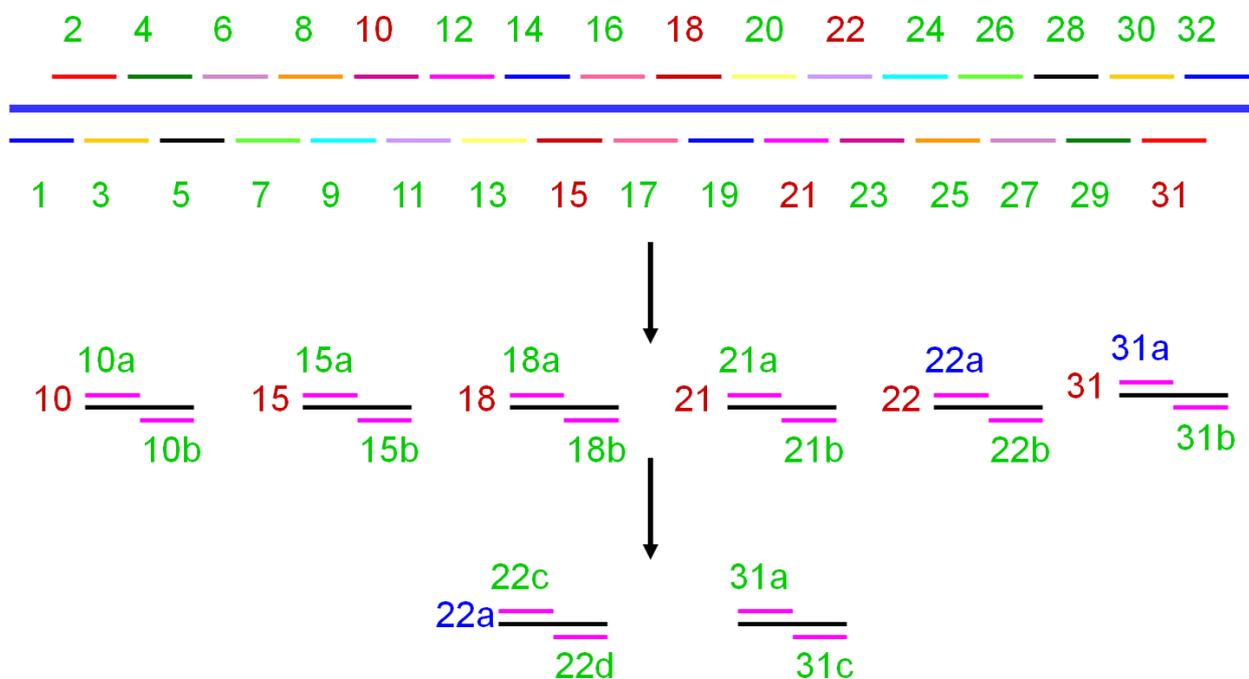


Figure 12 Schematic representation of library of fragments of sub-genomic 350 kb region on chromosome III. The numbers represent the small sub-fragments of 350 kb region that was PCR amplified. Arrows are representing further fragmentation of fragments where a PCR product was not obtained. The information about fragment numbers corresponding size and their overlapping size to next fragment are explained in Table14.

To this end, first the region of 350 kb oligonucleotides were designed for approximately 10 kb fragments with overlapping ends to the next fragment (Table 7). Out of 32 designed fragments, 26 fragments were amplified by PCR. Remaining regions, for which no product was obtained, were further targeted for amplification by using oligonucleotides spanning smaller regions. This process was repeated till all expected fragments were successfully amplified. The final number of amplified products was 41. All these PCR-based amplified products were analyzed via agarose gel electrophoresis. The products that corresponded to the expected size were purified from the agarose gel and quantified. These fragments were pooled in equimolar amounts and used for sequencing.

Table 7 Oligonucleotides used for PCR-based amplification of sub genomic region of interest on chromosome3

Fragment number	primers for amplification	Fragment size	Overlap size
1	GCTAACGAAAACAGAGCCGAAAAAATC CATTACAATGATCGGCGGTGAAGGT	10.9 kb	0.2 kb
2	CGATATCGAGTTTCGTGGAGATGGCTT CTCAAGCTGCAAACACTCGAACACCTT	9.1 kb	0.4 kb
3	GCCATCTACTCTTGACAGTTCCTGTT CTGGCTAACTACGCTCGAAATGTCGTC	10.6 kb	0.3 kb
4	CGGTTACACCTGACCCGTCGACAATT TTGTGACGAATAGTGAAAGGAGAG	13.0 kb	0.2 kb
5	ATCTGAAAATCCTTGCGTGAG ATCAGGCGAGTATTGGGATGACTCCTT	8.9 kb	0.4 kb
6	GTTGGGCTGAAGATCCTCTGGAATCTT TGCATGTCGCCCCATCAAACACACTC	14.0 kb	0.4 kb
7	TGGCTAAGGTACGTCTTGGTGAGCTT GCTGATATTCTACGGCAGCTGTTGGATT	9.5 kb	0.4 kb
8	GGTAACTTGGCGACAGTATTCTTGGTC GGGTTGAATTGTTTAGTAGTGGTGATT	11.7 kb	0.3 kb
9	GTGTGAATTAACCAGCTCGAGTCTTT CGTAAGTCGAATGCAGACCTATCTGCT	8.9 kb	0.3 kb
10a	TGCCTTGGCGAAGTTGTTGCCAAGGTC AGGACGATCTATGCTCATGAGGACACT	6.2 kb	0.7 kb
10b	GGAGAAGGATGATGACGGAATCTCACT TACTTTCCGATGAGATTAAGAGTT	5.3 kb	0.2 kb
11	TGTTCTCAAAAATATTGAATTGCCGATT ATCGGAAGCTATAAACAGCGCCGATT	12.3 kb	0.4 kb
12	GAACCAATGGATAGTGGTTCTAATGTC GGAAGTCGAAGGTCACATACCGGACTT	10.9 kb	0.04 kb

13	AGCAACACCACAGTTACCCTTAGGTCT ATAATCTCATTAGCCCACTCGAT	10.7 kb	0.2 kb
14	GATGGTTCTTGATGGTTGCTTCATCCT TCTTCGTCAGAGCGCTGACCCACCTT	11.3 kb	0.5 kb
15a	TGTCATTCATGAGGTGCGTCCTCAACT CTACTCTGTATAGACCTCACCAGTCACT	5.9 kb	0.6 kb
15b	CCGCCCTTTTCTAAACTGATCCCCT GAAGTCGCAGAGATTTACGATTCTCC	7.0 kb	0.5 kb
16	GACCTATCAATCAAGTAGACAGTGGTT TCCAAGTGCAGACTGCAGTTAGCACT	10.8 kb	0.4 kb
17	GGGAGCTCCACTAACTTGAGCAATCT ACTTGTAGCCTGTTAGTGTTCCTTGTT	11.8 kb	0.7 kb
18a	CCAGTTCGTCGATCCGTCTCAAGAGTT CAGAAGTCGAAGATGGACCCGTCGTT	6.5 kb	0.5 kb
18b	CAACACTAGGTTCTGCACTTGCATTGAT GTGAATGGTGGTGGAGCATATGGACAAT	6.9 kb	1.0 kb
19	GCCTGCATTTTCATGATCTCGGCCAACT* CACATGTATGTCACGTACAGTATGACC	12.0 kb	0.5 kb
20	CATAGTGCGGTACTCATCTAGTAGTT TAGTCTAGGAATCCTGATACCGTACCTT	13.6 kb	0.6 kb
21a	AGCCTACTTGCGGTATAACGGAGGTT CCAAATTGACTGTCTCCAGTACCATC	6.5 kb	0.3 kb
21b	TTGGAAACTCAGCCACAACATCCCTTC AACCTAACGAGCAGTACCTTGTTACCT	6.2 kb	0.5 kb
22a	GAAGCTCCTCGGTAGCGTTCTGTACTT GTTTGAGATCTGCGACATGTATGGATT	1.1 kb	0.1 kb
22b	TGTGTTATCTTCGCTGTGCGGACTT CAACTATGGGTCAGGTTATTCAGGTAT	1.8 kb	0.2 kb
22c	GTACGTACACGGTTGTAATTTGTGTCT GAGATACACGTACACGTGTATGACCTCT	2.0 kb	0.5 kb
22d	TCCAGTCATCTTCAACTTGAGTATCCTT TTAGGTGGATGCTTAGGTGGATAGGCT	6.1 kb	0.4 kb
23	GGGGATCTAATTCAGAGGCTGAGTCCT GGTTGGATCGAAGGCTTGAGGTAAGTT	11.1 kb	0.5 kb
24	TAACGAACCGCAATCTCATGAATCTCT GGGGACGTCGGTTGTATGGCAGAATT	11.5 kb	0.8 kb
25	AAAAGGCGGTCCTGCCTAGAGGCATT GAGATGGACCACCTATTCCAAGTCCTT	9.2 kb	0.5 kb
26	CAGAACATAAGGATCCCAAGACAAGTT CCTTCCATTCATACCCAAGCACCTTTC	16.2 kb	0.3 kb
27	CATACCGTTGCTGTGATTGCTTTCCCT GCACTAAGTTGGAATCTGTGCGGAGACT*	10.0 kb	0.8 kb
28	CACGGAGCAGATGTCCGAAGAACCTT* GGATTACGCTGCTGGATGCTCTGATC	10.8 kb	0.6 kb
29	TTGTAATTCACACCATGCTTCCCAGTT AGAGATTTGTGGTGGCCGTGTCTTCTT	10.7 kb	0.7 kb

30	CTTTAACCGCTTCGCAAACAAACCGCT CGATGCAACAATCGAGTCAATGCGTCC	10.8 kb	0.3 kb
31a	ATTTGGCTTAACCGGATTGGTCATCCT CAAATTGTTAGGAATGTCTTTCCAAAAC	3.5 kb	0.7 kb
31b	GTTACAAAACACTACCCAATTTAGGCTATT* GAATGAGGCTTACACGAGGTGGGATT	3.5 kb	0.6 kb
31c	GACGTTGCAGCATGAGATAGGTCGAT CCCGGATGATTGTTGATGAGATCGTGTT	5.3 kb	0.3 kb
32	GTTTCTTTGAACCCGAAGCTTCTCCATC CAAAAGATATTTAAGCGGAATCTTGCC	11.5 kb	

* sign indicates, this sequence matches more than once to the region of interest, nevertheless the second oligonucleotide sequence of the pair is unique and matches to specific sequence thus rendering PCR product specific for the region of interest.

2.1.A8 Solexa-based sequencing of fine mapped region of *sud2*

Successfully amplified and subsequently purified fragments were quantified using nano-drop nucleic acids measuring machine. 5 µg of an equimolar pool of all the fragments was subjected to Solexa-based sequencing by using services provided by the company GATC (Konstanz, Germany) in order to do the sequencing. The readout of single run was 10,682,567 short reads. The resulting final sequence average length was 40 bp per read.

2.1.A9 Sequence alignment

To identify polymorphic candidates present in *sud2* sequence region of interest, an alignment of Solexa sequencing obtained reads with wild type Columbia sequence was carried out in a collaboration work with Fritz Sedlazeck from the group of Dr. Arndt von Haeseler at Center for Integrative Bioinformatics Vienna, University of Vienna, Austria. The mutant sequence reads were aligned to the reference sequence of Col-0 which was obtained from NCBI database. The reference sequence was consisting of chromosome III region of interest covering genes from At3g44400 to At3g44900. To align the mutant sequence reads to the reference sequence, the program NextGenMap (Sedlazeck et al., submitted) was used. 92% of the total reads were aligned, and 96% of aligned reads mapped to a unique position within the region of 1350 kb. The sequence reads were subjected to pairwise alignment and the ones which exhibited identity of above 80% were considered for subsequent analysis. The identity was calculated by applying the following formula.

$$\text{Identity} = \frac{\text{Number of matches}}{\text{Length of the alignment}}$$

These reads were categorized into two classes depending on identity. The category 1 comprises reads that showed at least 80% or 97% (C-80, C-97) identity to the reference sequence. The category 2 includes the reads that showed 100% (C-100) identity to the reference sequence. The reads that were not supporting the category 1 and 2 requirement were considered as not informative to identify candidates as they might have been generated from unknown DNA contamination. Figure 13 shows the number of reads that were mapped based on identity to the reference sequence. Large number of reads can be mapped to the reference when the identity threshold is set to 70% or lower. As the required threshold for the identity is increased, the number of reads that can be mapped to reference sequence starts to decrease. When the identity limit is set to 97%, which corresponds to single mismatch per read, the graph shows a sharp drop reaching lowest point of the curve. This served as basis to identify reads with single mismatches, which in turn supports to identify candidates differing by 1 base from the Col-0 WT as expected from EMS mutagenesis.

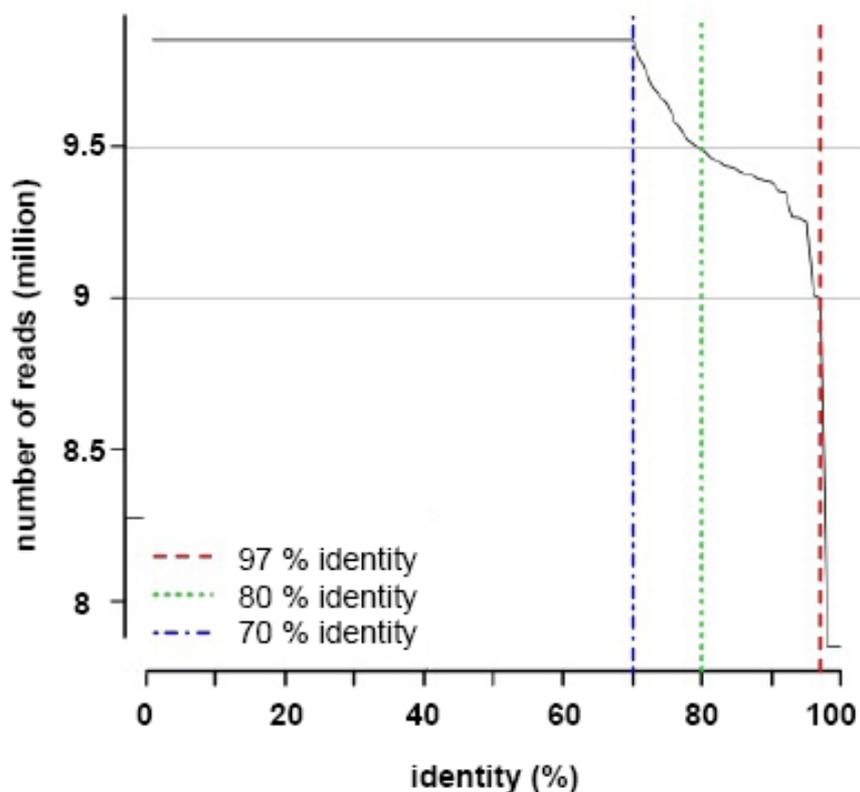


Figure 13 Graph represents relationship between the number of reads that can map depending on the value of identity. 70% identity serves as minimum and 97% as maximum threshold to identify mismatches in the mutant sequence reads. The number of reads that aligned showed certain percentage of identity to the reference sequence.

2.1.A10 Identification of SNP candidates

In the next step to identify SNP candidates, the relationship between the number of reads that can map depending on the value of identity was investigated for each category data set. To this end, the total number of reads that align to each base of reference region of interest was determined. Again they were categorized into C-80 and C-100. C-100 consists of 74% of the reads and 95% of them mapped to a unique position in the reference. C-80 comprised 94% of the reads and from these, 91% mapped to a specific position in the reference region of interest. Thus, for each base position in the reference sequence, the total number of C-80 and C-100 reads was obtained. To overcome varying coverage at different regions of the sequence, the position-wise C80/C-100 ratio was computed. This ratio value provides more robustness than absolute values for C-100 and C-80. The average ratio C-80/C-100 equaled 1.329 (median: 1.187). When most of the reads at a given position carry the reference base, the ratio will be close to the average value. If a particular position differs from the reference base on both chromosomes, the number of reads within the tolerated 20% mismatch level would not differ significantly from local neighbors. However, only few reads with perfect match can be expected at such position, as generation of reads with reference sequence at mutant positions requires actually a sequencing error or PCR mistake, thus leading to a comparatively high C-80/C-100 ratio. Therefore, the next step was focused on positions with a C-80/C-100 ratio > 50 . This corresponds to the 10^{-6} quantile of the empirical distribution of the C-80/C-100 ratio, which would indicate a significant difference in coverage for both categories (C-80 and C-100) at this position. In the 350 kb region of interest, 15 such spots were found, distributed over nine regions. Figure 14 displays the change in coverage for the 15 potential mutation spots. Five of the potential spots (2-6) showed a V-shaped drop in coverage for the C-100 category. This was exactly the change in coverage one would expect, if one nucleotide of mutant reads deviates from the reference sequence. The remaining 10 potential spots displayed a somewhat irregular behavior that cannot be explained by the presence of a single point mutation. 8 of the spots form two clusters with respect to their genomic positions, that is, spots 7, 8 and spots 10-15.

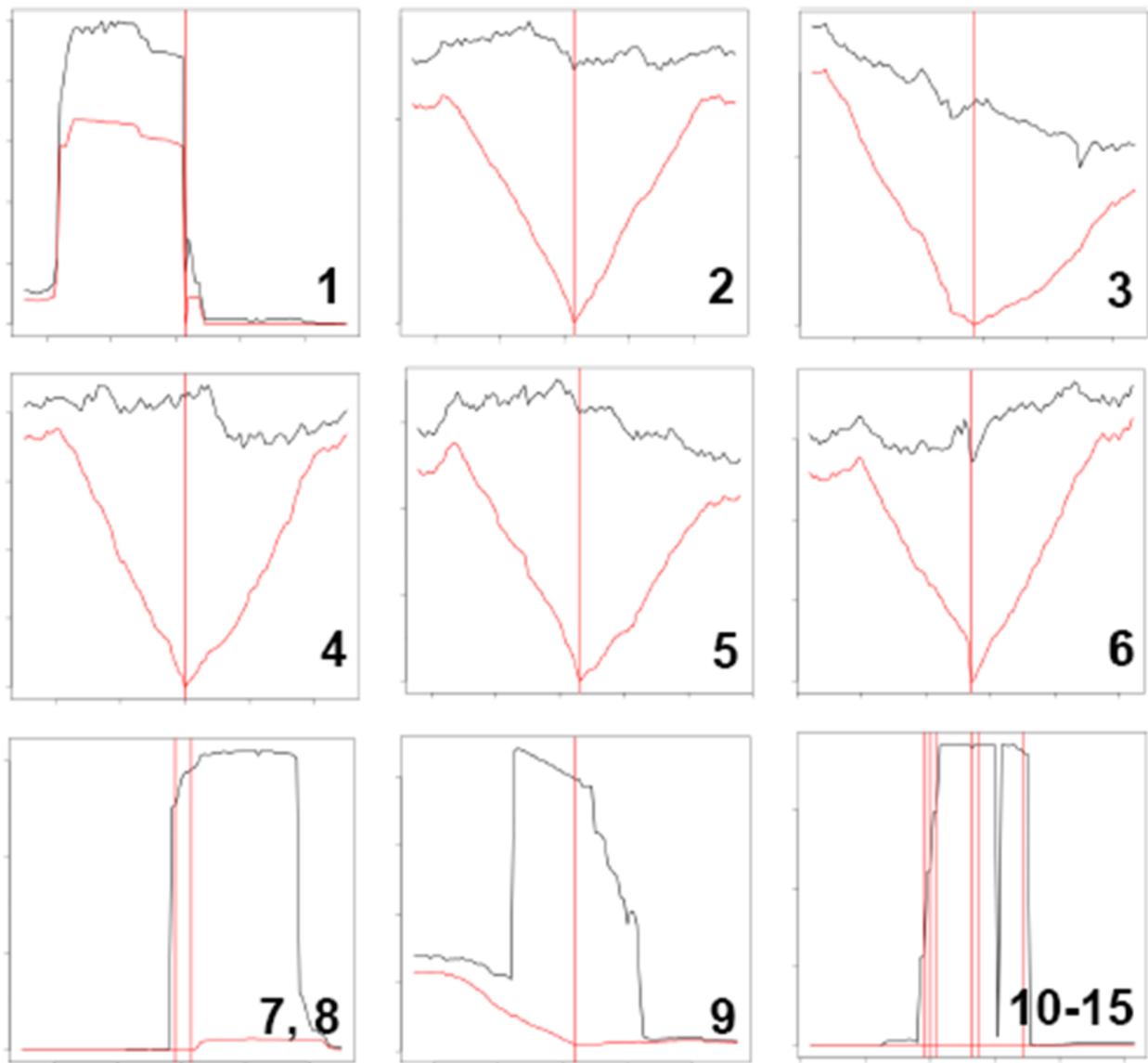


Figure 14 - C-80 and C-100 read frequencies at predicted mutant positions The number of reads was plotted on the y-axis for the C-80 (black lines) and the C-100 (red lines) set, respectively. Plots show the vicinity of predicted polymorphisms (x-axis). Scales: y-axis, marks indicate 200 reads for panels 1, 4, 5, 6, 7/8 and 10-15, and 1000 reads for panels 2, 3 and 9. Marks on the x-axis indicate 20 base pair distances. Positions of predicted polymorphisms are delineated by a vertical red line.

Table 8: Identified polymorphic regions and their coordinating position on chromosome III.

polymorphism number	Sequence context ^a	Frequency C-80 ^b				Coordination on chromosome 3 ^c
		A	G	C	T	
1	AGCTTCAGGGTTT AGCTTCAGGGTTT	14	0	151	6	16 051 815
2	TGGAAAAGAGGAAA TGGAAAAGGAAA	1233	0	4	0	16 161 395
3	TTCTGTTTTCTTT TTCTGTATTCTTT	1150	7	24	131	16 261 894
4	ATGGATTTTTCTT ATGGATCTTTCTT	1	419	1	1	16 308 852
5	TTCAGAAGCTTGA TTCAGACGCTTGA	0	621	1	1	16 339 238
6	CTACTAA TCGCCA CTACTAGCTTCGCCA	0	0	430	116	16 362 588
7	GAGGTCTATATTG GAGGTCTATATTG	2	0	505	0	16 373 512
8	CTATATTGAGTTT CGATATTGAGTTT	2	579	0	3	16 373 517
9	CGGCGGTGAAGGT CGGCGGTGAAGGT	25	65	3422	454	16 379 718
10	ATGCTAACATGTT ATGCTAACATGTT	3	1	1	221	16 380 572
11	GCTAACATGTTCT GCTAACATGTTCT	5	11	3	429	16 380 574
12	TAACATGTTCTTT TAACATGTTCTTT	0	11	2	581	16 380 576
13	TTTCTTTCTTTTTT TTTCTTTCTTTTTT	0	17	0	742	16 380 587

14	CTTTCTTTTTTCT CTTTCTTTTTTCT	0 – 0 – 742 – 25	16 380 589
15	TTCTCAAAGGGT TTCTCAAAGGGT	2 – 742 – 0 – 0	16 380 603

^a, The reference sequence context is written at the top, the sequence of the EMS treated plant line as determined by confirmatory (Sanger) sequencing is written below. The nucleotides in question are shown in bold.

^b, Number of reads in the C-80 set that contain a certain base at the position of interest.

^c, Coordinates are base positions on *Arabidopsis* chromosome 3, sequence version of September 11th, 2009 (www.ncbi.nlm.nih.gov).

2.1.A11 Validation of identified polymorphisms

In order to confirm the Bioinformatics based identified mutations, short fragments that cover the region of potential mutation were PCR amplified by using the DNA templates from *sud2* and progenitor line *RV86-5*. These fragments were subjected to Sanger sequencing. Table 8 shows the sequence context of the 15 positions and the frequency of each of the four bases for the C-80 reads at the putative mutant position, which allowed a hypothesis regarding the nature of the sequence change. The last column of Table 8 lists the position of the presumed mutation in the form of the number of the base on reference chromosome 3. The spots that showed clear V-shaped curve in the data set of C-100 (spots 2, 3, 5 and 6 in Fig 14, Table 8) were identified as true deviations from the reference sequence. The mutation at spot 2 showed polymorphism from G to A, whereas in spot 3 a substitution of T with A occurred. These changes were present in the *sud2* but not in the original mutant line *RV86-5* (Fig 14, Table 8). In the case of spots 5 and 6 the differing base pair already existed in the progenitor line *RV86-5*, therefore they might be pre existing SNPs and may not have altered the phenotype. In their polymorphic site 6, an A was replaced with GCT in the mutant sequence.

Table 9: Confirmation and annotation of identified polymorphisms

Polymorphism number	Annotation	Confirmation result
1	between annotated ORFs At3g44400 and At3g44410	not confirmed
2	between annotated ORFs At3g44580 and At3g44590	confirmed
3	within intron of gene At3g44713	confirmed

4	error in reference sequence of 2008 (now corrected)	confirmed ^a
5	within ORF (gypsy retroelement) At3g44796	confirmed ^b
6	within ORF At3g44820	confirmed ^{b, c}
7, 8	within ORF At3g44840	not confirmed
9 to 15	within ORF At3g44860	not confirmed

Genomic regions that showed polymorphism in alignment between *sud2* and reference sequence were PCR amplified and isolated from mutant and progenitor and subjected to conventional sequencing, the final comparison outcome are listed.

^a, sequence difference was confirmed, but the reference sequence (version of 2008 was used in this work) has since been corrected to the observed nucleotide.

^b, sequence was confirmed, but the differing nucleotide is present both in the mutated line and in its progenitor, suggesting that the nucleotide change cannot be the mutation causing phenotypic differences between progenitor and mutant lines.

^c, Sequence change in 6 is actually an A to GCT mutation, i.e. is a frameshift mutation.

The changes 1 and 7 to 15 were not confirmed. This is in agreement with the observation that they did not exhibit significant polymorphic curve, which indicated unexplainable deviation from the reference sequence. The spot 4 (Fig 14, Table 8) was actually a sequencing mistake in the sequence release used for comparison, which was reset in updated later version of *Arabidopsis* genome releases (Version of September 11th, 2009; NCBI). Table 9 shows summarized results after validation. These validated candidates would be of interest to check whether the identified modification can rescue the phenotype of progenitor line *RV86-5*. Table 10 listed short description of identified candidates. Some of these identified ones were further crossed into *RV86-5* for experimental validation.

Table 10: Identified suppressor screen candidates

Polymorphism number	Gene locus	Candidates short Description
1	AT3G44400	Disease resistance protein (TIR-NBS-LRR class) family; function: transmembrane receptor activity, nucleoside-triphosphate
1	AT3G44410	Pseudogene, disease resistance protein, putative, similar to disease resistance protein RPP1-WsB (<i>Arabidopsis thaliana</i>)
2	AT3G44580	<i>Arabidopsis thaliana</i> protein match is: <i>Arabidopsis</i> retrotransposon ORF-1 protein biological process unknown, molecular function unknown
2	AT3G44590	60S acidic ribosomal protein family; functions in: structural constituent of ribosome; involved in: translational elongation
3	AT3G44713	unknown protein; <i>Arabidopsis thaliana</i> protein match is: unknown protein. biological process unknown, cellular component unknown, molecular function unknown
5	AT3G44796	transposable element gene; gypsy-like retrotransposon family,
6	AT3G44820	Phototropic-responsive NPH3 family protein; functions in: signal transducer activity; involved in: response to light stimulus cellular component unknown, response to light stimulus, signal transducer activity, signal transduction
7 and 8	AT3G44840	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein; functions in : S-adenosylmethionine-dependent methyl biological process unknown, cellular component unknown, methylation, methyltransferase activity
9 to 15	AT3G44860	Encodes a farnesoic acid carboxyl-O-methyltransferase S-adenosylmethionine-dependent methyltransferase activity, biological process unknown, cellular component unknown, farnesoic acid O-methyltransferase activity, methylation, seedling growth

2.1.B1 Differential gene expression of *sud2* and *RV86-5*

As the mapping results showed no mutation in the open reading frame (ORF) of confirmed candidates. The mutations in non coding region can affect the expression levels of genes, to deduce if such expressional changes are prevailing in the mutant background of *sud2* in comparison to *RV86-5*, a microarray-based differential expression profiling was performed. Another noteworthy point to mention is that as the phenotypic changes observed in *RV86-5* come from expression of *ubK48R*, it is also very interesting to analyze what kind of transcripts are affected upon transgene induction. Towards this end, to identify differentially expressed genes between *sud2* and *RV86-5* lines, RNA samples from two week old seedlings 24h Dexamethasone induced and un-induced were extracted. The quality of these RNA

samples was examined. RNA quality was examined by microcapillary electrophoresis using the Agilent Bioanalyser. The Figure 15a and 15b depicts the qualities of RNA. These samples are clearly of high quality and suitable for further analysis. For every genetic background that needed to be tested for transcriptional differential expression, six biological replicates were generated. Information regarding biological samples and their induced or un-induced condition is provided in table 11.

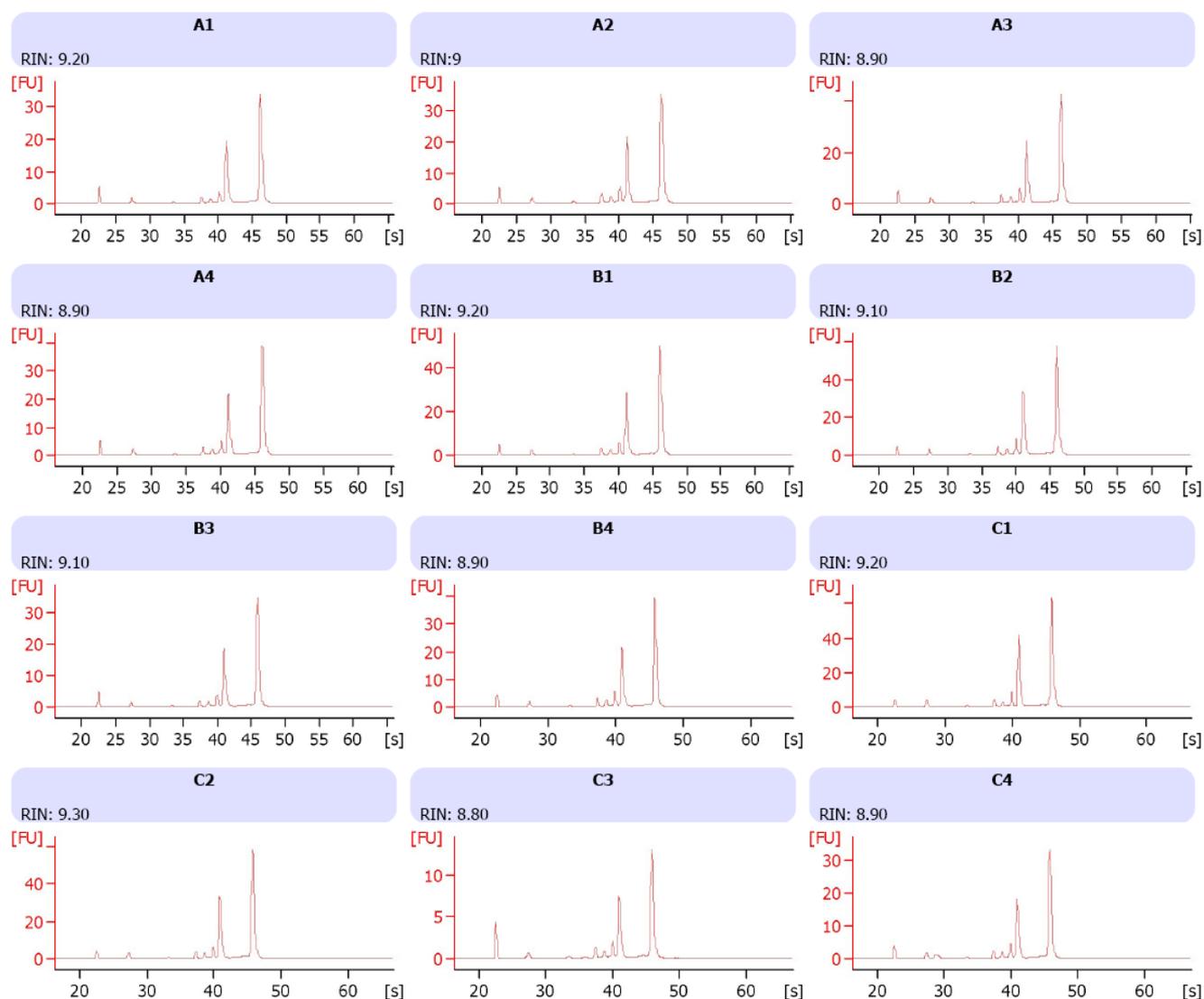


Figure 15a RNA quality. Every sample shows peak of 18S (at 42 S) and 28S (at 47 S).

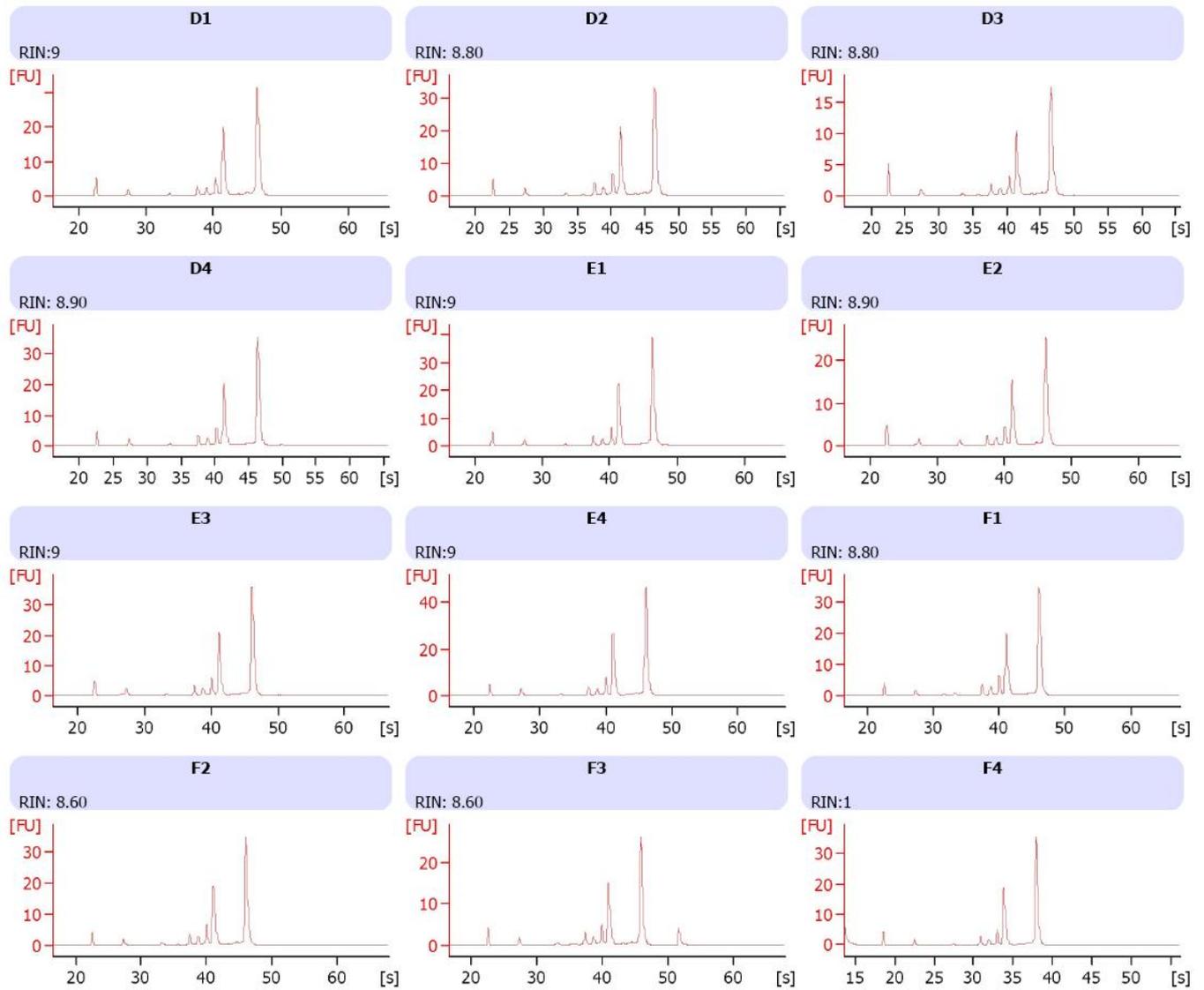


Figure 15b RNA quality. Every sample shows peak of 18S (at 42 S) and 28S (at 47 S).

Table 11 RNA Samples subjected for microarray ATH1 chip hybridization

Sample code	Source of the sample	Induced/uninduced
A1	<i>sud2</i> X RV86.5# 4-16-5-4-2-5	uninduced
A2	<i>sud2</i> X RV86.5# 4-16-5-4-2-5	uninduced
A3	<i>sud2</i> X RV86.5# 4-16-5-4-2-5	Induced
A4	<i>sud2</i> X RV86.5# 4-16-5-4-2-5	Induced
B1	<i>sud2</i> X RV86.5# 4-20-2-2-5	uninduced
B2	<i>sud2</i> X RV86.5# 4-20-2-2-5	uninduced

B3	<i>sud2</i> X RV86.5# 4-20-2-2-5	induced
B4	<i>sud2</i> X RV86.5# 4-20-2-2-5	induced
C1	<i>sud2</i> X RV86.5# 4-20-1-1-5	uninduced
C2	<i>sud2</i> X RV86.5# 4-20-1-1-5	uninduced
C3	<i>sud2</i> X RV86.5# 4-20-1-1-5	induced
C4	<i>sud2</i> X RV86.5# 4-20-1-1-5	induced
D1	RV86.5 # 1-2	uninduced
D2	RV86.5 # 1-2	uninduced
D3	RV86.5 # 1-2	induced
D4	RV86.5 # 1-2	induced
E1	RV86.5 # 2-4	uninduced
E2	RV86.5 # 2-4	uninduced
E3	RV86.5 # 2-4	induced
E4	RV86.5 # 2-4	induced
F1	RV86.5 # 5-1	uninduced
F2	RV86.5 # 5-1	uninduced
F3	RV86.5 # 5-1	induced
F4	RV86.5 # 5-1	induced

Quality checked RNA samples from RV86.5 and *sud2*xRV86.5 were analysed by hybridization to the probes present on ATH1 expression chip from Affymetrix in collaboration with Dr. Bruno Huettel at Max Planck Institute for Plant Breeding Research, Cologne, Germany. This chip consists of 21539 probes. A set of arrays with identical biological material is represented as a class. Thus 4 different classes were to be compared. Expression values of 21539 probe sets were generated by RMA with the affymetrix package. The results files were generated as CEL files and they were analyzed by using Bioconductor/R programs LIMMA and RANK product. To draw a biologically meaningful conclusion with respect to mutant phenotype and the responsible candidate, differentially expressed genes between treated mutant line (*sud2*) and treated progenitor line (RV86-5) would be of great interest. When these two classes were compared, the results showed 38

genes were altered in their expression. All candidates fulfill percentage of false positive predictions (pfp) threshold of 0.05%. A threshold for FDR is commonly agreed at 0.05. The higher expression of genes in the progenitor line might have caused the cell death phenotype. Such candidates were of interest for further investigation. Table 12 shows a list of selected candidates that show a higher transcript level in the progenitor lines in comparison to the mutant line when induced. Appendix 3 shows a pre-selected list of upregulated genes in *RV86-5* compared to *sud2*, both in induced conditions.

Table 12: Identified probable candidates for cell death in *RV86-5*. Fold change is the gene expression level ratio as calculated from microarray data analysis between *RV86-5* and *sud2* lines under induced condition.

Number	Mutant name	TAIR Locus	Fold change	Candidate short Description
1	<i>cand1</i>	AT5G57190	1.8755	Encodes the minor form of the two non-mitochondrial phosphatidylserine decarboxylase. Gene expression is low. Functions in N-terminal protein myristoylation, metabolic process, phosphatidylserine decarboxylase activity, phospholipid biosynthetic process, plant-type vacuole membrane.
2	<i>cand2</i>	AT4G31020	1.895	alpha/beta-Hydrolases superfamily protein; Functions in: N-terminal protein myristoylation, hydrolase activity, metabolic process
3	<i>cand3</i>	AT3G23450	1.894	Biological process and molecular function unknown
4	<i>cand4</i>	AT2G11810	2.095	MGD3 is the major enzyme for galactolipid metabolism during phosphate starvation. 1, 1,2-diacylglycerol 3-beta-galactosyltransferase activity, 2-diacylglycerol 3-beta-galactosyltransferase activity, cellular response to phosphate starvation, chloroplast outer membrane, fatty acid metabolic process, galactolipid biosynthetic process, galactolipid metabolic process, metabolic process
5	<i>cand5</i>	AT5G44580	1.714	Biological process and molecular function unknown
6	<i>cand6</i>	AT3G29970	2.2854	Biological process and molecular function unknown
7	<i>cand7</i>	AT5G39520	1.701	Biological process and molecular function unknown
8	<i>cand8</i>	AT5G20790	2.350	Biological process and molecular function unknown
9	<i>cand9</i>	AT1G19330	1.974	Biological process and molecular function unknown
10	<i>cand10</i>	AT1G62290	1.801	Saposin-like aspartyl protease family protein. Aspartic-type endopeptidase activity, lipid metabolic process, proteolysis, seed, seedling growth, vacuole.

2.1.B2 Experimental validation of microarray-based identified candidates

Mutants in the differentially expressed candidates from microarray analysis were obtained from stock centers. To elucidate if a mutation in these genes could rescue lethal phenotype of progenitor line (*RV86-5*), these mutants (*cand 1 to 10*) were crossed into progenitor line *RV-86-5*. F2 generations were examined for their survival on MS selection media supplemented with 0.7 μ M Dex. They were examined after two weeks. The following Figure 16a, b and c shows obtained result.

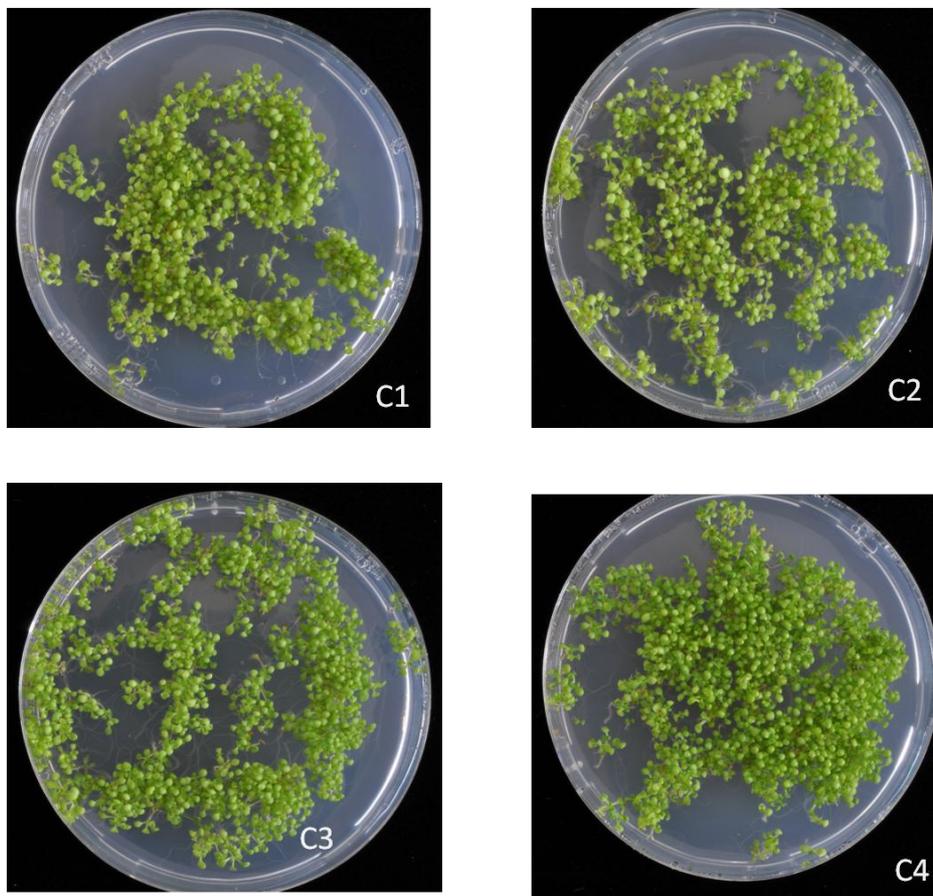


Figure 16a Rescue of lethal phenotype by identified candidate mutants. C1, C2, C3 and C4 indicate candidate number corresponding to the table 12. *RV86-5* was crossed into Candidate mutant lines; two-week old F2 generation seedlings were examined for survival on MS medium with Dex. All lines could alter the lethal phenotype of progenitor line.

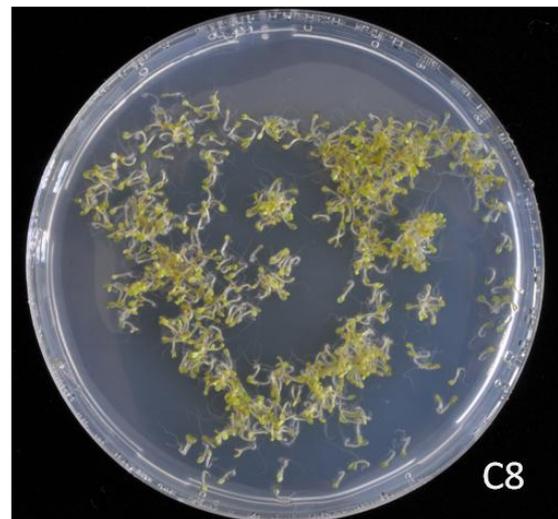
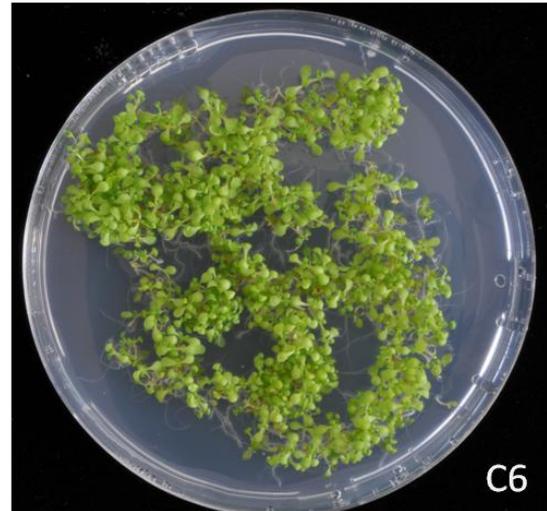
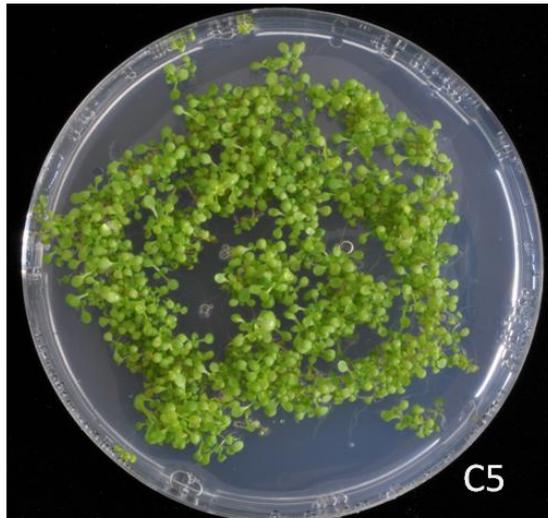


Figure 16b Rescue of lethal phenotype by identified candidates. C5, C6,C7 and C8 indicate candidate number corresponding to the table 12. *RV86-5* was crossed into Candidate mutant lines; two-week old F2 generation seedlings were examined for survival on MS medium with Dex. All lines except C8 could alter the lethal phenotype of progenitor line.

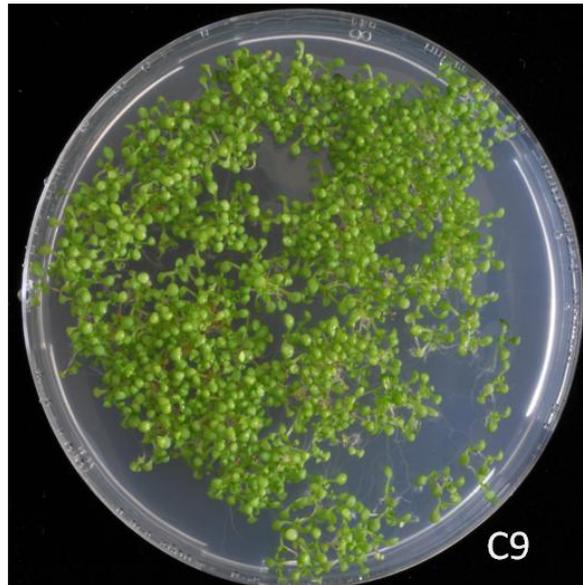


Figure 16c Rescue of lethal phenotype by identified candidates. C9 indicate candidate number corresponding to the table 12. *RV86-5* was crossed into Candidate mutant line; two-week old F2 generation seedlings were examined for survival on MS medium with Dex. Examined candidate 9 could alter the lethal phenotype of progenitor line.

All the tested candidate line could alter the cell death phenotype in the progenitor background. These results illustrate *cand1* to *cand 9* except 8 are negative regulators of cell death. *cand10* F2 generation is now ready to be tested. In case of candidate 8 either the crossing was not successful or the *cand 8* simply has no important role, in relation to the phenotype of progenitor line. The cross needs to be repeated before final conclusions. *cand 8* can be considered as negative control in this case. All the positive candidates further needed to be confirmed for the full transgene expression. One more layer of experiment with *sud2* as positive control and *RV86-5* as negative control and possibly with different levels of selection components in selection media needed to be tested. Taken together the preliminary results suggest the identified candidates do have a biological role in cell death, but more molecular details needed to be examined.

2.2 Results part 2 – N-end rule pathway

2.2.A1 Generation of reporter lines expressing test substrates

The N-end rule pathway targets substrates that are carrying destabilizing residue at the N-termini for degradation by the 26S proteasome. To analyze possible E3-ligase function of components of the plant N-end rule pathway, inducible reporter lines that harbor ubiquitin fusion protein constructs were generated. For this purpose open reading frames (ORF) consisting of a DHFR-HA-UB-X-lac-3HA-GUS fragment were inserted into a plant binary vector pER8 (constructs called pER-X-GUS). X-denotes single testable amino acid, in this case Cys (C), Asp (D), Glu (E), Phe (F), Leu (L), Met (M), Asn (N), Gln (Q), and Arg (R). The expression of these constructs was kept under a β -estradiol inducible promoter. These constructs were further used for agrobacterium-mediated transformation of Columbia wild type plants. When the expression of fusion proteins is chemically induced, they were transcribed and translated as a single entity. The ubiquitin processing enzymes cleaves it into two proteins, dihydroxyfolate reductase (DHFR) and β -glucuronidase (GUS) (Fig 17).

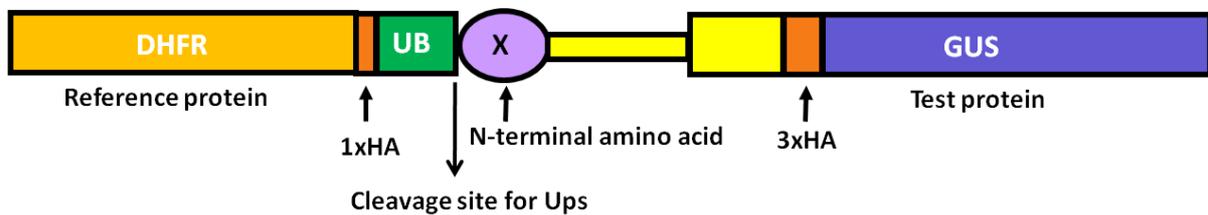


Figure 17 Ubiquitin fusion proteins. ORF of the construct showing GUS (coloured in violet to the right side of UB coloured in green) and DHFR reference protein (coloured in light orange to the left side of UB) containing a single HA tag (coloured in dark orange). X-denotes the aminoacid (coloured in light purple) present at the N-terminus of the GUS test protein which containing triple HA tag (coloured in dark orange). The flexible spacer is present between X and triple HA tag (coloured in yellow). Cleavage site for translated product by ubiquitin processing enzyme is present at the C- terminus of the ubiquitin.

The DHFR consist of single HA tag and is a stable protein, serves as reference protein. In contrast, GUS consists of 3xHA tag and specific N-terminal residues, is either stable or unstable depending on type of N-terminal residue, and serves as test protein.

The flexible spacer between N-terminal residue and GUS protein helps GUS protein to expose residue X in the folded GUS protein. If the tested GUS protein is stable in the tested plant line background depending on N-terminus residue, that plant line gives positive result in a GUS assay and if not, the GUS assay is negative. These GUS assay results are easily

visualized by eye for preliminary readout of the experiment. The stability of test protein versus reference protein can also be quantified by Western blotting.

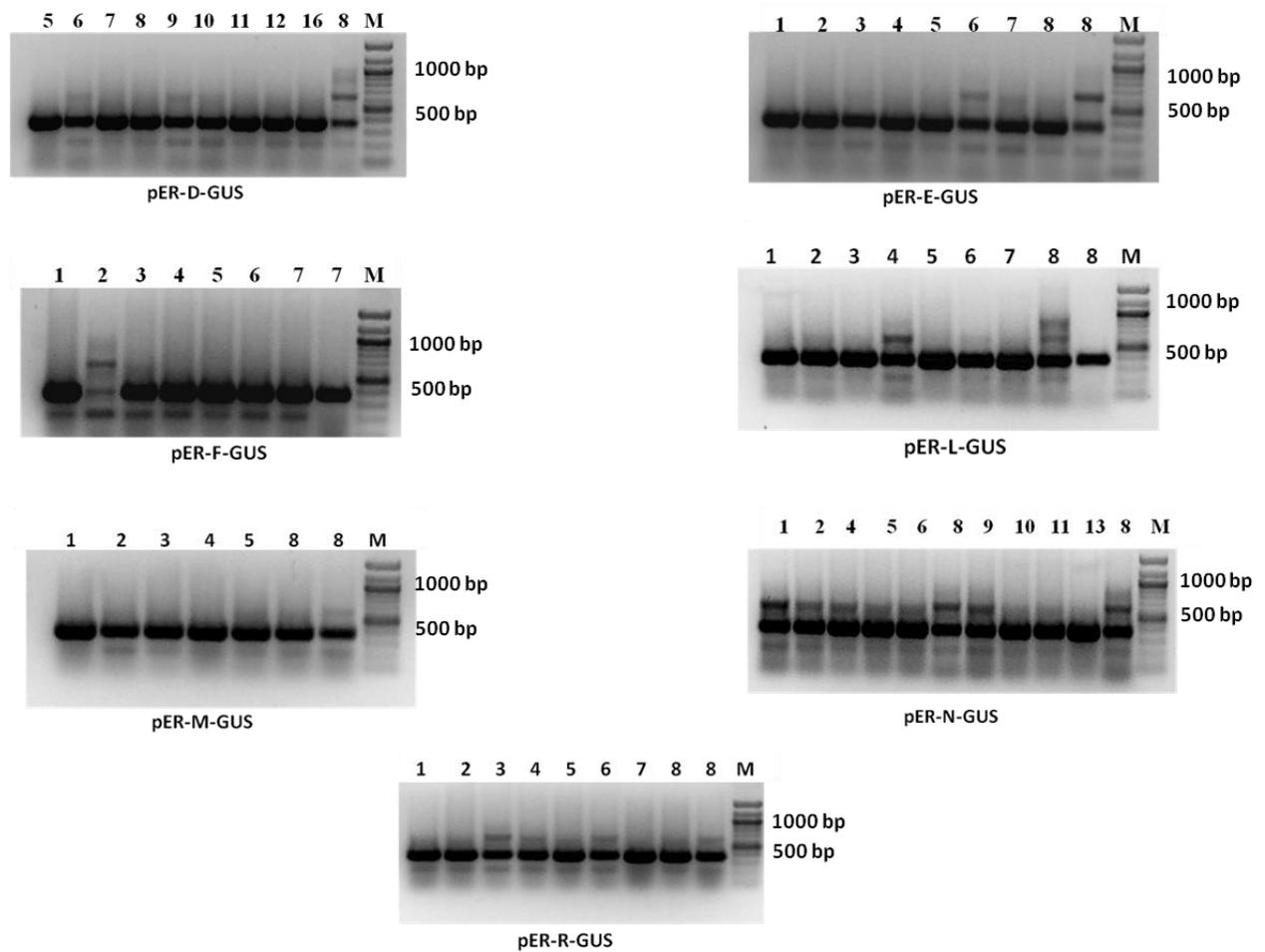


Figure 18 RT-PCR result of transgenic lines expressing test protein. Agarose gelelectrophoretic separation of RT-PCR's cDNA products of fusion protein expressing lines. For cDNA synthesis, GUS specific primers were used, and expected size of the product can be seen at 350 kb marker size. Number on top of every figure represents number of the transgenic line, the respective test protein name is written at the bottom. The letter D, E, F, L, M, N and R represents the amino acid present at the N-terminus of test protein GUS. The line present next to the marker line is an uninduced line and shows a weaker band.

Positive *Arabidopsis* transformants with desired test protein constructs were selected by analyzing their resistance on Hyg containing solid MS medium. Selected plant lines were grown on soil and selfed to obtain the next generations. The seeds of selected F2 lines were germinated on solid MS medium supplemented with Hyg and β -estradiol, in order to induce the expression of the transgene. The F2 lines were further confirmed by RT-PCR for the presence and correctly induced expression of the transgene. Results of RT-PCR are shown in Figure 18. The confirmed lines were further used either directly for experiments or crossed to

mutants in putative or known N-end rule pathway components to analyze their enzymatic function. The functional analysis was performed by testing stability of test protein either by GUS assay or biochemical analysis by Western blotting.

2.2.A2 Seed scale-up and EMS-mutagenesis of pER-L-GUS expressing lines

To perform EMS based chemical mutagenesis on L-GUS expressing reporter lines; 300 to 400 seeds from two selected pER-L-GUS F2 lines were germinated. The resulting lines were grown to the mature silique stage and seeds were collected and pooled to further use them for EMS mutagenesis.

In order to identify the E3 ligase responsible for degradation of the L-GUS test substrate, EMS- mutagenesis was performed. From the above mentioned L-GUS progenitor lines, 0.7 g of seeds were sterilized and subjected to EMS mutagenesis as mentioned in the Methods section. These EMS mutagenized seeds were sub-merged in 0.1 % agar for equal distribution of seeds onto soil. Two to three weeks after transferring seeds to soil, seedlings were examined for appearance of discolored sectors on leaves. This is a sign of mutation in photosynthetic components, and serves as a primary indication of successful mutagenesis. Seeds from all lines were grouped into nearly 50 pools; each pool consisted of seeds corresponding to about 150 to 200 individual plant lines. These pools were used for further screening.

2.2.A3 pER-L-GUS EMS-mutant screen by live tissue GUS assay

A GUS assay was used to find mutants, as the ubiquitin fusion transgene expresses GUS with Leu at the N-terminus. This L-GUS test protein will be stabilized in mutant lines that show dysfunction in destabilization of substrates with Leu at their N-termini, and such mutant lines are GUS positive. In order to identify the un-known E3-ligase that targets test substrate with Leu at the N-termini, live tissue GUS assay was performed. To drive the expression of the transgene and to identify mutant candidate lines, plants of the F2 generation of EMS mutagenized seed pools were germinated on selection medium. A total of 29 pools were examined. From each pool, 1500 to 2000 seeds were germinated on solid MS medium containing Hyg and β -estradiol (Fig 19, 20). 2-to3-three-week old seedlings were subjected to the live tissue GUS assay, as mentioned in the Methods section. After 24 h, mutants that showed positive GUS activity (blue colour) in roots were scored as putative mutant line of

interest as one can expect that only the mutant E3-ligase responsible for destabilization of L-GUS will show GUS assay positive result (Fig 19, 20).

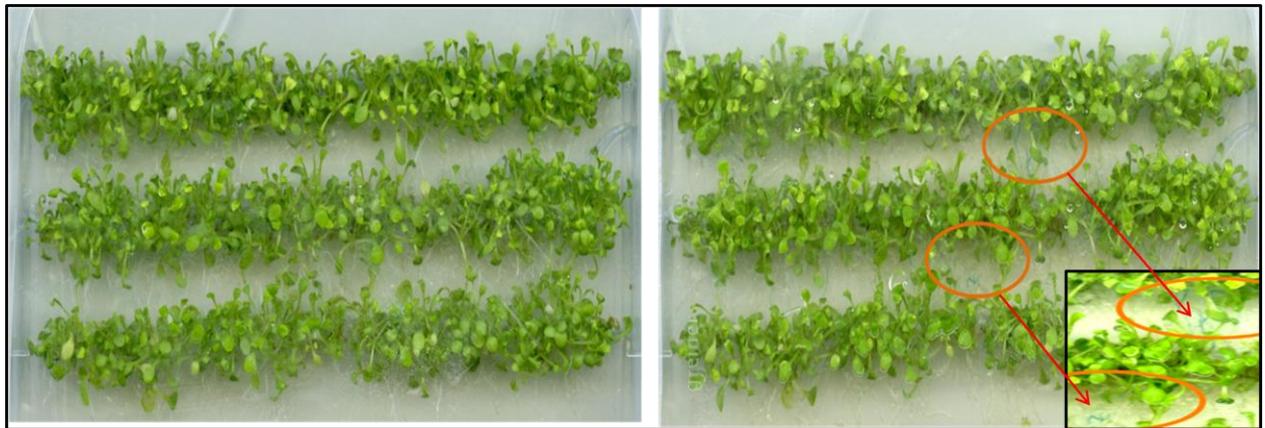


Figure 19 2-to 3- week-old seedlings from one of the 29 pools examined. The left panel shows seedlings before live tissue GUS assay, right panel shows seedlings after 24 h live tissue GUS assay. Circled seedlings are identified as putative mutants in an E3 ligase enzyme. Inlay shows close-up of the identified candidate mutant plants.

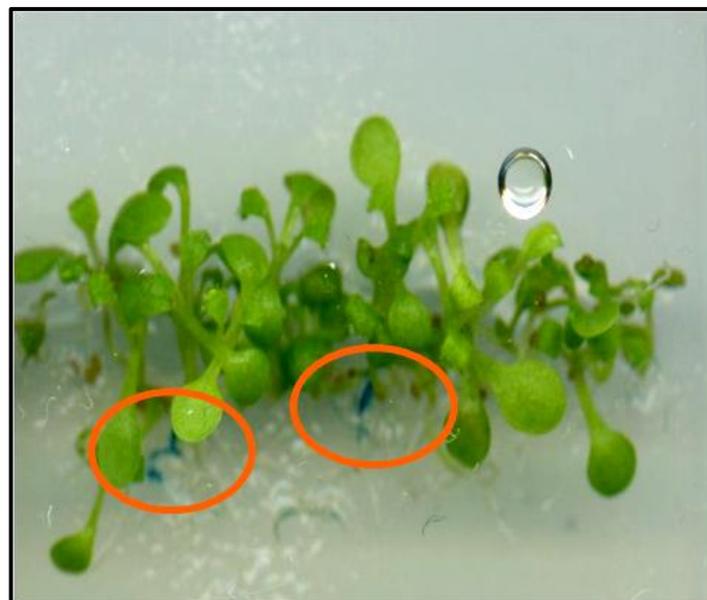


Figure 20 2-to 3- week-old seedlings from one of the 29 pools examined. After live tissue GUS assay, circled seedlings are identified as putative mutants in an E3 ligase enzyme that stabilize test substrate L-GUS.

From examined 29 pools, initially in the first screen 35 lines were selected as promising lines to identify mutants in E3-ligases of the N-end rule pathway. These selected lines were transferred onto soil and seeds were collected. The next generation was retested by the live tissue GUS assay. This was performed in order to select the reliable lines that show stable

results. There is a possibility that in some mutant lines EMS mutation may not be stable and in that case the lines that showed positive GUS assay results in first screen may not show reproducible result in second round of screen. Figure 21 and 22 displays some retested results.

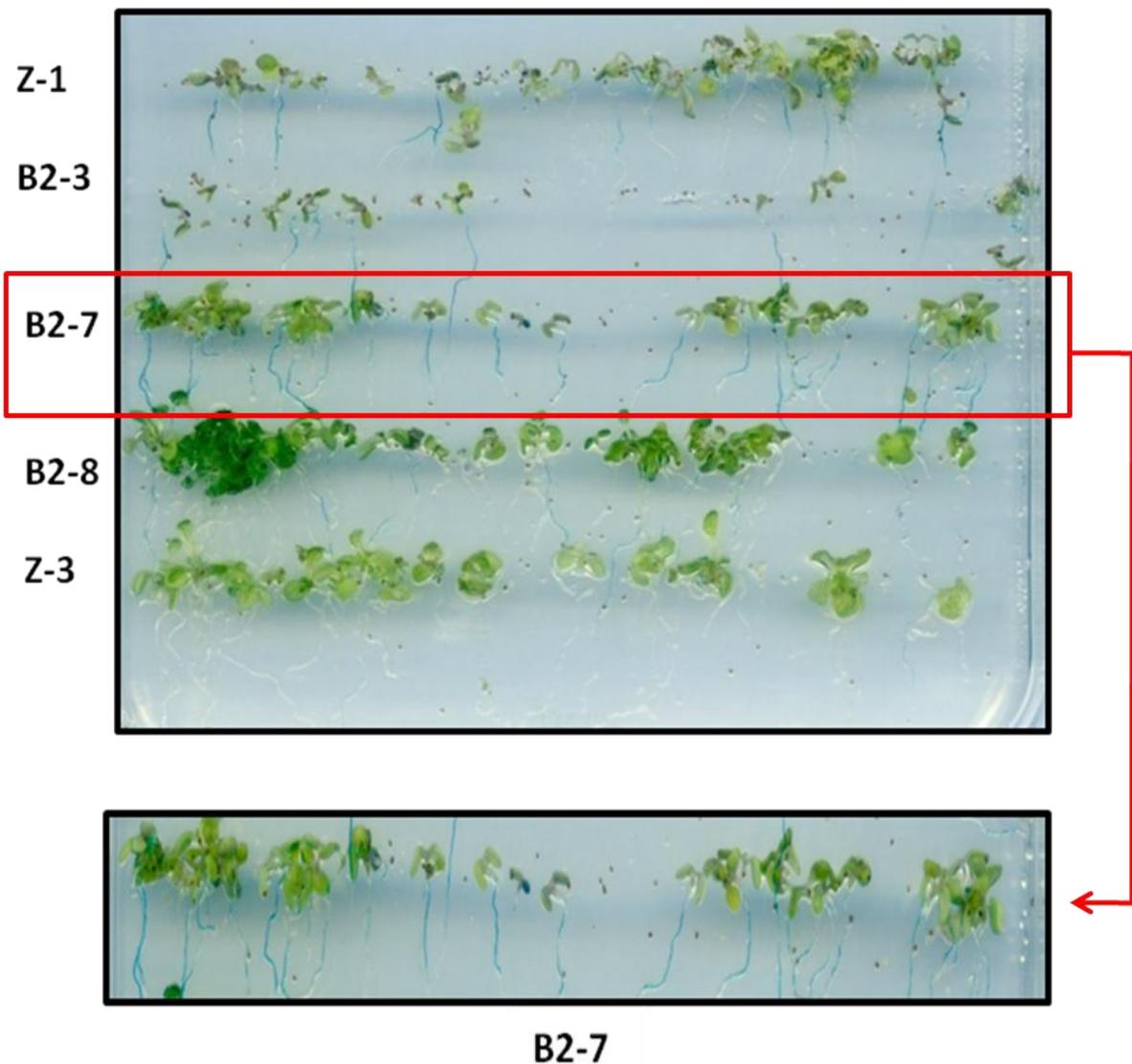


Figure 21 Identified mutants in putative E3-ligase function that targets test substrate L-GUS. Upper panel shows retesting of seedlings from pools Z, B2. In the lower panel pool B2-7 is highlighted (red boxed in upper panel) Mutant seedlings show blue color in roots after live tissue GUS staining.

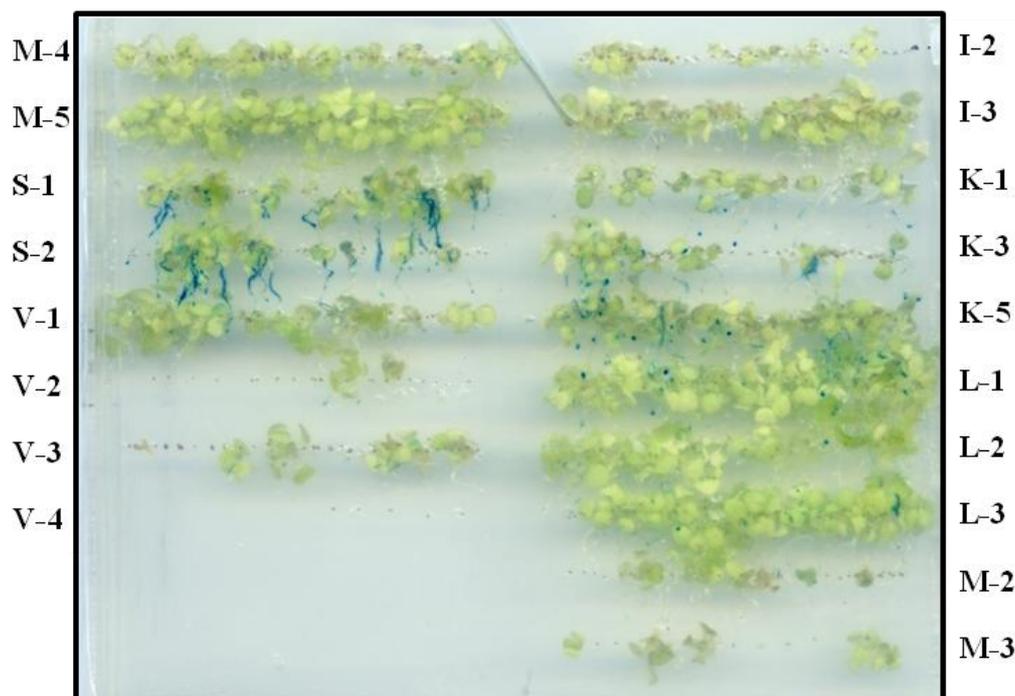


Figure 22 The second round screen of GUS assay. First round positive mutants test for reproducibility of GUS assay from pools M, S, V, I, K, L, M2 and M3. After live tissue GUS assay, putative candidate's roots show blue as a GUS assay positive. From the above retested lines lines from S and K pool candidate's shows reproducible results from first round screen.

In the second round screen, 20 candidates showed reproducible GUS assay results. Table 13 summarizes these selected lines from the second round of live tissue GUS screening for mutants that stabilize the L-GUS model substrate. These lines were further used in downstream experiments.

Table 13: Candidates for L-GUS stabilizing mutants derived from the second screening round. The lines selected as positive candidates, showed reproducible live tissue GUS assay results in the second round of mutant (F3 generation) screening.

Number	Selected mutants from pools based on GUS assay results
1	A-3-2
2	B2-4-1
3	B2-5-1
4	B2-7-1
5	B2-8-1
6	F-3-1

7	G-1-1
8	G-2-2
9	G-4-1
10	G-5-3
11	G-5-5
12	J-2-1
13	K-5-1
14	K-5-3
15	N-2-1
16	Q-1-1
17	S-1-2
18	S-2-2
19	S-2-3
20	Z-3-1

2.2.A4 Allelism test among the identified candidates

An allelism test was designed to examine if the stabilization of test substrate in the identified mutants resulted from mutations within the same gene or “in” independent genes. This test was carried out by crossing 19 (selected out of 20) individual mutants among each other. Table 15 shows the panel of the crosses. F2 generations from these crosses were germinated under inducing conditions and 2-week-old seedlings were subjected to GUS assay.

The results were examined if the crossed line showed the same GUS result as in previous first and second round of screen of individual mutants or deviated, depending on this they were either grouped into one or different complementation groups. If the mutation is in the same gene in the crossed lines it is expected, that GUS is still stabilized to the same level as in parental line; so in this case GUS pattern will remain, if not the same gene no allelism is expected and GUS should be undetectable or significantly weaker in comparison to the parental line. According to this hypothesis, lines were united in group, if their crosses did not show deviating GUS patterns from the parent lines, or separated in different groups, if the crosses showed altered GUS compared to parental lines. There were 4 independent alleles of

locus PROTEOLYSIS 8 (PRT8), which showed dark blue color after GUS assay. This group was termed *prt8* because of their putative proteolysis (PRT) function. Another 3 were independent alleles of locus proteolysis 9 (*prt9*), which showed light blue color after GUS assay. The observed results suggest that *prt8* and *prt9* might be independent mutations in different genes. The identified allelic groups and their individuals are listed in table 15.

Table 14: Schematic strategy of crossing among the mutants for allelism test. + and ✓ symbols represent the cross is planned and performed. The lines in green are first priority, those are in orange are second priority. Under flowering time – indicates late flowering and + indicates normal flowering time in comparison to Columbia WT. Blue boxed are crosses representing parental lines used for crossing. Identified mutants were also crossed to pER-L-GUS progenitor line for removal of unwanted mutations and to *Ler* to generate mapping populations.

Crossing scheme	flowering time	pER L-GUS #7	Ler	A-3-2	B2-4-1	B2-5-1	B2-7-1	B2-8-1	F-3-1	G-1-1	G-2-2	G-4-1	G-5-3	G-5-5	J-2-1	K-5-1	K-5-3	N-2-1	Q-1-1	S-1-2	S-2-2	Z-3-1	
A-3-2	+	✓	✓	■		+		++	+		+								+	+		+	+
B2-4-1	-	✓	✓		■														+			+	
B2-5-1	-	✓	✓			■			+	+			+			+						+	
B2-7-1	-						■																
B2-8-1	+	✓	✓					■						+					+				+
F-3-1	+	✓							■	+											-	+	
G-1-1	+	✓								■	+	+											
G-2-2	+	✓									■	+		+									
G-4-1	+	✓	✓									■	+			+							+
G-5-3	-												■										+
G-5-5	-	+	✓											■		+	+						
J-2-1	-														■								
K-5-1	-	+	✓													■	+						+
K-5-3	-	+															■						
N-2-1	-	✓	✓															■			+		+
Q-1-1	-	✓																	■		+		
S-1-2	-	✓	✓																	■		+	
S-2-2	-	✓	✓																		■		
Z-3-1	-	+	✓																				■

Table 15: Allelic groups and their individuals of identified putative E3-ligases of L-GUS test substrate

PRT8 GROUP	PRT9 GROUP
S-1/2	A-3
G-4/5	Q-1
B2-5/8	Z-3
K5-1	-

2.2.A5 Generation of mapping populations

In a forward genetic approach, to identify a mutant locus, a mapping population is required. Towards this end, a mapping population was generated by crossing identified mutant lines to *Ler*. Crossing of *prt8* and *prt9* to *Ler* background generated recombinants. These recombinants are the basis for further marker-based analysis to map the candidate gene locus. The crossing scheme for generation of the mapping population is showed in table 14.

2.2.A6 Phenotypic analysis of identified *prt8* mutants

To unravel whether the mutation in the L-GUS stabilizing putative E3-Ligase leads to any phenotypic deviation in comparison to wild type Col-0, seeds of identified mutants and Col-0 were germinated on selection medium and non-selection medium respectively. 2-week-old seedlings were transferred onto soil and placed in long day condition for further phenotypic analysis. The phenotypes of 10-week-old plants were scored. In primary phenotypic analysis of identified E3-ligases that stabilize test substrate L-GUS, they showed later flowering, and delayed senescence in comparison to Columbia WT plants, under long day condition.

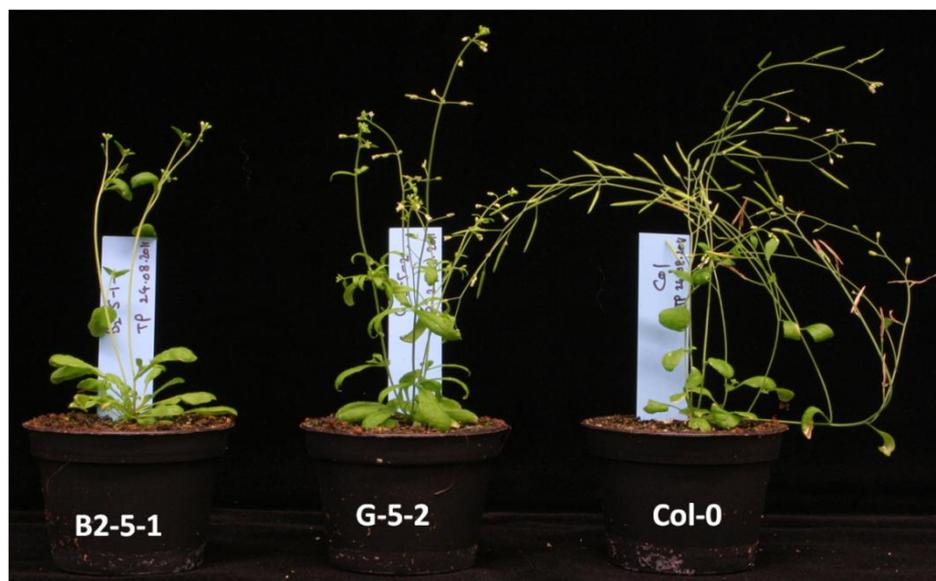


Figure 23 Phenotypes of *prt8* mutants and wild type Col-0. Mutant *prt8* lines B2-5-1 and G-5-2 show delayed development in comparison to wild type Columbia. The phenotypes were scored on 10-week-old plants. In Columbia WT, siliques already started ripening, whereas *prt8* mutants don't show any siliques in this period.

This phenotype is more prominent in *prt8* in comparison to other identified mutants. The mutants also showed decreased branching in comparison to wild type Col-0. Mutant *prt8*, B2-5 line showed a stronger phenotype. In Columbia it is clearly visible that siliques start to

ripen, whereas in the *prt8* mutant background siliques are not formed by this time. These phenotypic results suggest a possible role for PRT8 in development. The primary phenotyping results are shown in Figure 23. However, a closer analysis with outcrossed mutants remains to be done.

2.2.A7 Experimental evidence to show that L-GUS is a proteasome substrate, stabilized in *prt8* mutants

To illustrate that the L-GUS model substrate is a substrate of the proteasome, an experiment with proteasome inhibitor MG132 was designed. For this purpose, seeds from the marker line expressing L-GUS were germinated on solid MS medium supplemented with Hyg and β -estradiol. 2-week-old seedlings were subjected to MG132 treatment and without MG132 for control. After 15 h treatment, seedlings were further examined by GUS assay. The seedlings treated with MG132 showed GUS activity, whereas non-treated seedlings failed to stabilize the L-GUS model substrate (Fig 24 A and B). These results clearly indicate that the L-GUS test substrate is targeted for degradation by the proteasome.

The 3 tested lines containing different *prt8* mutant alleles were grown to compare the effect of MG132 treatment of the original line with the *prt8* mutants. The *prt8* mutants stabilize the L-GUS test substrate even in the absence of proteasome inhibitor. All the tested mutant lines, K-5-2, G-5-3 and B2-7, exhibited a similar result. All these lines could stabilize the L-GUS model substrate to the level it was stabilized in the MG132 treated marker line. These results directly support the idea that *prt8* is the novel E3 ligase of the ubiquitin dependent N-end rule pathway. Taken together the results clearly indicate that L-GUS is the substrate of the proteasome and mutant *prt8* probably corresponds to *Arabidopsis* E3-ligase that targets test substrate L-GUS. Figure 24 A, B, C, D and E illustrates these findings.

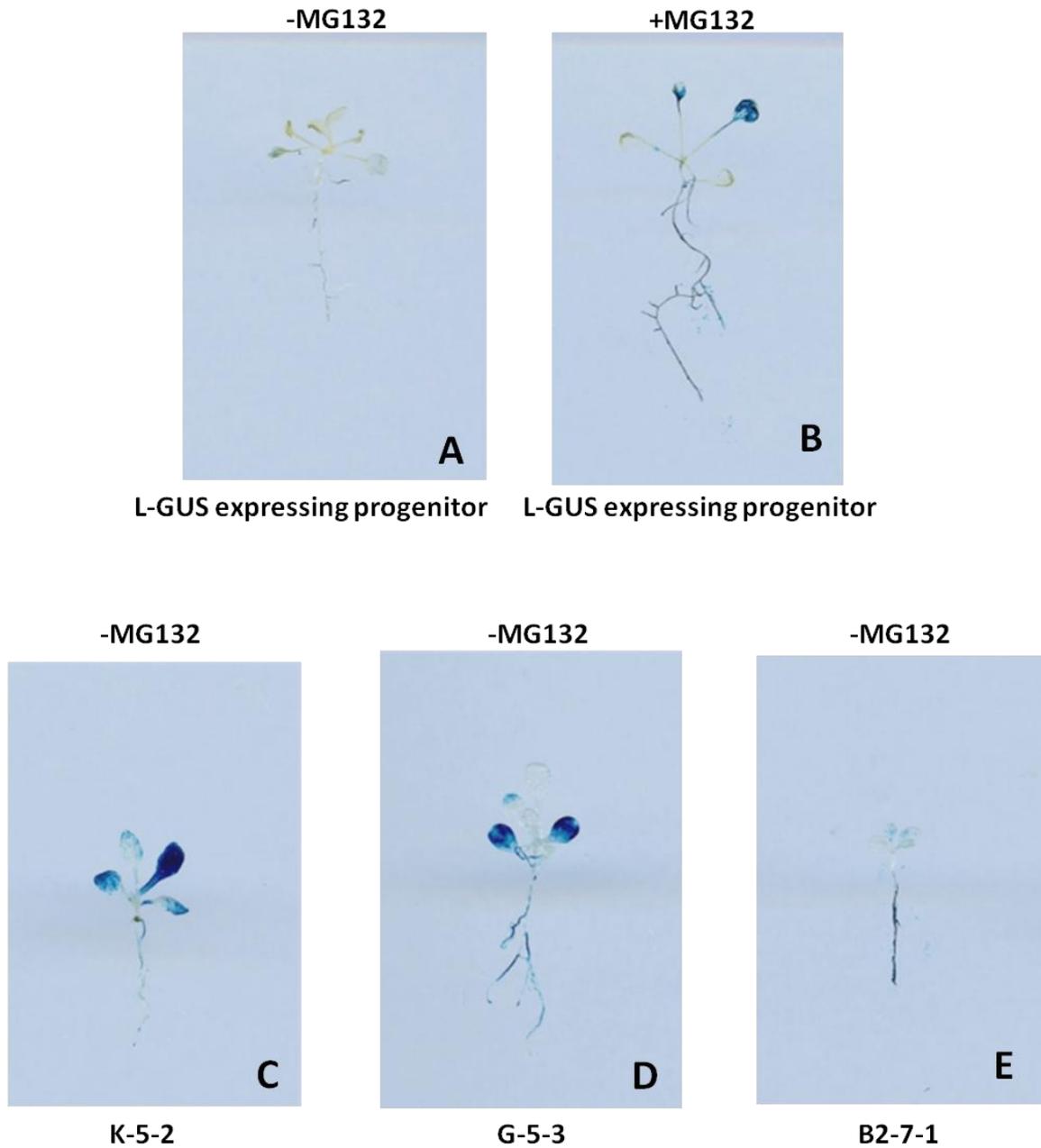


Figure 24 GUS assay with various genetic background lines treated or un-treated with MG132. Name of the plant line is written below and type of treatment is written above each panel. Panel A shows progenitor line. The L-GUS test substrate is destabilized in this line. In panel B, the progenitor line was treated with MG132 but has no mutation. The test substrate L-GUS is stabilized, indicating that L-GUS is the substrate of the 26S proteasome. Panel C, D, and E shows EMS-induced mutants in the PRT8 gene. These mutants stabilize the test substrate L-GUS without MG132 treatment. The blue colour indicates sterilized L-GUS visualized by GUS assay.

2.2.B1 Isolation of UBR-domain proteins of the N-end rule pathway by T-DNA library screening

Arabidopsis PRT7 shares homology to mammalian *UBR7*, a member of the UBR domain E3 ligase family of proteins. The functional importance of this protein in recognizing substrates of the N-end rule pathway is unknown. As many UBR domain containing proteins show E3-ligase function, it is interesting to examine whether *pri7* mutant can stabilize any test substrates. To deduce the possible function of *UBR7* as a plant E3 ubiquitin ligase of the N-end rule pathway, a T-DNA library (Koncz T-DNA library at Max Planck Institute for Plant Breeding Research, Cologne) was screened to identify a *pri7* mutant line (RIOS *et al.* 2002). This library consists of 39700 individual mutant lines. DNA samples are combined in basic pools of P-4000 and P-100, containing DNA from 4000 and 100 T-DNA lines respectively. Every individual DNA sample contained in a P-100 is present in two P-4000 pools (number and letter). To identify *pri7* mutant lines, two rounds of PCR-based screening was performed. In the first round, P-4000 pools were screened. Before performing PCR on the P-4000, a control PCR for primer specificity was performed by using the primer combinations given below with or without Col-wild type genomic DNA as a template.

P1 = T-DNA Left border primer (FISH1) + sense PRT7 gene primer

P2 = T-DNA Right border primer (FISH2) + sense PRT7 gene primer

P3 = T-DNA Left border primer (FISH1) + antisense PRT7 gene primer

P4 = T-DNA Right border primer (FISH2) + antisense PRT7 gene primer

P5 = sense PRT7 gene primer + antisense PRT7 gene primer

Primer combination P5 was used as a positive control for gene product size. The control PCRs using P1-P4 combinations with or without wild type DNA template did not yield any PCR product, hence these background free combinations were used for further screen

The above-mentioned P1-P4 combinations were used in the first round of PCR-based screening of P-4000 pools. A search for the reproducible PCR product size in number (11) and letter (K) P-4000 pools with corresponding primer combinations (p3) showed the presence of mutant/s containing a T-DNA insertion in the locus *PRT7* in P-4000 pools (Fig 25).

In a next step, in the second round of the screen, the same PCR product judged by corresponding size was being searched by performing PCR on the P100 pools which were part of the identified P-4000 pools. The second round screen was performed by using the same primer combination as in the first round. The analysis of these PCR products identified the presence of a *prt7* mutant in P-100 basic pools (501) (Fig 25).

Finally for the 100 individual lines belonging to the identified P-100 pool, seeds were sterilized and germinated on solid MS medium. The leaf material from 2-week-old seedlings was collected and material from 10 individuals was combined into a pool; forming total 20 pools (number pools 1-10, letter pools A-J, each plant is present in two pools). PCR was performed on these DNA pools to see if the corresponding band from previous results appeared or not, to identify the corresponding individual. PCR was performed using second round screen PCR protocol. A PCR product corresponding in size to the product obtained in the first and second round of the screen was found for two individual plant lines (Fig 25), indicating the success of the screen.

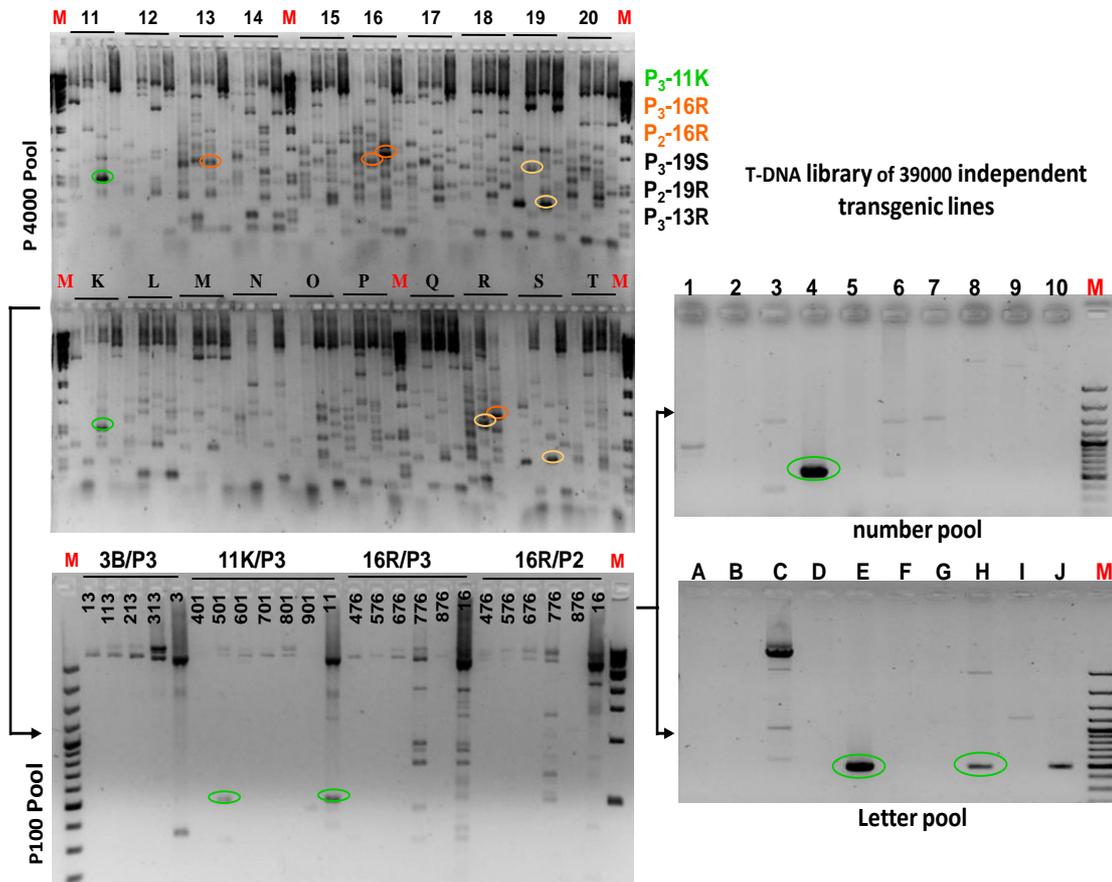


Figure 25 T-DNA library screen to identify *prt7* mutant. The left uppermost panel depicts an agarose gel electrophoretic separation of the PCR products resulting from the screening of the P-4000 pools (Top number pools and bottom Letter pools depicted). Under every number and letter pool, there are 4 lines present, they represent PCR products derived from 4 different primer combinations (P1 to P4). For example P3-11K indicates a same size PCR product resulted with P3 primer combination in number 11 and letter K P-4000 pools. Under every number and letter pool there are 4 lines present, each line contains PCR products derived from one specific primer combination (as there are 4 different primer combinations 4 lines represent P1 to P4). The left lowermost panel depicts an agarose gel electrophoretic separation of the PCR products resulting from the screening of the P-100 pools (Top number and Letter pools depicts the P-100 pools derived from that specific P-4000 pool for that specific primer combination, For example, 11K/P3 indicates P-100 pools of 401 to 901 derived from P-4000 pools of 11 and K for primer combination P3). In the second round DNA from pool 11 was used as control. In case of a true candidate line, the same size PCR product will appear in one of the P-100 pools. Circled in green are such PCR products. The right uppermost and lowermost panels depict an agarose gel electrophoretic separation of the PCR products resulting from the screening of the P-100 pool (501) DNA of 100 individuals which were pooled in two dimensional way resulting 1 to 10 and A to J pools each containing DNA from 10 T-DNA mutant lines (top number pools and bottom Letter pools depicted). A corresponding PCR product was identified in pool 4, E and H. Identified candidate is circled in green. Marker is represented in red. The expected PCR product size according to first round screen is 500 bp.

2.2.B2 Confirmation of isolated *prt7* mutant

The isolated *prt7* mutant PCR product was gel purified and subjected to sequencing for further confirmation. The sequencing result obtained is shown in Figure 23. The NCBI BLAST with the obtained sequence from T-DNA mutant search confirmed a T-DNA insertion in At4g23860. The BLAST result confirmed that the PCR fragment was derived

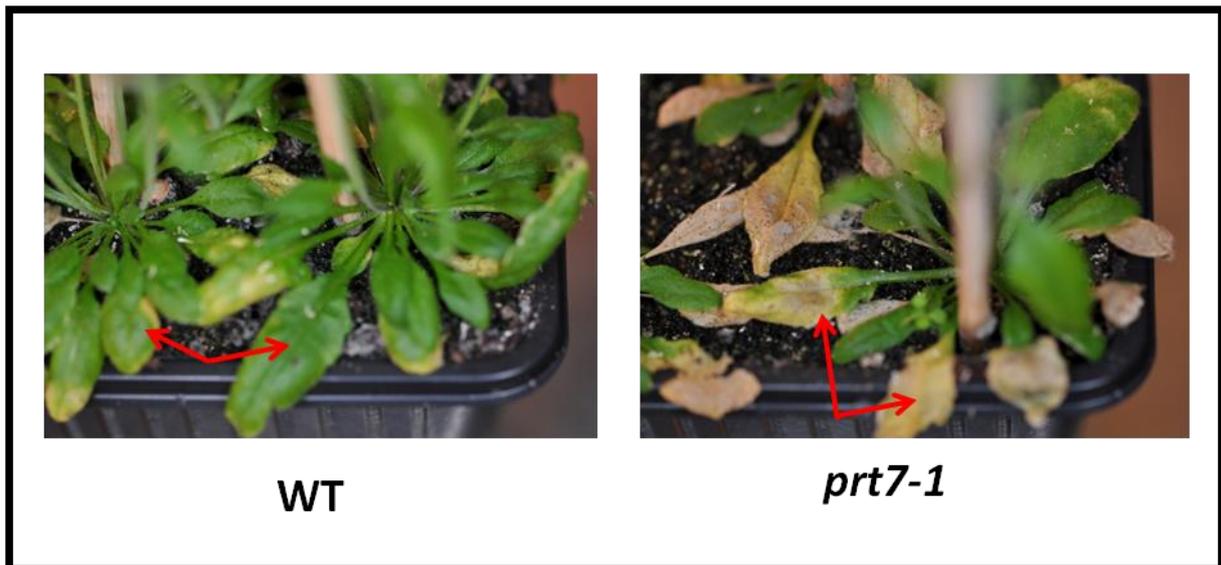


Figure 27 Mutant phenotype of the *prt7* line. The *prt7* mutant shows early leaf senescence in comparison to Columbia WT. Red arrows indicate the senescent or non-senescent leaves of *prt7* and WT, respectively.

2.2.B4 Enzymatic analysis of the *prt7* mutant

To decipher whether the *prt7* mutant shows stabilization of any of the N-end rule model substrates, the mutant was crossed into marker lines with R-, L-, D- and M-GUS substrates. GUS assay or biochemical analysis of stability of test substrate will reveal the functional importance of this protein. The test substrate will be stabilized provided the *prt7* is responsible for stabilization of that respective test substrate otherwise the test substrate will not be stabilized.

2.2.B5 Characterization of *Arabidopsis BIG*, homolog of a mammalian N-end rule pathway E3 ligase

Arabidopsis BIG shares homology with known functional mammalian N-end rule pathway E3 ligases. To understand the importance of this protein in *Arabidopsis*, a T-DNA insertion mutant in the *BIG* locus from the SALK collection was analyzed.

The mutant *big* lines were germinated and by using DNA prepared from these plants as template, PCR based genotyping was performed using T-DNA specific and gene specific oligonucleotides. The genotypically confirmed mutant lines were used for phenotypic observation. The result of this phenotypic study is summarized in Figure 28. The phenotype suggests that the *big* mutant shows among other phenotypes already described delayed leaf senescence in comparison to the wild type *Arabidopsis* (YOSHIDA *et al.* 2002). The mutant *big* also shows a strong late flowering phenotype compared to WT.

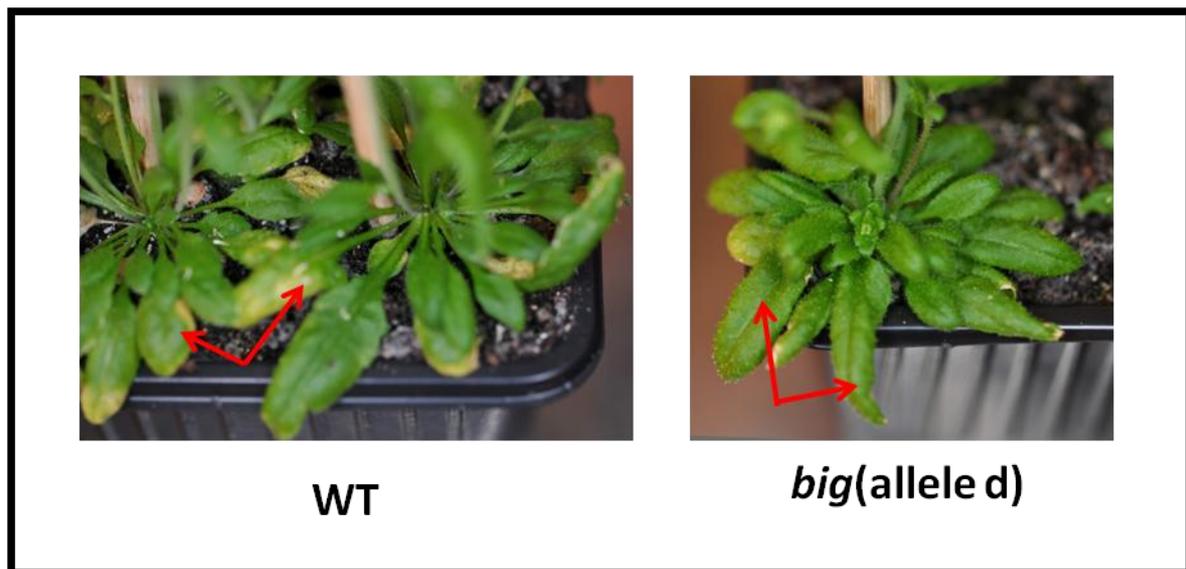


Figure 28 Phenotype of the *big* mutant. The *big* mutant shows delayed leaf senescence in comparison to Columbia WT. Red arrows indicate the senescence or its absence in leaves of WT and the *big* mutant.

2.2.B6 Enzymatic analysis of the *big* mutant

To understand the molecular function of BIG protein as *Arabidopsis* N-end rule pathway E3-ligase, the *big* mutant was crossed into a reporter line expressing the R-GUS test substrate. The homozygous state of the mutant line expressing the test substrate was verified by PCR-based genotyping. Protein was extracted from confirmed homozygous lines expressing the reporter R-GUS substrate and from a reporter line expressing the M-GUS as positive control. Columbia WT without transgene was used as negative control. The isolated proteins were subjected to biochemical analysis by western blot. The preliminary results from this experiment do not support that the BIG gene product functions as *Arabidopsis* E3-ligase (Fig 29). The R-GUS transgene was not stabilized in the mutant background of *big*. An E3-ligase PRT6 with UBR domain is known to stabilize R-GUS. The stability of R-GUS in this mutant background was not totally supported to the level of stable M-GUS. This indicates that the function is shared by another E3-ligase. To find out whether *big* shares E3-ligase function with PRT6, a double mutant was analyzed to see whether this combination can stabilize R-GUS test protein levels. In the double mutant *prt6 big*, the protein R-GUS showed stabilization, but not more than to the level stabilized in single *prt6* mutant. The stabilization of R-GUS in double mutants resulted only from the presence of *prt6* mutant allele. This preliminary result suggests that BIG does not target test substrates with basic residue at the N-termini.

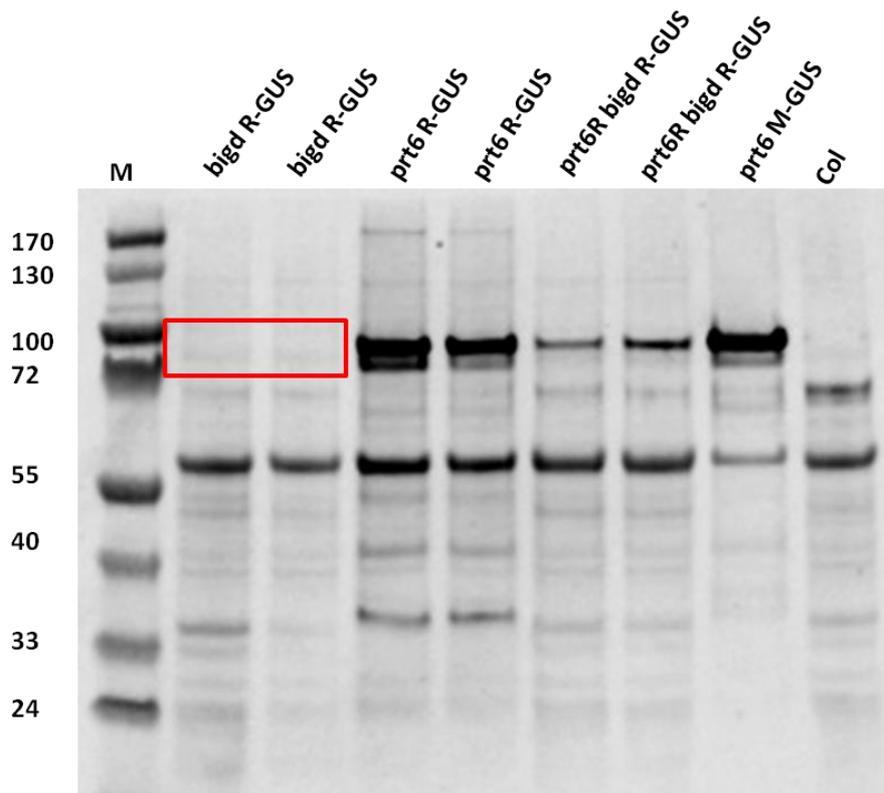


Figure 29 Identification of the N-end rule pathway test substrate by protein blotting and fluorescent based secondary antibody excitation and detection. Protein extracts from different genetic backgrounds were first separated by SDS-PAGE and subsequently blotting was performed, anti-HA tag was used as primary antibody and IR800 infrared fluorescence dye coupled secondary antibody was used for detection of fusion protein parts. Boxed region shows R-GUS destabilization in *big* mutant background. In the *prt6* mutant background, stabilization of R-GUS substrate is noticed. Prt6 M-GUS served as positive control. Col-0 served as negative control for the transgene. For detection, equal amounts of protein extracts were loaded. The expected test protein band should show up between 80 and 100 kD, the reference protein band between 25 and 35 kD.

2.2.C1 Deamidation components of the *Arabidopsis* N-end rule pathway

Many components in *Arabidopsis* share distinct homology to mammalian N-end rule pathway proteins. Deamidation of amino-terminal Asn or Gln is the first step in degradation of substrates with N-terminal Asn (N) or Gln (Q). In mammals, the process employs the proteins NTAN1 and NTAQ1, which convert N-terminal Asn and Gln into Asp (D) and Glu (E), respectively. The *Arabidopsis* genome contains At2g44420 and At2g41760, encoding putative homologs of the mammalian enzymes NTAN1 and NTAQ1, respectively. To decipher the functional importance of these proteins in the *Arabidopsis* N-end rule pathway, mutants in the respective genes were isolated. *Arabidopsis* mutants *ntan1-1* to *ntan1-7* are TILLING lines, generated in Col *er* background obtained from the Seattle the TILLING

project. Another deamidation branch component *ntaq1-3* was obtained from the GABI-KAT T-DNA insertion mutant collection.

2.2.C2 Genotyping and phenotyping of *Arabidopsis ntan1-1* and *ntaq1-3*

The *ntan1-1* mutant carries an EMS-induced C to T mutation, which resulted in the generation of a stop codon (TAA). To genotype the *ntan1-1* mutant, the region covering the mutation was PCR amplified and the PCR product was digested with enzyme *PsiI* enzyme. This enzyme cuts inside the amplified fragment only in the mutant background but not in WT Col. The homozygous mutant shows a single digested smaller band in comparison to the WT control *Ler* or Col background. In the heterozygous case, two bands, one corresponding to mutant and another one corresponding to WT, will be present. Figure 30 shows genotyping results after digestion of PCR products in the *ntan1-1* mutants and control Col. The results support that lines 1 to 10 are homozygous for the *ntan1-1* mutation.

The mutant shows very subtle phenotype. The gene model from TAIR database suggests NTAN1 consists of 9 exonic and 8 intronic regions (Fig 31).

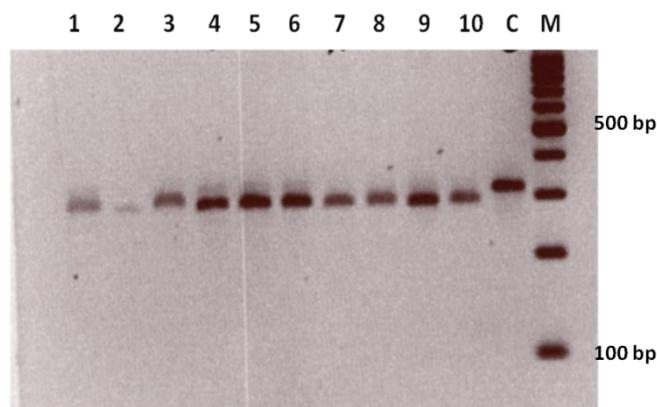


Figure 30 Agarose gel electrophoretic separation of PCR fragments after enzymatic digestion. Genotyping of *ntan1-1* lines. 1 to 10 correspond to individual *ntan1-1* progeny plants, C denotes the WT control (Columbia), M is DNA marker. Lines 1 to 10 are showing homozygous state for the *ntan1-1* mutation. The homozygous mutants are expected undergo cleavage by *PsiI* and result in a fragment corresponding to 300 bp marker size. Wild type DNA will not be cleaved by the enzyme and results in a larger fragment size in comparison to the homozygous mutant line.



Figure 31 Gene model for *ntan1-1*. *Arabidopsis* At2g44420 shows 9 exons (depicted in yellow) and 8 introns (depicted in purple between yellow regions) regions. The triangle shows the C to T modification site, which generated the stop codon in the 7th exon.

The mutant *ntaq1-3* has a T-DNA insertion, was genotyped by PCR product analysis. PCR was performed using DNA template prepared from mutant plants and wild type plants, by using mutant and WT specific oligonucleotides. The results are depicted in Figure 31. All the analyzed lines were homozygous for the mutant locus because PCR using mutant specific primers showed bands in all these samples (samples numbered in black in the Fig 31). In control PCR for the wild type allele, the mutant samples did not show any band, whereas only wild type samples showed positive result (samples numbered in red color in the Fig 32). These results clearly indicate that the mutant lines are homozygous for the T-DNA insertion. These mutants showed very subtle phenotype like in the case of *ntan1-1*. According to the TAIR database, the NTAQ gene has six exons and five intron encoding regions in the sequence (Fig 33).

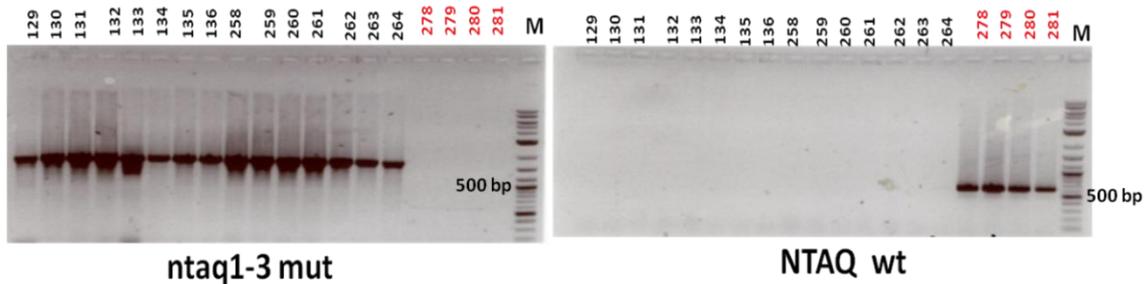


Figure 32 Genotyping of *ntaq1-3* mutants. The left panel shows the result of PCR analysis of the mutant lines after gel electrophoretic separation. Mutant lines (numbered in black) gave a PCR product, the control lines (numbered in red) did not yield a PCR product due to the usage of T-DNA specific primers for PCR amplification. The right panel shows the result of PCR analysis of the WT allele in all lines after gel electrophoretic separation. Mutant lines (numbered in black) showed no PCR product, the control lines (numbered in red) showed a PCR product. The mutant lines did not yield a PCR product due to the interruption of WT DNA by the T-DNA. M is a DNA marker.



Figure 33 Gene model for *Arabidopsis* NTAQ gene. The sequence exhibits 6 exons (depicted in yellow) and 5 intron regions (in blue internal places between yellow areas). The triangle shows the T-DNA insertion position in the 5th exon.

2.2.C3 Analysis of enzymatic function of *Arabidopsis* Ntan and Ntaq

To elucidate enzymatic activity of *Arabidopsis* Ntan, mutants in At2g44420 are being analyzed. As the mutant line was generated by EMS-induced mutation, it was several times back-crossed into Col to remove any undesired mutations. A resulting line was crossed into a reporter line expressing N-GUS. Biochemical analysis of this line will provide information on this function of *Arabidopsis* Ntan. To understand the function of the enzyme Ntaq, *Arabidopsis* lines mutated in At2g41260 were crossed into a reporter line expressing Q-GUS as a model substrate. After generation of homozygous progeny from this line, stability of the substrate can be analysed by GUS activity measurement or biochemical analysis.

2.2.D NO-mediated modification in N-end rule pathway in *Arabidopsis*

The N-end rule pathway of mammalian systems is known to be involved in a non-enzymatic process. This modification results in generation of substrates to be degraded in a proteasome dependent manner. This non-enzymatic modification involves NO and O₂ in mammals and converts Cys into Cys-sulphinic acid (CysO₂H) or Cys-sulphonic acid (CysO₃H). This oxidized form of Cys resembles Asp acid residue and gets arginylated by tRNA-Arg protein transferase. This arginylated Cys serves as primary destabilizing residue and targets for degradation by E3-ligase. Misregulation of these steps leads to deregulation of apoptosis in *Drosophila*. This implies that NO signaling and N-end rule pathway dependent protein degradation are important for proper apoptosis. In *Arabidopsis*, this modification is not characterized. As this reaction is a non-enzymatic process, there is a high probability that in plants such processes may also exist, to recruit substrates to the proteasome. To unravel this process in *Arabidopsis*, a β -estradiol inducible Cys-GUS line was tested by treatment with NO under oxic and anoxic conditions, respectively. In *Arabidopsis*, Cys-GUS is a stable substrate, whereas treatment with the NO-donor, Na-nitroprusside (0.5 mM) for 15 h resulted in destabilization of the otherwise stable Cys-GUS substrate (Fig 34a). The C-GUS substrate was still stable if the NO treatment was performed in anoxic condition (Fig 33a). NO

treatment or exposure to anoxic conditions had barely any effect on the control test protein M-GUS (Fig 34a and b), indicating stability of the control. Further, when NO treatment in oxic condition was accompanied by treatment with proteasome inhibitor MG132; the degradation was inhibited (Fig 34a). Taken together, these results suggest that NO and O₂ dependent Cys modification functions in a similar way in *Arabidopsis* as in sophisticated mammals.

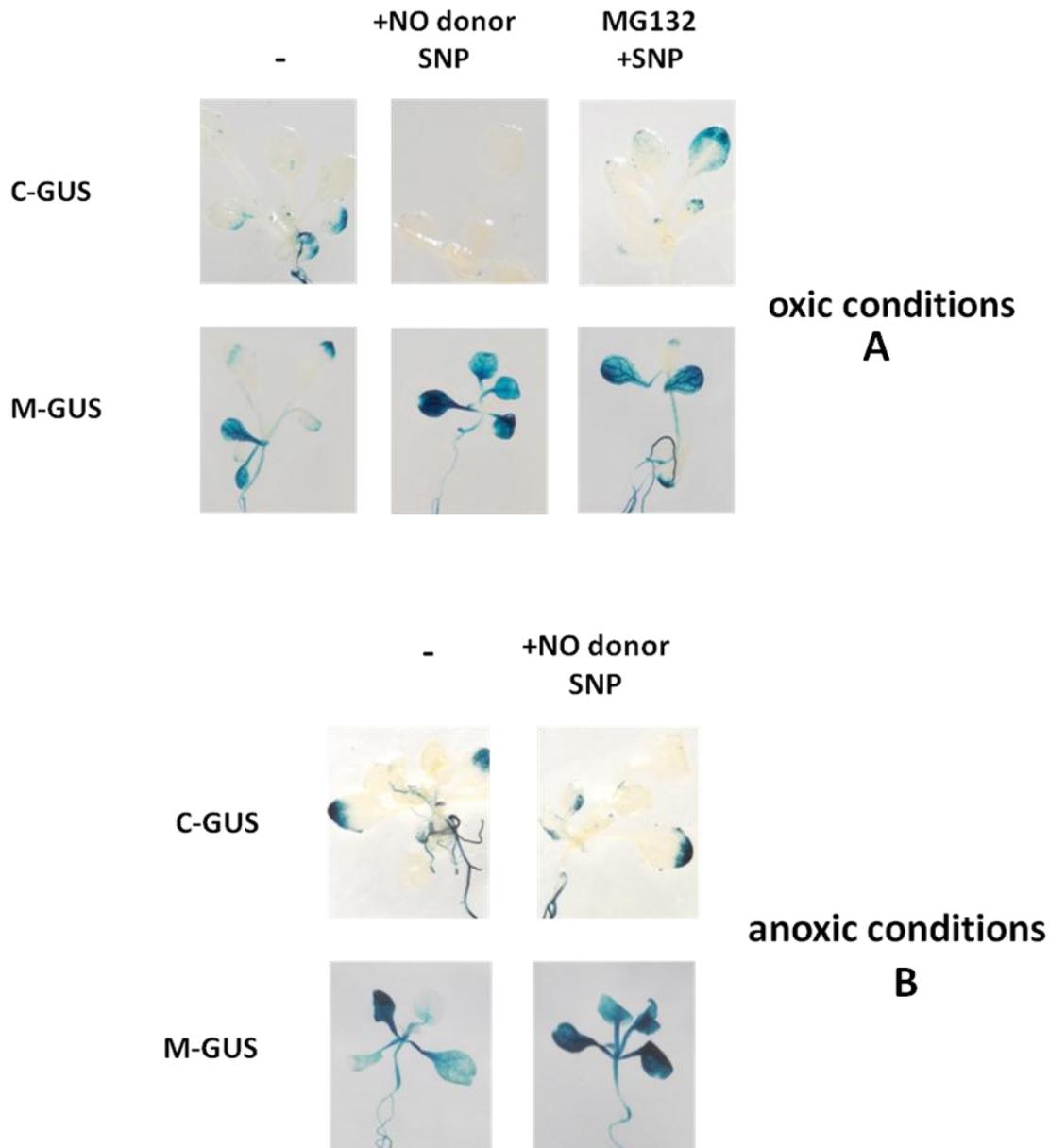


Figure 34 C-GUS stability test under oxic and anoxic condition. The Figure shows results of 15 h treated and un-treated C-GUS and control M-GUS test substrate GUS activity. Panel A shows activity of GUS substrate under treated and untreated oxic condition. Test substrate M-GUS was used as control for treated and untreated oxic condition. Test substrate C-GUS is stable under untreated oxic condition. It becomes unstable when treated with NO in oxic condition. C-GUS test substrate is stabilized in Col-0 line even after NO treatment in the presence of MG132 in oxic condition. It implies that C-GUS is a target of the 26S proteasome. In contrast to C-GUS, M-GUS expressing line has not shown any noticeable difference in GUS activity under treated or un-treated oxic condition. The panel B shows C-GUS and M-GUS stability under anoxic condition with or without NO treatment. The results reveal that C-GUS test substrate and M-GUS control test substrates are fairly stable in NO treated and untreated anoxic condition. This implies that for degradation of C-GUS, NO should be accompanied by O₂ for efficient degradation.

3 DISCUSSION

3.1 Discussion part 1

3.1.1 The major consequence of ubiquitination inhibition in *Arabidopsis* is cell death

In this research work inhibition of ubiquitin K48 linked mediated poly-ubiquitin chain formation was used as a tool to understand ubiquitination downstream molecular events and its links to the cell death and other processes. In general, to understand the biological and functional importance of any pathway or components of a pathway; 2 main classical approaches can be taken. The first one is forward genetics that includes mutant generation followed by identification of genes responsible for the mutant phenotype. The second approach involves reverse genetics which starts with a mutation in the desired gene and which aims for the identification of a phenotypic deviation for that given mutation.

To understand the importance of the ubiquitination pathway in *Arabidopsis*, in a forward genetic approach ubiquitination was perturbed. To perturb the ubiquitination process, ubiquitin Lys 48, which is a main determinant of substrate poly-ubiquitination and stability, was modified to Arg. Hypothetically this variant inhibits poly-ubiquitin chain formation, which is an essential determinant of substrate degradation via the proteasome. Thus one could assume that expression of this ubiquitin variant would give the same result as inhibition of the proteasome. This can serve as a powerful tool to study downstream signaling effects of the ubiquitin proteasome dependent proteolysis pathway, without the proteasome being inhibited. This gives the possibility to uncover important players that are dependent on the ubiquitin dependent protein degradation pathway.

Lesions on the leaf surface are the main visual sign of cell death in plants. The leaf lesions phenotype was also noticed in the rice *spotted leaf11* (*spl11*) mutant, which has a mutation in a gene encoding a U-box/ Armadillo repeat protein with E3 ligase activity (ZENG *et al.* 2004). In yeast it has been shown that ubiquitin K48 serves as linkage site for formation of multi-ubiquitin chains that is required for the degradation of some substrates of this pathway. Further it was also shown that, expression of ubiquitin K48R showed cell cycle arrest with a terminal phenotype which was evidenced by replicated DNA, two-lobed nuclei and mitotic spindles. This has clearly indicated that ubiquitin K48 dependent substrate degradation is vital for regulation of developmental processes in yeast (FINLEY *et al.* 1994).

In the adult *Arabidopsis RV86-5* line, when variant ubK48R was expressed, lesions on leaves were noticed which is a cell death phenotype in plants. When the variant ubK48R was expressed in the germination stage, the cell death took place at the seedling stage. Presumable reason for the cell death phenotype (in *RV86-5*) would be failure of poly-ubiquitin chain formation. The noticed lesions in the ub K48R expressing *Arabidopsis* suggest that there might be a possible link between the cell death process and ubiquitin dependent protein degradation. In the ub K48R expressing line, deregulation of poly-ubiquitin chain formation on substrates, which require poly-ubiquitin chain formation by K48 chains to be recognized by proteasome might have failed. Substrate accumulation might be lethal for the plant and as a consequence the cell death takes place. It might be that multiple substrates are stabilized that are supposed to be poly-ubiquitinated via ub K48 and the noticed cell death phenotype is caused by multiple substrate stabilization rather than stabilization of a single substrate.

3.1.2 Suppressor of ubK48R expression, *sud2*, rescues lethal phenotype

In order to elucidate the molecular mechanism behind the cell death phenotype in *RV86-5*, a suppressor of the *RV86-5* mediated cell death generated by EMS, *sud2*, was isolated. The *sud2* mutant can rescue the lethal phenotype of *RV86-5*. When testing these *RV86-5* and *sud2* plant lines on selection media for survival *sud2* plants can withstand the lethal effect of the ubK48R transgene expression whereas the WT allele *sud2* (*RV86-5*) cannot (Fig 9). In *sud2* plant lines a component that caused cell death in the *RV86-5* background might have been mutated by EMS mutagenesis and this mutation consequently might have rescued the phenotype in the *sud2* background. The suppressor *sud2* phenotype also showed Mendelian monogenic segregation. This finding further suggests that the observed phenotype might have resulted from accumulation of a single component.

3.1.3 Mapping - a way to hunt for *sud2* locus: results suggest *sud2* position on chromosome III

Map-based cloning is extensively used to identify candidate genes responsible for mutant phenotypes derived from EMS mutagenesis in *Arabidopsis*. In this study by taking advantage of well developed polymorphic markers between *Col* and *Ler*, the *sud2* mutant locus was initially confirmed on chromosome III. 192 recombinants were tested with four SNP markers. The results were in good agreement with results from rough mapping, indicating the mutant position on the south arm of chromosome III, which was formerly identified by the previous

graduate student Marcus Garzon. The analysis of 192 recombinants also further narrowed the mutant locus position on chromosome III to the region from between markers MUO22 (11.4Mb) and CIW4 (18.9Mb) to between markers T32N15 (16.36 Mb) and T6H20 (17.2 Mb).

3.1.4 Does larger population help to overcome the limitation of low recombination?

The mapped region suggested low recombination in the region of interest. There are two ways to overcome such undesired situation in mapping. The first one is to make backcrosses and generate new recombinants. These new recombinants can be used for marker-based analysis, which is a time consuming process. The second option is simply to generate a large mapping population and analyze markers for fine mapping. In the present research work, the latter option was chosen for two reasons: one is that generation of a large mapping population is easy and less time consuming when compared to generating new recombinants and to continue mapping. The second reason is that this mapped region contains three repeated segments and because of this the backcrossing strategy may not support further fine mapping of the mutant locus.

A systematic analysis of several SNP and dCAPS markers on a population of 1239 individuals has delimited the mutant locus to a 350 kb region on the chromosome III. This covered the region from At3g44400 to At3g44900. It is noteworthy to mention that the *sud2* phenotype co-segregated with T32N15, suggesting the *sud2* mutant locus is present in the vicinity of this marker. Out of 1239 analyzed candidate recombinant lines, 11 were informative regarding position of the mutant. Hypothetically one could expect that in an ideal case the closer one moves towards the mutant locus, on the one hand homozygosity increases and on the other hand heterozygosity decreases. This was exactly the scenario noticed in *sud2* mutant recombinant lines that were informative in fine mapping. This finding has significantly reduced the region where the mutant locus can be found but three repetitive sequence regions spanning this mapped region have proved to be a bottle neck for further marker based mapping to the candidate gene level.

3.1.5 Fragment library construction: an alternative route to reach to the *sud2* locus

The technical advancements, for example the platforms for well-developed markers for mapping and the development of next generation sequencing methods have enormously speeded up the mapping procedure. There are still biological issues in the mapping process that hinder the identification of a gene responsible for a particular phenotype. One such issue is a leaky mutant phenotype and another one is for instance the mutation is located near the centromeric region. The latter case results in low recombination, when crossed to other ecotypes to generate polymorphisms and thus precludes molecular marker analysis between two different ecotypes. In this research work, in *sud2* fine mapping demonstrated low recombination and on top of it, the mapped region had around 35% sequence repeats. In this area, analysis of more markers would not show new polymorphisms. As a way out to this complex problem, a fragment- based library was designed to PCR amplify the whole region of interest, in this case a 350 kb sequence area on chromosome 3. Though many highly efficient DNA polymerase enzymes exist in the market, many of them are not suited for amplifying fragments with a size beyond 10 kb. Because of these practical reasons, the delimited area of interest on chromosome 3 was PCR-amplified by generating a library of sub-genomic fragments of about 10 kb in size, covering the region of interest on chromosome 3. A total of 41 fragments with overlapping areas to the respective adjacent fragments were generated. Despite the difficult area with many repeats, because of the specificity of primers combination which is resulted in obtaining mostly a clear single bands and even if not, by purifying the band corresponding to the desired size of band one could maintain specificity and amplify the region of interest. This process proved to be a way to overcome the undesired problem generated by the biological material due to low recombination.

3.1.6 Next generation sequencing: an excellent tool for mapping process for the identification of genes of interest

The forward genetics approach is destined to link a certain phenotypic trait with a certain allele or mutated version of a locus. A general strategy in forward genetics is first to generate mutants exhibiting a certain phenotype of interest by EMS mutagenesis followed by marker-based mapping. The recent improvements in sequencing methods made available the whole genome sequence information of several *Arabidopsis* ecotypes and made comparatively easily accessible the genomic information of any mutant line and this facilitates the finding of single nucleotide changes generated by EMS-mutagenesis which are generally G to A and C to T.

The main advantage of the new technologies is that the number of reads generated will provide high quality of information. This huge array of sequence data provides the possibility to answer long-standing questions related to phenotypic traits.

The mutant *sud2* DNA prepared by sub-genomic library PCR amplification was subjected to Solexa sequencing. The generated high quality sequence reads were used for alignment to the reference sequence to identify mutations caused by the EMS treatment. For theoretical reasons, one can also use whole genome sequencing but for the present situation the whole genome sequencing was not required. This was because the mutant locus has already been mapped to a small region on chromosome III. When one has narrowed down the position of the mutant locus between two markers, it is extremely helpful to use only the region of interest for sequencing, provided it is manageable. This would save one from dealing with an undesired data load. The data generated in this way do not occupy so much space on the analyzing machine as in the case of whole genome sequencing, and it is easier to achieve high coverage.

The number of short reads derived from Solexa sequencing was above 10 million reads. Most of the reads could be aligned to a specific sequence region of the reference sequence. This is indicating that derived short reads are of high specificity. Every single base of the aligned reads obtained was scored for alignment with reference sequence. When a single mismatch in the *sud2* template precludes perfect alignment, the number of bases for that given reference sequence is expected to be either zero, or low as long as mismatches are not allowed. When a limited number of mismatches are allowed in alignment, the number of aligned reads is equal to the surrounding bases. The ratio between reads with tolerated 20% error and 100% fitting reads (that is with 0% error) was calculated for each base. Base positions where at least 50 times more reads with 20% error matches than 0% error were taken as putative targets to find a polymorphism. The 20% mismatch tolerance could also detect insertion and deletion mutations. Polymorphism numbers 2, 3, 4, 5 and 6 (Fig 14) were identified as such polymorphic changes, indicating the sensitivity of the method to detect polymorphisms at the base level.

3.1.7 Analysis method provides graphic quick view to identify candidates showing polymorphism

One very interesting observation is that the analysis method applied for alignment of sequence reads to reference sequence generates a clear graphical view. Looking at the

graphical view of the C-100 dataset alignment one can easily predict whether sequence variation is present or not between the compared sequences. As it is shown in Figure 14, wherever a “V” shape appeared, such a spot was correlated to a predicted nucleotide change. The falling number of reads at the predicted mutation position provided the plot of the spot with a distinctive “V” shape. This graphical visualization can be used as criterion to determine mutation sites in a given sequence. Table 9 shows the confirmed candidates that were identified using the “V” shape criteria in the graphical view.

3.1.8 Microarray identified differentially expressed candidates

The main aim of performing microarray-based analysis of differential transcript abundance between *sud2* and *RV86-5* was to identify candidate genes responsible for the cell death phenotype. With microarrays, one can measure gene expression on the whole genome level and generate functional data for a given mutant background for many genes at a given time point. Microarrays measure expression of transcript level of genes whose probes are available. The ATH1 array consists of 21539 probes and was employed in this work to determine gene activity in *sud2* and *RV86-5* backgrounds by using the expression data. The microarray expression patterns in these two backgrounds can reveal which gene expression is associated with the cell death phenotype in the *RV86-5* mutant background and the expression of which gene is associated with the rescued phenotype in the *sud2* background. This method also facilitates the association of unknown genes with known genes depending on co-regulation of co-expressed genes. In *Arabidopsis*, for data mining and analysis co-regulated genes study a co-response database (CSB.DB) built based on publicly available expression data is available (STEINHAUSER *et al.* 2004). A recent analysis of co-regulation identified genes involved in cellulose synthesis by using publicly available *Arabidopsis* arrays of cellulose synthase genes (PERSSON *et al.* 2005). Also, co-regulation analysis of known signaling components was carried out by using the open source identified brassinosteroid-related genes (LISSO *et al.* 2005). These findings show the potential use of mining publicly available expression profiles for co-regulation studies and the discovery of gene functions. *Arabidopsis* expression data analysis tools are available at NASC. The web based Genevestigator is an analysis platform and an integrative visualization tool for expression data (ZIMMERMANN *et al.* 2004). Mapman was developed to visualize genomics data in the form of diagrams of biological processes and metabolic pathways, which helps to make connections between gene functions and biological pathways (THIMM *et al.* 2004).

Microarray expression data output mainly depends on quality of the sample used for expression study. Thus it is worth paying attention to the quality of the samples used for hybridization to the ATH1 chip.

To make biologically meaningful connection between the cell death phenotype in *RV86-5* and downstream signaling of the ubiquitin proteasome dependent protein degradation pathway, two main comparisons of expression data were made. The first comparison includes expression data between un-induced *RV86-5* and induced *RV86-5*. In this case it gives information about transcripts that are differentially expressed in induced *RV86-5* background and those might be the direct or indirect cause for the cell death phenotype. The second comparison includes expression data generated between induced *RV86-5* and *sud2*. This comparison provides information about direct candidates involved in the lethal phenotype in the *RV86-5* background and such candidates may be higher or lower expressed in *RV86-5* background in comparison to *sud2*.

3.1.9 Lessons from expression comparison between un-induced and induced *RV86-5*

When comparing datasets, 855 genes are down regulated minimum 2.4 fold and 780 are up-regulated to minimum 2.4 fold in addition, 121 genes are up-regulated between 2.2 and 2.4 fold. Thus overall number of up-regulated and down regulated is nearly same. This should also be true for gene classes belonging to various biological processes and this criterion serves to normalize comparison of net changes of transcripts belonging to specific biological pathways. List of selected differentially expressed gene are shown in Appendix 1 and 2.

Are there changes in genes of ubiquitin – dependent protein degradation?

In this category, 22 genes are 2.4 to 7 fold down-regulated and 20 are up-regulated from 2.4 to 10 fold. Most of these differentially expressed ones are cullin type ligase subunits. No proteasome subunits are listed in these groups. One anticipates a role for these components in the cell death phenotype, but the data suggest there is no feedback from these components in *RV 86-5*.

Protein phosphorylation and dephosphorylation

These post translational modifications are often signals for ubiquitination or de-ubiquitination processes (JOO *et al.* 2008; YOO *et al.* 2008) or for diverse stresses. 36 kinases are up-regulated more than 2.4 fold. Ten of them contain a Leu rich repeat domain (LRR) domain.

One kinase interacting protein and one kinase inhibitor are up-regulated and 6 phosphatases are up-regulated. 23 kinases are down regulated more than 2.4 fold, 12 of them contain LRR domains. 7 phosphatases are down regulated. There are more kinases induced than down-regulated. Those with low transcription are receptor type kinases, consistent with down-regulation of some defense components. It might be that phosphorylation influences the protein turnover, as many of the ubiquitination substrates are either phosphorylated or dephosphorylated before being targeted for ubiquitination.

Chloroplast targeted or encoded genes

This is the largest group of genes coordinately changed. 234 genes are down regulated, 211 of them are altered in their expression between 2.4 and 10 fold. These include components of transcription, translation chaperones, amino acid biosynthesis and of the photosynthetic apparatus. In contrast to down-regulated genes, only 37 are up-regulated.

Mitochondrial targeted or encoded genes

In contrast to chloroplast genes (more down-regulated Vs less up-regulated), for mitochondria, 21 were down-regulated, and 60 were up-regulated. Because of down-regulation of many genes in the chloroplast, as an alternative for the required energy, plants may rely on mitochondria under an energy crisis situation. When there is such an energy crisis, plants decrease growth. If partial inhibition of protein degradation would result in growth arrest as the primary event, with lowered energy requirement as secondary consequence, one will not anticipate that mitochondrial genes show significant up-regulation.

Light signaling

Why would the deregulation of ubiquitin dependent protein turnover lower chloroplast performance and turn on the mitochondrial machinery for energy production? One possible hypothesis is that in *RV86-5* when *ubK48R* is expressed, plants may face impaired light signal transduction. One well-studied E3-ligase in light perception is CONSTITUTIVE PHOTOMORPHOGENIC 1 (*COP1*), mutation in which leads to a de-etiolation phenotype. These mutants develop as if light is available even in the absence of light. In *RV86-5*, it might be that the opposite scenario exists and chloroplasts may be adjusted to lower light than actually present. The expression profiling data also suggests that in *RV86-5* background there are problems with blue light perception.

In *RV86-5*, blue light photoreceptors expression is significantly altered at the transcriptional level, phototropin1 and 2 (PHOT1 and 2) being 2.3 and 2.7 fold down-regulated, respectively. An NONPHOTOTROPIC HYPOCOTYL3 (NPH3) family protein and an NPH3 family member with BTB domain (which was noticed in suppressor mapping result) are 7.3 fold down-regulated. A PHOT1, 2 interacting protein is down regulated 5-fold. A transcription factor induced by high light is 3-fold down-regulated and also the ELONGATED HYPOCOTYL 5 (HY5) homologous transcription factor HYH is down-regulated. While all these are -regulated, no blue light component is up-regulated. In contrast, no red light response component was down- regulated, only one red-light signal transduction component, far-red-light-insensitive 1 (FRE1), is mildly up-regulated (2.6 fold). With this it is easy to hypothesize that blue light signaling is critically dependent on protein turnover and that decreased turnover may be perceived as decreased light intensity. If energy crisis due to misregulation of chloroplast hypothesis is correct, plants must lower their growth rate or even stop growth. One can observe down-regulation of growth-associated genes.

Translational machinery (ribosome assembly, mRNA processing etc)

There is a clear down regulation of the translational machinery genes including components of ribosome biogenesis and mRNA maturation. 57 genes of this category are down regulated; in contrast only 9 genes of this category are up-regulated. In addition, chaperone gene expression is decreased, 13 cytoplasmic or ER chaperone genes are down-regulated, whereas 4 are up-regulated. This is in good agreement with the decreased need of chaperones for newly synthesized proteins, and it is again in contrast to an accumulation of denatured proteins (due to the deregulation of the ubiquitin-dependent degradation machinery), which would be normally followed by up-regulation of mechanisms to remove denatured proteins. However, the plants may interpret accumulation of denatured proteins as heavy metal intoxication (as heavy metals are known as protein denaturants).

Plant growth hormones

Plant physiological changes co-incide with plant hormone synthesis, spatial and temporal distribution or response. Thus it is interesting to know the hormonal gene expression in *RV86-5* background.

Auxin

From auxin biosynthetic enzymes group, 5 are down regulated, but only 1 is up-regulated. One enzyme of auxin desensitization is up-regulated. From auxin carriers group, one auxin

influx carrier, LAX3, is down-regulated and one up-regulated. Two efflux carriers, PIN1 and PIN4, are up-regulated. Eight auxin responsive transcriptional regulators are down-regulated, 4 are up-regulated. 6 of the SAUR proteins are down-regulated but only 2 are up-regulated. Taken together these changes may result in a net decrease in intracellular auxin concentration and auxin response.

Cytokinin

Among the enzymes involved in cytokinin-biosynthesis, 2 are down-regulated; one enzyme of cytokinin turnover is up-regulated. 3 cytokinin-responsive transcription factors are down-regulated, one is up-regulated. Four genes of cytokinin response regulators, ARR proteins, are down-regulated. With these observations it seems that there is a decrease in cytokinin biosynthesis and response.

Brassinosteroids

Four brassinosteroid (BR) responsive genes are down-regulated, a change that includes the transcription factors BR enhanced expression 2 (BEE2) and BEE3. Two genes of the BR response are up-regulated; one of them is listed as BR and cytokinin-regulated. One enzyme supposedly involved in BR biosynthesis is down-regulated. These results suggest that brassinosteroids may contribute to the observed cessation of growth in *RV86-5* background.

Gibberellin

One gene involved in steroid biosynthesis, and two genes of Gibberellic acid (GA) biosynthesis are up-regulated. No GA biosynthesis genes and GA regulated genes are down-regulated. 3 GA responsive genes are up-regulated. These results suggest that GA is the only growth promoting hormone with a positive co-regulation with induction of the ubiquitin variant. These changes in components of hormone signaling result is consistent with the noticed decrease in translation process related components of processes necessary to withstand the “energy crisis”.

Defense-related genes

47 genes with possible connection to defense responses are down-regulated, 44 are up-regulated upon induction of the dominant negative ub variant. Thus this output indicates no clear trend.

There are more jasmonic acid (JA) induced genes up-regulated than down-regulated: 12 are up-regulated and 7 are down-regulated. Moreover, one of the down-regulated genes is WRKY70, a suppressor of JA responses. This change thus may have the effect of up-regulating JA responses. Likewise, there are three enzymes of ethylene biosynthesis which are down-regulated, while 3 are up-regulated. Many ethylene-responsive genes are also JA-responsive. Among ethylene-responsive genes that are not annotated as JA responsive, 6 are up-regulated and four are down-regulated. Thus, there is apparently a slight net-increase in JA / ethylene responses.

Many genes are adversely influenced by JA and salicylic acid (SA). One such gene is WRKY70, is down-regulated and contributes to favoring JA responses over SA responses. In addition, 5 genes which are annotated as being SA responsive are down-regulated, and 7 are up-regulated. The general impression remains that changes in defense-related genes are governed more by JA and ethylene than by SA.

Cell wall remodeling related genes

In the group of up-regulated genes, 6 genes are up-regulated more than 10-fold, 33 genes between 10- and 2.4 fold, and six between 2.4 to 2.2 fold. In the group of down-regulated genes, 3 genes are down-regulated more than 10-fold and 28 genes are down regulated between 10- and 2.4 fold. Among the down-regulated genes are inhibitors of cell wall modifying enzymes, among the up-regulated genes are enzymes of cell wall modification (e.g., pectin esterases). There is high induction of the genes of lignin biosynthesis.

The above results support the hypothesis that the cell wall is strengthened.

Biosynthesis of UV protecting flavin compounds and defense-related isoflavonoid compounds is down-regulated, indicating that newly synthesized phenylpropanoid compounds are channeled into lignin biosynthesis, not into the flavone/isoflavone branch. This is also consistent with the hypothesis that plants “think” to be in low light, so that no additional UV protection is necessary.

2 genes of flavonoid biosynthesis are more than 10-fold down-regulated, 7 genes between 10- and 2.4 fold down-regulated; only one Anthocyan modifying enzyme is up-regulated, and one transcription factor, ANTHOCYANLESS 2, is mildly up-regulated (2.3-fold). This is to be seen in relation to the up-regulation of the phenylpropanoid pathway genes such as

phenylalanine-ammonia lyase (PAL) enforces the impression that lignin biosynthesis is up-regulated, not flavin biosynthesis.

Neither the alterations in defense-related genes, nor the “energy crisis” have an obvious connection to the finding that cell walls are significantly remodeled upon ub K48R induction. One could thus hypothesize that cell wall integrity and/or strength is regulated by a separate circuitry that depends on ubiquitin-mediated protein turnover.

Red-ox homeostasis (cytoplasm)

It was previously shown that cell death associated with ub K48R expression co-occurs with increased intracellular presence of reactive oxygen species in *Arabidopsis* (SCHLOGELHOFER *et al.* 2006). It was therefore of interest to see whether induced plants strengthen their enzyme machinery dealing with reactive oxygen species (ROS). Surprisingly, this seems not to be the case. One interpretation is that induced plants expect that chloroplasts are no longer a source of ROS to be taken care of (without light, photo-oxidative processes would indeed be no source of oxidative stress). A decreased synthesis of red-ox homeostasis enzymes may therefore be the result of co-regulation with chloroplast genes.

Among the group of genes that could potentially contribute to red-ox homeostasis, we included the Thioredoxins and Glutathione S transferases, oxidoreductases with Rossmann fold, Fe and 2-oxoglutarate-dependent redox enzymes. 40 genes that could potentially contribute to red-ox homeostasis (chloroplast and mitochondrial enzymes not counted) are down-regulated, but only 16 are up-regulated. These numbers may be compared to the distribution of changes for Cytochrome P450 type oxygen consuming red-ox enzymes, which are not known to contribute to redox homeostasis, but are involved in multiple biosynthetic pathways. There are 11 up- and 11 down-regulated P450 enzymes.

De-toxification (efflux carrier, heavy metal induced genes)

There is no indication that inhibition of ub-dependent proteolysis is interpreted as intoxication by xenobiotics. 12 MatE, EamA or ABC transporters that could potentially explore xenobiotics are down-regulated, whereas 7 are up-regulated. This contrasts with genes that counteract heavy metal intoxication: proteins that export heavy metals or are involved in synthesis of heavy metal ligands such as Nicotinamine are up-regulated (12 genes), and only 2 genes are down-regulated. This can be seen in the context of heavy metals as protein denaturants. Partial inhibition of ubiquitin-dependent proteolysis may therefore lead to a mild increase of denatured proteins, which is interpreted by the plant as heavy metal intoxication.

Genes associated with cell death

As indicated above, there is no clear trend regarding induction of defense-related genes. There may be a need to down-regulate “unnecessary” genes with respect to the energy crisis caused by shut down of chloroplasts, but signs of stress may exist that would suggest stress responses (after all, plants die some time after induction of ubiquitin K48R).

In order to address the question why plants with an inhibited ubiquitin system die, mis-expression of genes previously associated with cell death processes was specifically analyzed. 7 genes with annotation as senescence-induced or -related are down-regulated, and 6 genes of this class are up-regulated. Among the down-regulated genes are ORE1, SEN1, and WRKY22, whereas ORE7 and WRKY45 are among the up-regulated genes. Regarding genes with potential connection to fast cell death programs, there are 4 down-regulated and four up-regulated genes. Interestingly, two of the down-regulated genes presumably act in a pro-apoptotic fashion in defense responses (DND1 and DND2), and one of the up-regulated genes is a member of the BAX inhibitor family, which may raise the cell death threshold. A cysteine protease with role in xylem differentiation (proto-xylem cell death) is up-regulated, whereas another member of this class is down-regulated. Metacaspase 1 (AMC1) is 2.5-fold up-regulated. Taken together there is no obvious trigger of fast cell death.

The overall comparison of data suggest that ub K48R variant induction leads to increased phosphorylation, which is anticipated as many of the E3-ligase substrates of hormone biosynthesis and perception and other substrates undergo phosphorylation either for destabilization or for stabilization. Thus phosphorylation has an effect on the ubiquitin dependent protein degradation pathway. *RV86-5* transgene induction also showed an effect on light signaling, as the line showed phototropic responsive behavior, many of the chloroplastic genes are down regulated and many of the mitochondrial genes are up-regulated. This contrast suggests an “energy crisis” situation. Furthermore, the data also suggest that cell wall is strengthening and lignin biosynthesis is activated.

3.1.10 Experimental validation of identified candidates from microarray analysis

From the comparative expression analysis between induced *sud2* and *RV86-5*, we identified several candidates, candidate 1 and 2 are suspected to be involved N-terminal protein myristoylation (Table 12). It is not known if K48 linked multi-ubiquitin chains have any role in myristoylation. The identified candidates are resulting from failure of K48 multi-ubiquitin

chain formation. It would be interesting to know if these two candidates have any role in regulating subcellular localization of components of ubiquitination machinery and thereby influence the cell death phenotype in ub K48R variant induced line and its rescue in *sud2*. Many of the identified candidates are unknown genes, with a putative role either as substrates or as E3-liages of ubiquitin dependent protein degradation pathway and influence cell death processes.

3.1.11 Biological importance of identified candidates

One could also raise the question why so many candidate mutations can prevent cell death phenotype in *RV86-5* background, where ubK48R is expressed. One possible reason would be that many substrates are supposed to be poly-ubiquitinated via ubiquitin Lys48 chains and designated for degradation. The identified candidates might belong to such a substrate pool and might accumulate in *RV86-5* because of failure of poly-ubiquitin chain formation caused by variant ubK48R. If these identified candidates are substrates, one more interesting question to pursue is to see whether these are targets of a single E3-ligase or of different E3-ligases.

The *sud2* suppressor screen candidate BTB-NPH3 belongs to a BTB family which forms a class of CRL that uses CUL3a/b protein as the scaffold. In contrast to CUL1, where adaptor and substrate recruiting functions are performed by separate proteins, in CUL3, BTBs are known to have adaptor and substrate recruiting functions in one protein (PINTARD *et al.* 2003). BTB proteins are clustered depending on their protein structure. Depending on corefold comparison BTB- NONPHOTOTROPIC HYPOCOTYL3 (NPH3), T1, Skp1 and Elongin C are clustered as closely related in comparison to other forms of BTBs.

The *Arabidopsis* genome contains 21 BTB-NPH3 proteins. These proteins seem to be plant specific BTBs. These BTBs are known to be involved in phototropism via signal transduction pathway by light activated Ser/Thr kinase NHP1. NPH3 protein has three distinct domains, known as N-terminal BTB (broad complex, tramtrack, bric a brac), centrally located NPH3 domain (Pfam, PF03000) and a C-terminal coiled-coil domain (PEDMALE and LISCUM 2007). It is known that BTB-NPH3 and ROOT PHOTOTROPISM2 (RPT2) form heterodimers (MOTCHOULSKI and LISCUM 1999; SAKAI *et al.* 2000) . The coiled-coil domain of NPH3 has been shown to interact with phototropin 1 (PHOT1) that contains light, oxygen and voltage sensing (LOV) domain, known to function as protein-protein interacting domain with a binding site for FMN (INADA *et al.* 2004).

Research work in plant light signaling has provided functional evidences of E3-ligases in regulation of plant growth and development. The human homologs of *Arabidopsis* COP1 and DET1 E3 ligases were shown to be crucial negative regulators of human tumor suppressors P⁵³ (WERTZ *et al.* 2004; YI and DENG 2005). One can assume that BTB-NPH3 might have played such an important role in the lethal phenotype in *RV86-5* background probably by targeting positive regulators of cell death in a light mediated way. The results from microarray and suppressor screen do demonstrate that BTB-NPH3 might have a very important role in changing chloroplast gene expression in induced *RV86-5*. This hypothesis is not sufficient to explain the explicit molecular mechanisms how the plants are switched from “light” to “lowlight” chloroplast gene expression condition even in the presence of light via protein turnover. In this context, taking the observed data into consideration, one can say that light signaling is affected in the *RV86-5* background because of disturbed poly-ubiquitination and may lead to cell death. The *sud2* background survived the lethal effect of ubK48R may be because a mutation in the BTB-NPH3 gene prevented interaction with negative regulators of cell death.

Some of the BTBs are involved in the rate limiting step of the ethylene biosynthetic and perception pathway. In ethylene biosynthesis, 1-aminocyclopropane-1-carboxylic acid (ACC) forms as precursor to ethylene. There are three types of such precursor forms, type1, type2 and type3. Type 1 and 2 are known to be short-lived when ethylene is not present. For type-1 ACSs, in the presence of ethylene their C-terminal motif is assumed to be phosphorylated, and thus leading to inhibition of recognition by an as yet undiscovered E3-ligase. In case of type-2 ACSs, which include ACS4, ACS5, and ACS9, ethylene may function in a similar way as in case of type-1 ACSs and block their recognition by the CUL3/BTB E3s ETHYLENE-OVERPRODUCING1 (ETO1), ETO1-LIKE1 (EOL1) or ETO1-LIKE2 (EOL2) (CHAE *et al.* 2003). Mutants in ETO1 show elevated levels of ACS5 which, results in ethylene overproduction, thus exhibiting a constitutive ethylene response phenotype (WANG *et al.* 2004). The two ACSs, type 1 and type 2 are stabilized in the presence of ethylene and function in a positive feed-back manner. For type-3 ACSs it is not known whether they are regulated by BTB type E3s.

It could be that the newly identified Phototropic-responsive NPH3 family BTB also functions in a similar way to the known ETO1, EOL1 or EOL2. If that is the case, one would expect that mutants in this BTB/POZ gene should result in constitutive ethylene response phenotype. This is still remains to be examined to see if the mutant in BTB/POZ causes a known ETO1, EOL1 or EOL2 type phenotype and whether it targets any ethylene biosynthetic enzyme/any

ACSs and leads to their stabilization. It is also interesting to see if the cell death phenotype in *ub K48R* coincides with elevated levels of ethylene, as it is noticed in senescence stage, which is a kind of slow form of cell death in plants. Figure 35 illustrates the hypothetical function of the BTB/POZ protein in ethylene biosynthesis.

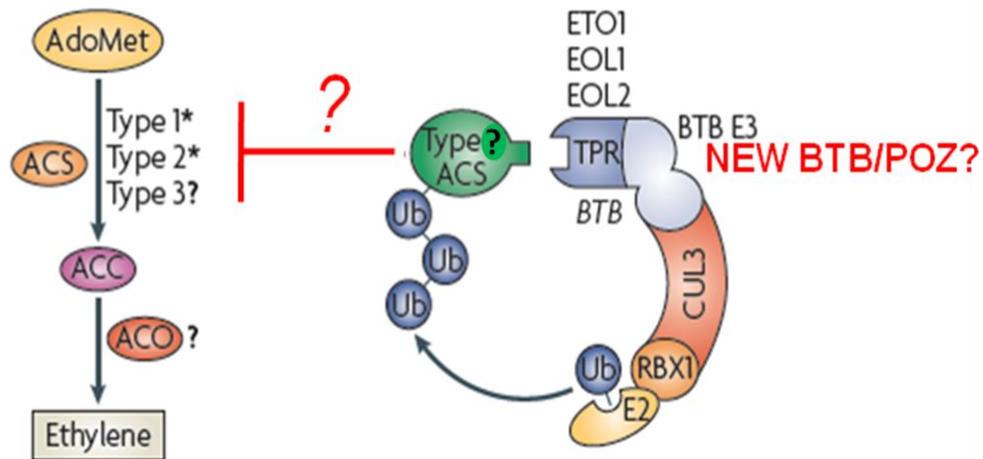


Figure 35 Control of ethylene synthesis by the UPS. Taken from Vierstra, 2009 and modified (VIERSTRA 2009). Ethylene is synthesized from *S*-adenosylmethionine (AdoMet) by the sequential action of 1-aminocyclopropane-1-carboxylic acid (ACC) synthases (ACSs) and ACC oxidases (ACOs). The stability of ACSs and possibly ACOs is under the control of the ubiquitin–26S proteasome system (UPS), especially for type 1 and type 2 ACS families, the degradation of which is blocked by ethylene. Turnover of type 2 ACSs requires ubiquitination by a family of bric-a-brac–tramtrack–broad complex (BTB) E3s assembled with the ETHYLENEOVERPRODUCING1 (ETO1), ETO1-LIKE 1 (EOL1) and EOL2 recognition proteins that have a tetratricopeptide repeat (TPR) motif. The identified BTB/POZ protein may function in the same capacity as the BTB ETO1, EOL1/2. ? Indicates unknown targets of novel BTB/POZ.

Another alternative hypothesis for BTB/POZ functional significance of the identified BTB protein is possible. There is one well known RING E3-ligase important to photomorphogenesis called COP1. The dark dependent translocation of COP1 into the nucleus targets turnover of the transcription factors LONG-HYPOCOTYL5 (HY5) and CONSTANS, an output of the plant circadian clock that controls the photoperiodic dependence of flowering. It could be that the newly identified BTB protein functions opposite to COP1, as it is predicted to be involved in phototropic-responsive. It would be interesting to see what kind of phenotype is observed in the double mutant with COP1. It is also interesting to see whether a plant with mutated AT3G44820 shows any changed response to ethylene. Nonetheless the result suggests that the newly found protein shows traits to already known BTB E3-ligases.

Another BTB E3-ligase, NONEXPRESSOR OF PR GENES1 (NPR1), seems to be a key player in plant responses to pathogens triggered by the hormone salicylic acid. It is suspected that NPR1 might be regulated by redox changes generated by pathogen attack. HR type cell death is linked to ROS signaling. One can also hypothesize such a role for the novel NPH3 protein as well, and the noticed cell death phenotype in the line expressing ubiquitin variant line that fails to make ub K48 linked poly-ubiquitin chains may be linked to this pathway.

The above hypothesized functions for the NPH3 protein need further experimental investigation in order to explicitly determine the functional importance of it, what kind of biological role it has as E3 ligase and through which pathway it is acting and leading to cell death phenotype in the ubiquitin variant background line. A connection between energy crisis cell wall strengthening and cell death in respect to involvement of identified candidates needs to be investigated more deeply. Nonetheless the results clearly helped to understand unknown connections between cell death and ubiquitin dependent protein degradation pathway.

3.2 Discussion – part 2

3.2.1 Importance of the N-end rule pathway in development

Degradation processes are crucial in any living organism. In this respect, proteolysis plays crucial roles in the regulation of a variety of cellular processes. The N-terminus of every protein can define its stability in the cell, yielding a rule known as N-end rule degradation pathway. Many substrates have been reported to be degraded via the N-end rule pathway. Among them are proteins from yeast, *Drosophila*, mammals and plants. These known substrates need first to be cleaved or otherwise processed by distinct factors, and the cleavage products display the N-terminus to be recognized for degradation. For example, in yeast the cohesin subunit SCC1 (Sister Chromatin Cohesion Protein 1) gets cleaved by separin, and the cleaved SCC1 product is degraded via the N-end rule pathway. The deregulation of this process inhibits sister chromatin separation, which causes higher percentage of lethality (RAO *et al.* 2001). In *Drosophila*, *Drosophila* inhibitor of apoptosis1 (DIAP1) gets cleaved by caspase and is further degraded by the N-end rule pathway. Its degradation is crucial for correct regulation of apoptosis (DITZEL *et al.* 2003). The mammalian regulator of G protein signaling (RGS) proteins RGS4, RGS5, and RGS16, are oxidized and arginylated at their N-terminus, which marks them as substrates for the N-end rule pathway (HU *et al.* 2005).

In plants, *Arabidopsis* protein RIN4 is a negative regulator of plant immunity, can be cleaved by a bacterial pathogenicity factor and as a result is further targeted for rapid degradation via the N-end rule pathway. In plants, also mutations in known components of the pathway lead to an aberrant phenotype, such as delayed leaf senescence caused by mutation in arginyl-tRNA: protein arginyltransferase (ATE1/ATE2) and germination defects caused by mutation in *PROTEOLYSIS6* (*PRT6*) were found to be linked with hypersensitive response to exogenously added abscisic acid (ABA), emphasizing the importance of this pathway in crucial developmental processes (HOLMAN *et al.* 2009).

However, compared to other organisms, in plants the components of this pathway and their roles are less well understood. This work contributes to the identification of novel components and provides new insights in this relatively new area in plant research.

3.2.2 Target specificities of N-end rule E3-ligases and their role in development

The N-end rule pathway is a ubiquitin-dependent protein degradation pathway. The substrates of this pathway are proteins with a bulky amino acid at the N-terminus such as Arg. In plants, *PROTEOLYSIS6* (*PRT6*) and *PROTEOLYSIS1* (*PRT1*) are responsible for recognizing proteins with N-terminal basic and hydrophobic residues, respectively (GARZON *et al.* 2007; POTUSCHAK *et al.* 1998). *PRT6* shows homology to yeast E3-ligase, UBR1, and also mammalian E3-ligase, UBR1, and targets substrates with same N-terminal residue for degradation by the 26S proteasome (TASAKI *et al.* 2005). In addition to a Ring finger domain, which is characteristic for E3-ligases, the yeast UBR1 contains two distinct recognition domains for targeting specific residues in the N-termini of N-end rule substrates: the UBR-domain, which binds basic N-termini, and the ClpS-domain, which binds hydrophobic residues (BARTEL *et al.* 1990). The plant homolog of yeast UBR1, *PRT6*, contains only one of these recognition domains, namely the UBR-domain, and thus can target basic residues, but not hydrophobic residues, as evident from the structure (GARZON *et al.* 2007). It can be assumed that the missing function (targeting hydrophobic residues) is performed by other E3-ligase candidates. Indeed, *PRT1*, which is structurally distinct from known E3-ligases, is known to be involved in turn-over of aromatic hydrophobic residues. However, neither *PRT6* nor *PRT1* are able to destabilize aliphatic hydrophobic residues, such as Leu, which is known to be an unstable residue in plants and other eukaryotes (GARZON *et al.* 2007; POTUSCHAK *et al.* 1998). These facts underpin the hypothesis of the existence of another E3-ligase that targets aliphatic hydrophobic residues.

3.2.3 Transgene-based screen led to the identification of the novel E3-ligase PRT8

As the known plant E3-ligases do not account for binding aliphatic hydrophobic residues, such as Leu, in order to find an E3-ligase that can target Leu residue, a mutant screen using Ubiquitin-GUS-fusion protein was performed. This technique allows to analyze a mutant population with regard to the stability of a test substrate containing a Leu residue at the N-terminus. A plant line that is mutated in a Leu-targeting E3-ligase retains the test protein due to inability to degrade it. The test protein contains a GUS-reporter-protein, the activity of which is visually detectable. This allows the analysis of a large population of mutants to identify the responsible E3-ligase. The conventional GUS-activity assay is performed in fixed plants. In contrast, the newly developed live tissue assay allows the mutant screening at the

early seedling stage and gives the possibility to keep the identified GUS-positive individual alive after GUS assay and the detected GUS-positive plant line thus can be further propagated. This is crucial for further characterization of the identified mutant. It is worth to mention that it not only speeds up the process but also reduces the amount of work tremendously. This method further can be explored for any kind of EMS screens of mutants that include a reporter construct as a tag to follow. One can also use other reporter tags instead of GUS, for example GFP, RFP, YFP or any fluorescence fusion constructs for screening of mutant, such tags can be visualized under microscope that allows fluorescence detection. One would observe the respective signal depending on stability of the expressed construct in the plant line genotype background for that given purpose. Thus this method serves as a very powerful tool for EMS mutant screening in a researcher friendly way.

Taking advantage of this newly developed technique, the EMS-mutagenized reporter-lines, which were harboring the GUS test substrate containing a Leu residue at its N-terminus, were analysed. Seeds were germinated on selection media supplemented with the required chemical to induce the expression of the Ubiquitin-GUS-fusion protein. This led to the identification of the *pri8* mutant, which showed a GUS-positive signal in the root, reflecting the fact that PRT8 could be a Leu-targeting E3-ligase. The scheme of this screen is depicted below (Fig 36).

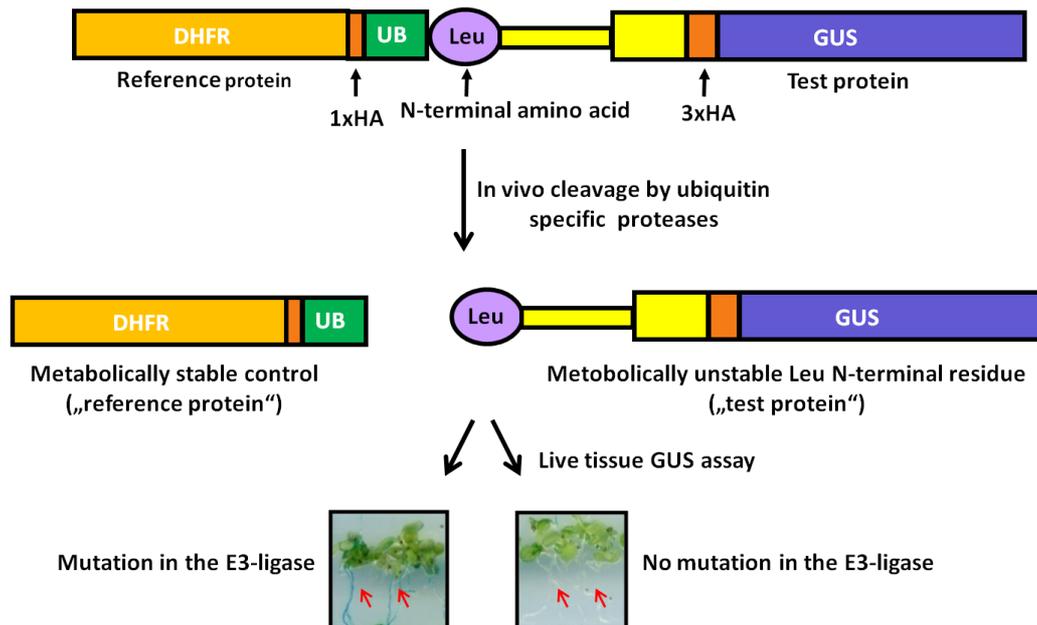


Figure 36 Ubiquitin fusion protein and live tissue GUS assay: a tool to decipher enzymatic functions of N-end rule E3-ligases. ORF of the construct showing GUS coloured in violet to the right side of ubiquitin (UB, coloured in green) and the DHFR reference protein (coloured in light orange to the left side of UB) containing a single HA tag (coloured in dark orange). X-denotes the amino acid (coloured in light purple) present at the N-terminus of the GUS test protein which contains a triple HA tag (coloured in dark orange). The flexible spacer is present between Leu and the triple HA tag (coloured in yellow). A cleavage site in the translated product recognized and cleaved by the ubiquitin processing enzyme is present at the C-terminus of the ubiquitin. The single downward pointing arrow indicates in vivo cleavage of the fusion protein by ubiquitin specific proteases and cleavage of translated product into two separate proteins, reference protein (DHFR) and test protein (GUS with Leu at the N-terminus). The 2 arrows pointing down in the lower half of the figure indicate the GUS assay results of tested plant lines. If the tested plant line has a mutation in the responsible E3 ligase, it shows a positive GUS assay; if not, GUS assay gives negative results, as exemplified in the Figure. The red arrows in the left panel points to GUS positive tested plant roots and in the right panel red arrows points to GUS – negative roots of the lines tested with Leu substrate.

To check whether a protein containing a Leu-residue at its N-terminus is degraded in a proteasome-dependent manner, the reporter line containing the construct expressing the test substrate was treated with a proteasome inhibitor. The GUS assay on these treated lines was positive, indicating that the test substrate is stabilized. This demonstrates that Leu-test proteins are substrates of the proteasome. In contrast to this, non treated progenitor plants have destabilized the test substrates due to the active proteasome.

It could be argued that the positive GUS-signal in the identified mutant plants might be due to mutations in other components of the degradation pathway, namely the proteasome itself. This hypothesis is based on the fact that EMS generates mutations randomly in the genome. However, if this was true, one would expect severe or even lethal phenotypes, as the proteasome is known to be involved in degrading a wide range of substrates. In fact, some of

the other GUS-positive plants that were detected could not survive, presumably due to such a mutation of the proteasome, which was detrimental to the plant. But the *prt8* mutant could survive, indicating that indeed not the proteasome, but maybe an E3-ligase is affected. This is further supported by the phenotypic similarity of the *prt8* mutant to known mutants in other components of the N-end rule pathway, such as the delayed leaf senescence mutant *dls*. The latter is affected in the t-RNA-Arg-transferase, an enzyme that attaches Arg to the secondary residues and converts them into primary substrates to be recognized by the downstream E3s (YOSHIDA *et al.* 2002).

To verify that the observed phenotype is indeed due to the *prt8* mutation and not any artifacts caused by EMS mutagenesis, this line has been outcrossed and checked for several generations. The mutation co-segregated with the noticed phenotype and stabilization. The originated descendants are ready to be analyzed for final confirmation.

It is still not known whether PRT8 is specific only for aliphatic hydrophobic residues. To address this, *prt8* mutant should be crossed with other reporter lines that harbor test substrates containing residues other than Leu, for example basic or aromatic residues.

3.2.4 Do plant homologs of mammalian UBR-domain proteins exhibit E3-ligase function?

A recent report revealed the family of UBR proteins in mammals consisting of 7 members, namely UBR1-7. *Arabidopsis* has two homologs of UBR-domain proteins, in addition to the previously mentioned PRT6. These are BIG and PRT7, which are closely related to the mammalian UBR4 and UBR7, respectively. The functional importance of BIG and PRT7 is not known yet (TASAKI *et al.* 2005).

To deduce whether BIG possesses E3-ligase activity, a T-DNA mutant line for this gene was obtained from the SALK collection and crossed with a reporter line expressing an Arg-GUS test protein. Arg was chosen because the UBR-domain is known to target basic residues. As control, the *big* mutant was also crossed to a Met-GUS reporter expressing line, where Met serves as a stabilizing residue and leads to a stable GUS-control. Preliminary results from GUS assays using these lines suggest that BIG has no E3-ligase function in turnover of the Arg-specific test protein, as Arg-GUS could be fully degraded in the *big* mutant background. This can be due to the fact that PRT6 is fulfilling this function. However, in the *prt6* mutant it was recently shown that Arg-GUS is not completely stabilized (GARZON *et al.* 2007). This suggests that another player is involved. To see whether BIG and PRT6 act together on Arg,

double mutants were analyzed. The results revealed no redundant function between PRT6 and BIG.

To analyze the function of PRT7, its mutant was isolated in frame of this work from the Koncz T-DNA library, and the same approach was applied as for BIG. The GUS-assays are in progress. Even though functions of BIG and PRT7 remain to be further examined, their phenotypes indicate that there might be a possible connection to the N-end rule pathway. The *big* mutant displays a late senescence phenotype, which is in agreement with known phenotypes of N-end rule pathway components. Interestingly, *prt7* mutant shows an early senescence phenotype, suggesting a negative role in the same pathway. It would be interesting to see whether the double mutant of *big prt7* displays any defect in degradation of test substrates with a basic residue at the N-terminus. Nonetheless the observed phenotypes of these homologs imply antagonistic functions in the N-end rule pathway.

3.2.5 Do plants process tertiary residues via deamidation?

In mammals, the N-terminal tertiary residues Asn and Gln undergo deamidation during the processing to become substrates that are directly recognized by E3-ligases. This de-amidation is performed by NTAN1 and NTAQ1 for deamidation of Asn and Gln, respectively, in mammals (GRIGORYEV *et al.* 1996; WANG *et al.* 2009). The *Arabidopsis* genome contains distant homologs of (these components named NTAN and NTAQ. It needed to be examined whether they also exhibit the deamidation function as shown for their mammalian counterparts. In this work therefore the mutants in these genes were analyzed. The *ntan* mutant was further crossed to a reporter line with Asn-GUS, and the *ntaq* mutant was crossed to a reporter line expressing Gln-GUS, in order to deduce a potential function in the deamidation from analysis of the test protein stability of the test-protein by GUS-assay or Western blot. These lines are ready to be analyzed for their enzymatic function. No aberrant phenotypes were observed in *ntaq* mutant plants. This might be the case for several different reasons. For example, maybe the amount of substrates that fails to be deamidated for final degradation in the *ntaq* mutant might be below the threshold to cause any obvious detrimental defect.

3.2.6 Does the N-end rule pathway play a role in NO signal perception?

Nitric oxide (NO) is generated in eukaryotes mainly by NO synthases. NO is known to be involved in posttranslational modification of proteins under a wide range of physiological conditions. NO can react with the certain amino acid residues, mainly Cys and Tyr, and leads to the formation of nitrosothiols. These modifications, in turn, change protein functions or activity. Since this protein modification by NO is reversible, it can serve as a component in signaling pathways.

The requirement of NO and O₂ for proteolysis through the N-end rule pathway of protein degradation was exemplified for regulator of G protein signaling (RGS) proteins RGS4, RGS5, and RGS16. RGS proteins carry at their N-terminus a Cys as the second residue following Met. The cleavage of Met exposes the Cys, which allows it to be modified through NO-mediated S-nitrosylation.

It was not clear whether also in plants NO, and O₂, thus S-nitrosylation is involved in the N-end rule pathway, even though a role of NO in other plant processes is already known. This hypothesis was investigated in this work by using a reporter line expressing a test substrate with an N-terminal Cys. When an NO-donor was added to these plants, the test-protein was completely destabilized, while without NO, degradation was at least in part non functional. To ensure that this effect is caused by Cys-specific modification, a control line expressing a Met-test substrate was used in parallel, where the stability of Met-substrate was affected neither with nor without NO. This means that the degradation of the Cys-substrate is specific for Cys-substrates and NO dependent. This kind of modification can be anticipated in plants as NO is present. One more close observation to point is that in non-treated condition Cys-GUS is partially destabilized, which could be explained by the presence of the internal NO levels, which might have contributed to this residual degradation. Further, in this work it could be shown that this NO-mediated degradation requires O₂, as in anoxic conditions NO alone was not sufficient to account for the complete degradation of the Cys-test substrate. These findings provide strong evidence that NO indeed plays a role in the N-end rule pathway in *Arabidopsis*. This could be a starting point to learn more about natural substrates of NO based signaling. Although levels of NO are not easy control, as there are ca 200 *Arabidopsis* ORFs expected to have Cys at as their second position, enhancing the possibility to find among them natural targets of this branch.

3.2.7 Conclusion and outlook

This thesis work made use of several artificial test-substrates as a tool and successfully identified novel components of the N-end rule pathway in *Arabidopsis*. Figure 37 summarizes the findings of this work.

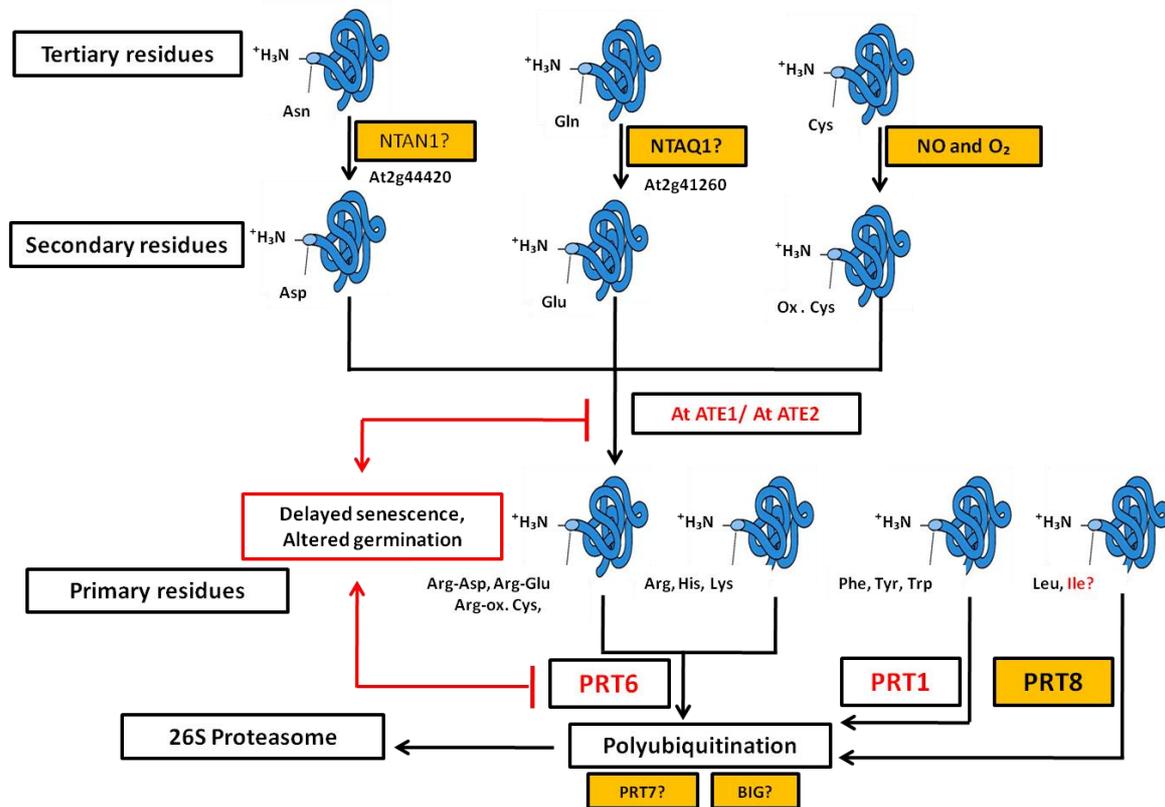


Figure 37 Overview of identified components of the N-end rule pathway in *Arabidopsis thaliana*. Components marked with orange boxes were identified in this work and the ones in red lettering were previously identified. Substrates are depicted in blue colour and with specific N-terminus. Components that target specific substrates are mentioned in the respective boxes. The final degraded protein is depicted below the proteasome box. Arrows indicate flow of modifications. NTAN1 and NTAQ1 are deamidating enzymes that convert tertiary residues into secondary residues. Cys is a tertiary residue converted into a secondary residue in the presence of NO and O₂. The secondary residues are converted into primary residues with the help of the enzymes ATE1 and ATE2.2 E3-ligases, PRT6 and PRT1 are already known E3 ligases with specificity for basic and aromatic hydrophobic residues, respectively. PRT8 is a novel E3 ligase identified in this work that shows specificity for Leu test substrates. The function of BIG as an E3 ligase is not clear. PRT7 was isolated via a T-DNA library screen, and is suspected to have a role as E3 ligase because of its strong early senescence phenotype.

These findings will serve as a starting point for further studies aiming to identify the natural targets of the pathway. One possibility would be to perform yeast-two hybrid assays using the identified gene products to find interacting partners. Another way to address this would be by inducible expression of the novel proteins fused to detectable tag in the background of the respective mutant plant line. This serves as a powerful tool to pull-down or co-immunoprecipitate in vivo interaction partners and/or substrates.

4 MATERIALS AND METHODS

4.1 Material

4.1.1 Chemicals, kits, antibodies

Chemicals and restriction enzymes used in this thesis work were purchased from following companies and their purity level is as certified by respective companies.

Amersham (Germany)

Duchefa (The Netherlands)

Invitrogen (Germany)

Roth (Germany)

Sigma-Aldrich (Germany)

NewEnglandBiolabs (Germany)

Roche (Germany)

Fermentas (Germany)

Difco Laboratories (USA)

Fermentas (Germany)

Qiagen (Germany)

Merck (Germany)

Kits

Nucleospin Plasmid® (Macherey-Nagel)

BioSprint 96 DNA Plant Kit (Qiagen)

PureYieldTMPlasmid Miniprep System (Promega)

RNeasy Mini Kit (Qiagen)

NucleoSpin® Extract II (Macherey-Nagel)

Wizard® SV Gel and PCR Clean-Up System (Promega)

Antibodies

Rat anti-HA antibody (Roche)

IRDye 800-conjugated goat anti-rat IgG antibody (Rockland)

4.1.2 Oligonucleotides, markers, enzymes

Oligonucleotides

Oligos were purchased from

Isogen Life Science (De Meern, The Netherlands)

Microsynth AG (Balgach, Switzerland)

DNA and Protein Markers

GeneRulerTM 100 bp DNA Ladder (Fermentas/NEB)

GeneRulerTM 1 kb DNA Ladder (Fermentas/NEB)

PageRulerTM Prestained Protein Ladder (Fermentas)

DNA polymerase enzymes

Phusion Hot Start High Fidelity DNA

polymerase (Thermo Fisher Scientific)

GoTaq[®] (Promega)

*LA Taq*TM (Takara)

4.1.3 Bacterial strains and binary vectors

Escherichia coli strain **XL1-Blue** (Stratagene)

This strain was used for cloning experiments.

Agrobacterium tumefaciens **C58C1 pCV2260**

This strain was used for cloning and in-planta transformation, it was obtained from D. Staiger (Eidgenössische Technische Hochschule, Zürich, Switzerland).

Binary vectors

pER8

This vector is a binary vector, has a β -estradiol-inducible promoter, and provides resistance against Spectinomycin for Agrobacteria and Hygromycin for plants.

p3

This vector is a binary vector, has a constitutive 35S promoter with three enhancer regions and has two selection markers to provide resistance against Kanamycin for Agrobacteria and Hygromycin for plants.

4.1.4 Plants

The following *Col/Ler* mutant lines, either generated in this work or received from mutant stock centers, were used for experimental purpose.

Table 16: *Arabidopsis thaliana* mutant genotypes that were used in this Thesis work

Genotype	Mutation type	Genetic Background	Source
<i>RV86-5</i>	ubK48R	Col	Schloegelhofer et al., 2006
<i>sud2</i>	EMS suppressor of ubK48R	Col	Schloegelhofer et al., 2006
<i>prt7</i>	T-DNA insertion in	Col	Isolated in the current work
<i>big</i>	T-DNA insertion	Col	SALK
<i>prt6</i>	T-DNA insertion	Col	SAIL
<i>ate1</i>	T-DNA insertion	Col	SALK
<i>ate2</i>	T-DNA insertion	Col	SALK
<i>ntaq</i>	T-DNA insertion	Col	GABI-KAT
<i>ntan1-1 to 1-7</i>	TILLING	<i>Ler</i>	Seattle TILLING project
<i>prt8</i>	EMS on test protein with L-GUS	Col	Isolated in the current work
<i>prt9</i>	EMS on test protein with L-GUS	Col	Isolated in the current work

Wild type Columbia (**Col-0**) and Landsberg *erecta* (**Ler**) were also used for control, transgenic line generation and crossing purpose.

4.1.5 Buffers and solutions

Seed sterilization solution

Ca(ClO)₂ 15 g was dissolved in 500 ml of dH₂O.

Plant DNA extraction buffer

Tris 200 mM (pH 7.5), EDTA 25 mM, NaCl 250 mM, and SDS 0.5 %.

Ferguson buffer for plant protein extraction

Tris-Cl 50 mM (pH 6.8), SDS 4% and β-Mercaptoethanol 10%.

Vitamin mix (500X)

Biotin 10 mg, thiamine 1 g, myo-inositol 5 g, nicotinic acid 50 mg, dissolved in final volume of 100 ml ddH₂O

GUS buffer

Na-Phosphate, pH 7 100 mM, EDTA 10 mM, K-Ferricyanide 0.5 mM, K-Ferrocyanide 0.5 mM, Triton- X 100 0.1 % and shortly before use X-Gluc 1 mM was added.

TAE buffer (50X)

Tris base 242 g, Na₂EDTA·2 H₂O 37.2 g and Glacial acetic acid 57.1 ml.

Buffers for SDS-PAGE and western blot

Sample loading buffer (LSB 2X)

This is prepared by using Glycerol 50 %, DTT 20 mM, SDS 2 %, Tris-Cl pH 6.8 125 mM and Bromophenolblue 0.003 %.

Electrophoresis buffer (5 x)

Tris 7.55 g, Glycine 36 g and SDS 2.5 g final volume was adjusted to 500 ml with dH₂O.

PBS 10 x

Na₂HPO₄ (anhydrous) 10.9 g, NaH₂PO₄ (anhydrous) 3.2 g and NaCl 90 g were added and final volume made up to 1 l with dH₂O, pH was adjusted to 7.2.

Transfer buffer

This buffer was prepared by using Glycine 190 mM, Tris 25 mM, Methanol 20 % and SDS 0.05 %.

Separating gel

Separating gels (for 12% PAGE) were prepared by using following chemicals and volumes.

Table 17: Components of separating gel for PAGE

Component	Volume
Acrylamide 30 %	2 ml
Tris 1.5 M (pH 8.8)	1.3 ml
SDS 10 %	50 µl
APS 10 %	50 µl
TEMED	4 µl
dH ₂ O	1.6 ml

Stacking gel

Stacking gels (for 12% PAGE) were prepared by using following chemicals and volumes.

Table 18: components of stacking gel

Component	Volume
Acrylamide 30 %	330µl
Tris 1M (pH 6.8)	250µl
SDS 10 %	20 µl
APS 10 %	20 µl
TEMED	2 µl
Bromophenol blue 0.003%	1.6 ml

4.1.6 Media

***Agrobacteria* YEB medium**

Beef extract 5 g, yeast extract 1 g, peptone 5 g, sucrose 5 g, with or without agar 15 g, depending on solid or liquid media, dissolved in 1 l dH₂O, adjusted pH to 7.2 with NaOH. Antibiotics final concentration of Rifampicin (100 µg/ml), Kanamycin (25 µg/ml), Spectinomycin (50 µg/ml), were added depending on the selection marker on the construct of interest. MgSO₄ 2 mM was either added or not after autoclaving media.

***Escherichia coli* LB medium**

Tryptone 10 g, yeast extract 5 g, NaCl 10 g, with or without agar 15 g, depending on solid or liquid media made up to 1 lit with dH₂O, adjusted pH to 7.0 with NaOH (Sambrook and Russel, 2001). Antibiotics at a final concentration of Chloramphenicol 50 µg/ml, Ampicillin 100 µg/ml, Kanamycin 25 µg/ml were added depending on the selection marker present on the construct of interest.

***Arabidopsis* MS medium**

MS salt 4.3 g, MES 0.5 g, sucrose 10 g, with or without agar 8 g, depending on solid or liquid media made up to 1 lit with dH₂O, adjusted pH to 5.7 with KOH. After autoclaving just before the use added vitamin mix (1X). For selection and induction hygromycin (25 µg/ml), β-estradiol (5µM), Dexamethasone (0.7µM) and Claforan (200 µg/ml) were either added or not depending on the type of transgene present or not.

4.2 Methods

4.2.1 Transformations

***E.coli* transformation**

To transform *E.coli* with the plasmid DNA of interest, 100 µl of XL1 blue cells were thawed on ice. 10 µl of plasmid DNA of interest was added and the mixture was incubated on ice for 30 min. The cells were shifted to 37°C for 2 min for heat shock and immediately placed back on to ice for 30 sec, after which 750 µl of LB medium were added. Cells were transferred to 37 °C at 750 rpm shaking condition for 1 h and cells were spun down at 3000 rpm for 3 min. Supernatant of 750 µl was discarded and the pellet was resuspended in the rest of 100 µl of

solution and spread on to solid LB medium with Ampicillin (100 µg/ml) or Kanamycin (25 µg/ml) or Spectinomycin (50 µg/ml), depending on the selection marker present on the plasmid DNA of interest. Plates were incubated overnight at 37°C for selection of positive colonies.

Agrobacterium transformation

Agrobacterium competent cells of 200 µl for each transformation were thawed on ice and 5-7 µl of each construct to be transformed (miniprep-based purified from *E.coli*) was added. These tubes were frozen in liquid nitrogen for 1 min and shifted to 37°C for 5 min. Immediately thereafter 1 ml of YEB medium with Rifampicin (50 µg/ml) was added, incubated at 28°C on a shaker with 750 rpm for 2 h. Cells were centrifuged at 10000 rpm, 1.1 ml of supernatant was discarded and the pellet was resuspended in the remaining 100 µl of solution. This was placed on solid YEB medium with Rifampicin (50 µg/ml), and either Kanamycin (25 µg/ml) or Spectinomycin (50 µg/ml) depending on the selection marker present in the transformed vector. These plates were incubated at 28°C for 2-3 days and colonies were selected for further examination of true transformants.

4.2.2 Plant methods

Plant genomic DNA extraction (Manual procedure)

30-50 mg of fresh plant material was deeply frozen in liquid nitrogen in a 1.5 ml reaction tube. This was used as starting material for DNA extraction. This plant material was mixed with small amount of fine quality quartz (sand), 200µl of DNA isolation buffer and thoroughly homogenized with an IKA-Mixer/glass pestle. The homogenized solution was centrifuged for 5 min at 14000 rpm. The supernatant was transferred into a fresh sterile 1.5 ml reaction tube and precipitated with 200 µl of Isopropanol by mixing for 5 min. The solution was centrifuged for 5 min at 14000 rpm at room temperature, the supernatant discarded and the pellet was washed with 500 µl of 70% EtOH. Further, the sample was centrifuged for 3 min at 14000 rpm; the supernatant was discarded and the pellet was dried either at RT or by speed vacuum pump. The pellet was resuspended in 60 µl of sterile dH₂ O, incubated for 5 min at 65°C and centrifuged for 2 min at 14000 rpm. Finally the supernatant was transferred into a fresh sterile 1.5 ml reaction tube and 1 to 2 µl of this DNA was either directly used for PCR reaction or stored at -20°C for further PCR based applications. Unless mentioned, all the steps were performed at room temperature.

Plant genomic DNA extraction (Bio sprint automated DNA extraction)

For genotyping of a very large fine mapping population, high quality DNA purification was performed by using *DNeasy® 96 Plant Kit* (Qiagen, Germany) and a fully automated *BioSprint® 96* machine. Fresh plant leaf material (2 to 3 week old) of 30-50 mg was used as starting material and collected into 2 ml safe-lock collection microtubes (each rack of 96 tubes) containing 300 µl lysis buffer. Before closing the tubes with caps, two 3 mm tungsten-carbide beads were added per tube to ensure proper disruption of leaf material. The racks were placed between the adapter plates of the tissue lyser adapter set 2 x 96 (Qiagen, Germany) and fixed firmly into the tissuelyser clamps and the samples were homogenized for 1 min at 30 Hz. The racks were centrifuged at 6000 x g for 5 min at room temperature. Further steps were performed according to "Purification of DNA using the Bio Sprint 96" protocol included in the Bio Sprint DNA plant handbook 03/2005. For the final DNA elution in 96-well microplate, 100 µl of dH₂ O per well was used. Eluted DNA was stored at -20°C till the analysis of the sample was completed.

Plant RNA Isolation

Two to three week old induced or un-induced seedlings were used as material for isolation of total RNA. Fresh seedlings of maximum of 100 mg were used as starting material for RNA isolation by using the *RNeasy Plant Mini Kit* (Quiagen). The Protocol for Isolation of Total RNA from Plant Cells mentioned in *RNeasy Mini Handbook 06/2001* was used. Digestion of DNA during RNA isolation was performed by on-membrane DNase digestion with RNase free DNase as mentioned in the protocol. To elute, 40 µl of RNase-free water was used and eluted RNA was used for downstream applications such as RT-PCR and expression-chip analysis.

Plant protein extraction

Transgenic induced or un-induced and wild type seedlings at an age of 2-3 weeks were selected for protein isolation. From these plants, 100 mg of fresh leaf material was collected in 1.5 ml reaction tubes and frozen in liquid nitrogen. A volume of 200 µl of prewarmed (5 min at 37°C in a heating block) Ferguson protein extraction buffer and a small spoon of fine quality quartz was added to the leaf material and the mixture was well homogenized with an IKA-Mixer/glass pestle. The samples were centrifuged at 14000 rpm for 1 min at room temperature and the supernatant was transferred into a fresh sterile tube and incubated in a heating block at 95°C for 10 min. The samples were inverted 2-3 times during incubation.

Samples were centrifuged for 10 min at 14000 rpm, the supernatant was transferred to a new reaction tube and immediately 1 vol of 2 X LSB was added. These samples were either directly used for Western blotting or stored at -20°C for further experimental use. In case of direct use, protein samples were incubated at 99°C for 5 min before being separated on polyacrylamide gel.

4.2.3 Plant growth

Plant growth on solid media

To grow *Arabidopsis* WT/mutant lines on solid media, solid MS with 1% sucrose supplemented with vitamin mix (1X) was used. Additional antibiotics and other inducing chemicals were either added or not depending on requirement. If they were added the final concentrations were, Hyg 25 µg/ml, Dex 0.7 µM, Mtx 100 µg/l and β-estradiol 5 µM. Sterilized seeds of required amount (20 to 50 in general) were placed on to the plates. These plates were placed at 4°C for three days. Plates were shifted to long day light condition. Seven to twelve days old seedlings were collected and used for further experimental procedures.

Plant growth in liquid media

To grow *Arabidopsis* WT/mutant lines in liquid media, 24- or 6-well plates were used. These plates were filled with 2 to 10 ml (depending on well size) liquid MS with 1% sucrose supplemented with vitamin mix (1X). Additional antibiotics and other inducing chemicals were either added or not depending on requirement. If they were added the final concentrations were, Hyg 25 µg/ml, Dex 0.7 µM, Mtx 100 µg/lit, SNP 100 mM, MG 132 100 µM and β-estradiol 5 µM. Sterilized seeds of required amount (20 to 50 in general) were placed into 24- or 6-well plates with media. These plates were placed under long day light condition either on a slowly rotating or stable platforms. Seven to twelve day old seedlings were collected and used for further experimental procedures.

Plant growth conditions

Sterilized seeds were placed on selective or non-selective solid MS media with 1% sucrose containing plates. These plates were sealed with parfilm and kept at 4°C for 3 days to vernalize. Afterwards, the plates were placed under light condition for 1 to 2 weeks. The seedlings were shifted onto soil and grown to maturity in long day conditions (16 h of light and 8 h of darkness). In case when plants were supposed to be used for transformation

experiments, they were grown in short day conditions (8 h of light and 16 h of dark) on soil from seedling stage. After two months of growth they were shifted to long day to induce flowering. When they started to bolt, the main shoot was cut in order to provoke more branches and thus to increase number of flowers to increase potential number of transformants. After these plants were transformed they were grown in long day condition till they reached seed maturity.

4.2.4 Plant genetic methods

Crossing of plants

Crossing was done to transfer the reporter constructs from one genetic background into another, to examine possible functional effects of different mutant alleles on each other's phenotypes (genetic interaction), to make allelism tests among EMS mutants and to perform complementation analysis. Two parental lines with suitable stage of inflorescence were selected. From these branches siliques, flowers and side branches were removed except the selected 2-3 flower buds that were to be crossed. From parental female flowers to be crossed, the petals, sepals, and all the anthers except stigma were removed (it was ensured that no anther with pollen came in contact with the pistil while removing them). The male parent's selected flowers that were recently opened and showed visible pollen were squeezed gently at the base with forceps, sepals petals and pistil were removed, and anthers with pollen were carefully taken out and brushed gently on the surface of the stigma. At the end of the crossing, pollen was noticed as yellow dust on the stigma. The crossed flower was covered with a small plastic bag in order to avoid other undesired pollen contact. The crossed flowers were labeled to allow unequivocal identification. After 2-3 weeks, seeds from the cross were collected.

Plant genotyping

Genomic DNA purified from two to three weeks old mutant and wild type plant lines was used for genotyping. The T-DNA insertion and WT alleles were detected by PCR-based amplification of the gene products using genotype allele specific forward and reverse oligonucleotides. In case of TILLING lines, the PCR products were digested with respective restriction enzymes before being analyzed. The status of genotype was determined by presence or absence of allele PCR product.

Table 19. Primers for genotyping plants with different genetic backgrounds

Name of the gene	TAIR number	Genotype	PCR primers
UBR7	At4g23860	Mutant	CTG GGA ATG GCG AAA TCA AGG CAT C GAC TCC TAC AAA ACC AAC AAC GAA TCA AGT CTT
		WT	CTC CAT CAA TAA CCT GGT AAT GGT CCG ATT GAC TCC TAC AAA ACC AAC AAC GAA TCA AGT CTT
BIG	At3g02260	Mutant	TGG TTC ACG TAG TGG GCC ATC G AGC TGC CAC ACA TGC CTG GAC ATT
		WT	GAA ATG GCA GAT GAC TTG GCG AAT AGC TGC CAC ACA TGC CTG GAC ATT
PRT6	At5g02300	Mutant	GCC TTT TCA GAA ATG GAT AAA TAG CCT TGCTCC GTT TCT TGT TCT GGG GAG GAT GGT TT
		WT	AGG ACA ATA GGT ACA TAC TCA TTT GTT GTT TCT TGT TCT GGG GAG GAT GGT TT
NTAQ	At2g41760	Mutant	GGG CTA CAC TGA ATT GGT AGC TC GCT TTT CGC AGA GTA CCA GAG GTA ATC
		WT	AAG ACA TTG GAA TGC TAA GGA AGC TT GCT TTT CGC AGA GTA CCA GAG GTA ATC
NTAN	At2g44420	Mut/WT	CTT GGG CAC AAT ACC AAG TTG GAT TTA ATA AAG AAA CAT GGT TAC GCT GAT T

Phenotyping of fine mapping population

Fine mapping population individuals (1237) were germinated under induced conditions on solid MS media with 1% sucrose and supplemented with 0.7 μ M Dexamethasone. After 2-3 weeks, seedlings were scored for survivors (homozygous mutant), non-survivors (homozygous wild type) and semi-survivors (heterozygous).

Marker-based genotyping

SNP and dCAPS markers were amplified by PCR to identify polymorphisms between Col-0 and *Ler*. These generated PCR products were either directly analyzed or subjected to cleavage by a respective restriction enzyme before being analyzed. In both cases analysis was performed on normal (0.8) or high (2.5-3) percentage agarose gels depending on fragment size difference between polymorphic products. The result of these markers analyses helped to decide whether there was homozygosity or hemi/heterozygosity (present) for that given marker locus on that specific chromosome for that given recombinant line.

4.2.5 Model substrate generation and stability assays

Ubiquitin fusion protein construct generation

A fragment consisting of DHFR-X-GUS, fusion protein construct ORF was excised at *XhoI* and *XbaI* sites from vectors pUPR. These purified fragments were inserted into inducible plant binary vector pER8 or p3, keeping ORF expression under control of β -estradiol inducible or CaMV 35S promoter. These pER8-X-GUS and p3-X-GUS vector were used for transformation of *Arabidopsis* plants.

GUS assay

Plant seedlings to be examined for GUS activity were grown on selection or non selection media for 2-3 weeks. Three to five seedlings from each plant line were placed into a 1.5 ml reaction tube with 1 ml of freshly prepared GUS buffer. These tubes were subsequently infiltrated under vacuum for 5-10 min and incubated overnight at 37°C for. After 24 h, GUS buffer was removed from tubes and seedlings were incubated in 1 ml of washing solution (70% EtOH) overnight at 37°C. The next day, the samples were washed 2-3 times with 70% and 75% EtOH to ensure removal of pigments of the tissue. The results were analyzed visually for staining of tissue.

Live tissue GUS assay

This method was newly developed for identifying candidate mutants of the EMS screen. Plant lines to be examined for live tissue GUS assay were germinated on selection MS solid medium square shaped plates in 4 to 5 rows by maintaining equal distance between each line. Seedlings were grown by placing plates vertically, allowing roots to be exposed on the surface of the media. Freshly prepared GUS buffer (lacking Triton-X 100) of 3 to 5 ml was gently sprayed onto the plate by using a solution sprayer and plates were immediately sealed with Parafilm in order to avoid any kind of contamination on plates and to the plants. All the steps were performed under sterile conditions only. These freshly sprayed plates were incubated overnight at room temperature. GUS positive lines were directly selected and transferred onto soil and allowed to grow to maturity. The next generation was used in further experiments.

4.2.6 Plant treatments

The Proteasome inhibition treatment

To treat *Arabidopsis* seedlings with the proteasome inhibitor, MG132, seedlings were grown for up to 2 weeks in liquid MS media supplemented with Hyg and β -estradiol. These seedlings were transferred into fresh liquid MS media supplemented with Hyg and β -estradiol and 100 μ M MG132 for treatment, or without MG132 but with equal volume of DMSO for control. After 15 h, treated and control seedlings were further examined.

Nitric oxide (NO) treatment

For treatment of *Arabidopsis* seedlings with NO, donor sodium nitroferricyanide (III) dihydrate (SNP), seedlings were grown for up to 2 weeks in liquid medium with Hyg and β -estradiol. These seedlings were transferred into fresh medium supplemented with Hyg and β -estradiol and 100 mM SNP for treatment, or without SNP but an equal volume of DMSO for controls. After 15 h, seedlings were used for further analysis.

Plant anoxic treatment

To expose *Arabidopsis* seedlings to anoxic conditions, seedlings were grown up to 2 weeks in liquid MS medium with Hyg and β -estradiol. Fresh MS medium supplemented with Hyg, β -estradiol and SNP, MG132 were added together or separately (depending on designed experiment) was degassed for at least 4 h in vacuum. Two week old seedlings were transferred into 1.5 ml reaction tubes and submerged totally by filling tubes with the degassed medium and immediately closing the lids and sealing them air-tight with parafilm. These 1.5 ml tubes were kept in darkness for 15 h. For control experiment all the steps were performed in the similar way but with only one change that is in place of degassed liquid medium, normal not degassed liquid MS medium was used.

4.2.7 Purification methods

Plasmid DNA purification

An overnight culture of 5 ml from *E.coli* or *Agrobacterium* was used for purification of the plasmid DNA. Nucleospin Plasmid® (Macherey-Nagel) or A Wizard®*plus* SV Minipreps DNA Purification System (Promega) and centrifugation protocol from users guide was followed and final elution was performed in 40 to 50 μ l of nuclease-free water. Eluted DNA

was used for enzymatic digestion to confirm presence of insert construct of interest and for confirmation by sequencing. In case of later experimental use, the DNA was stored at -20°C.

PCR product purification

PCR amplified products of desired samples were purified from agarose gel by using the Wizard®SV Gel and PCR Clean-Up System from Promega. During purification, instructions provided by the manufacturer were followed. Final elution was done in 40 µl of either sterile dH₂O or Nuclease-free water provided by the PCR clean-up system. Purified fragments were subjected to sequencing and use in downstream applications.

4.2.8 Standard enzymatic reactions

Restriction reaction

For restriction of vectors of interest (pER, p3, pUPR) the following reaction composition was used.

Table 20: components of a standard restriction mix

Volume	Component
2.00 µl	Vector of interest
0.25 µl	Enzyme I of interest
0.25 µl	Enzyme II of interest
2.00 µl	Compatible buffer
10.50 µl	dH ₂ O
Total Vol 15.00 µl	

The volumes of standard reaction mix components were modified in order to obtain desired results (in some cases). The reaction mix was incubated for overnight at temperature suitable to enzymes.

Ligation reaction

For ligation insert/construct of interest into vector of interest (pER, p3, pUPR) the following ligation mix was prepared. For every ligation reaction a control reaction was prepared by omitting the insert component.

Table 21: Standard ligation reaction components

Volume	Component
7.00 µl	insert of interest
3.00 µl	Vector of interest
1.00 µl	DNA ligase T4
2.00 µl	10x ligase buffer
7.00 µl	dH ₂ O
Total Vol 20.00 µl	

Table 22: Standard control ligation reaction components

Volume	Component
-----	-----
3.00 µl	Vector of interest
1.00 µl	DNA ligase T4
2.00 µl	10x ligase buffer
14.00 µl	dH ₂ O
Total Vol 20.00 µl	

These basic ligation reaction component volumes were modified to obtain desired results (in some cases). Ligation reaction was performed in general by incubating at 16°C overnight.

Digestion reaction

For mini-prep DNA or PCR products and PCR-based amplified marker products were digested using the following reaction mix.

Table 23: A Typical DNA digestion reaction components

Volume	Component
3.00 µl	DNA of interest
0.40 µl	Enzyme of interest
2.00 µl	Compatible buffer
14.60 µl	dH ₂ O
Total Vol 20.00 µl	

Minor changes to the volumes of components were made when desired results were not obtained directly (in some cases). The digestion reaction mix was incubated generally overnight at a temperature suitable for the restriction enzyme used in the reaction.

4.2.9 Nucleic acids synthesis and quantification

cDNA synthesis

Total RNA extracted from induced and un-induced transgenic lines was used for synthesis of complementary DNA (cDNA). The SuperScriptTM II Reverse Transcriptase (Invitrogen) enzyme was used in this experiment. For each sample, 5 µl of RNA was added to 6.4 µl of dH₂O and kept at 65°C for 5 min. The reaction tubes were placed immediately on ice. The reverse transcription enzyme mixture of 8.6 µl (RT buffer (5 x) 4 µl, RNasin 0.6 µl, dNTP (10 mM) 2 µl, oligo dT (100 µM) 1 µl and reverse transcriptase 1 µl) was added. This final 20 µl reaction volume containing micro tubes were incubated at 42°C for 1 h and from this cDNA 1-2 µl were used as template for subsequent PCR reactions.

Quantification of RNA

Purified total RNA from plant lines of induced or un-induced origin was quantified using nano-drop machine.

Quantification of DNA

Purified DNA was quantified using nano-drop machine.

4.2.10 Databases and Bioinformatics tools

Bio informatics tools, databases, and internet based resources

The following web sites were used for gene sequence analysis, sequence alignment, BLAST search, SNP and dCAPS search and designing.

<http://www.Arabidopsis.org>

<http://www.tigr.org>

<http://www.ncbi.nlm.nih.gov>

<http://helix.wustl.edu/dcaps/dcaps.html>

<http://www.Arabidopsis.org/Cereon/>

4.2.11 Other plant-related methods

Seed sterilization

For each plant line, seeds of required amount were kept in 1 ml sterilization solution in a 1.5 ml eppendorf tube and kept on a shaker for 15 min at room temperature. They were centrifuged briefly and the supernatant was discarded. Seeds were washed 2 times with 1 ml sterile dH₂O. These seeds were dried under the laminar hood. These sterilized seeds were used for growing plants in non-selective or selective MS solid or liquid media.

Plant transformation

Three days prior to the plant transformation, sequence confirmed single colonies of *Agrobacterium* containing the construct with the insert to be transferred into plants were used to inoculate 20 ml YEB medium with Rifampicin (50 µg/ml), and either Kanamycin (25 µg/ml) or Spectinomycin (50 µg/ml) depending on the selection marker present on the transformed vector. These cultures were incubated on a shaker with 200 rpm at 28°C for 2 days. 2 ml of the culture was transferred into 100 ml fresh YEB medium with respective antibiotics and incubation was continued under the same conditions overnight. This culture was centrifuged at 5000 rpm for 15 min. The supernatant was discarded and the pellet was resuspended in 200 ml dH₂O with 5% sucrose and 0.05% silwet 77. *Arabidopsis* wild type plants with unopened buds and partially open flowers were dipped for 30 sec to 1 min. Dipping was repeated 1 more time with a 30 min gap and the plants were placed into well-covered trays. They were shifted into long day growth condition and allowed to grow up to the stage of ripe silique. Transformants were selected by checking for resistance to respective selection on solid MS medium supplemented with 45 µg/ml claforan.

Western blotting

Leaf protein extracts of equal amount were separated on 12% polyacrylamide gels. To detect and quantify protein of interest, separated proteins were transferred onto BA85 nitrocellulose membranes by blotting for 1 h. The membrane was further processed by using Odyssey blocking buffer according to the protocol of Li-Cor Biosciences. The blot was incubated at 4°C overnight with rat anti-HA antibody (Roche). The blot was washed with buffer containing 1xPBS and 0.1% Tween and incubated with IRdye 800-conjugated goat anti-rat IgG secondary antibody (Rockland) for 1 hr at room temperature. The blot was dried and protected from light. The Odyssey Infrared Imager (Li-Cor) was used for protein detection

and quantification. The protein bands were excited at wavelength of 780 nm and emission was recorded at 820 nm. For background signal, emission at a void point was measured.

DNA sequencing

The required DNA samples of plasmids, mutant and wild type *Arabidopsis thaliana* and PCR products were sequenced using services of Automated DNA Isolation and Sequencing service (ADIS) at Max Planck Institute for Plant Breeding Research in Cologne, Germany or AGOWA genomix of Berlin, Germany and LGC genomics Berlin, Germany.

For Solexa sequencing DNA was sent to GATC of Konstanz, Germany.

Sample preparation for Gene expression by Microarray

To compare differential expression of transcripts in different genetic backgrounds and conditions, RNA samples were extracted from two week old seedlings either 24h treated or untreated. These samples were subjected to Microarray based detection of differential expression of genes.

T-DNA library screen

To identify mutations in the gene of interest, in this case PRT7, the Koncz collection (Max Planck Institute for Plant Breeding Research in Cologne, Germany) of T-DNA tagged *Arabidopsis thaliana* plants, a library consisting of 39700 individuals, was screened. By using gene specific and T-DNA border specific oligonucleotide combinations, a PCR-based screen was performed.

EMS seed mutagenesis of Arabidopsis

The reporter line seeds to be mutagenised were first generated in large numbers. For seed mutagenesis, 0.7 g of sterilized seeds were taken in to a 50 ml Falcon tube and suspended in 50 ml of dH₂ O. Then 150 µl of EMS was carefully added. In order for every possible seed to get exposed to EMS, the falcon tube was kept on a rotator. The seeds were treated with EMS for 12 h and then thoroughly washed with water 6-8 times and at every wash seeds were transferred into a new Falcon tube and the old ones were submerged into NaOH containing solution. After the final wash, seeds were submerged in 0.1% agar to get homogenized distribution of seeds, and these seeds were spread onto soil to grow to maturity and to collect F₂ seeds for downstream experimental purpose. During mutagenesis, care was taken to avoid any kind of EMS contact to the body.

5 ABBREVIATIONS

ABA- abscisic acid

APC/C- Anaphase Promoting Complex/Cyclosome

ARM- Armadillo repeats

ASK1-*Arabidopsis* S-phase Kinase associated Protein1

ATE1- Arg t-RNA protein transferase

BEE2-BR enhanced expression 2

BTB/POZ -Bric a brac, Tramtrack and Broad complex/Pox virus and Zinc finger

cM-centi Morgan

COP1- Constitutively Photomorphogenic-1

CP-core particle

CRLs- Cullin Ring Ligases

dCAPS-derived cleaved amplified polymorphic sequence

Ddi1- DNA-damage inducible 1

DET1- De-Etiolated-1

Dex-Dexamethasone

DHFR-dihydroxyfolate reductase

DIAP1- *Drosophila* inhibitor of apoptosis1

DUBs- De-ubiquitinating enzymes

E1- activating enzyme

E2- conjugating enzyme

E3- ligase

EMS-Ethyle methanesulfonate

FRE 1-far-red-light-insensitive1

GA- gibberellic acid

GUS- β -Glucuronidase

HDM-Hyg, Dex, and Mtx

HECT E3s-Homology to E6AP C-Terminus
HR -hypersensitive response
HY5-ELONGATED HYPOCOTYL5
Hyg -Hygromycin
JA-jasmonic acid
MG132-Proteosome inhibitor
Mtx-Methotrexate
NERP -N-end rule pathway
NO -nitric oxide
NO-Nitric oxide
NPH3 – NONPHOTOTROPIC HYPOCOTYL3
NTAN-N-terminal Asn deamidase
NTAQ-N-terminal Gln deamidase
PAL- phenylalanine-ammonia lyase
PCD -Programmed cell death
PHOT-Phototropin
PRT1- PROTEOLYSIS1
PRT6- PROTEOLYSIS6
PRT6-proteolysis6
PRT7-Proteolysis 7
PRT8-Proteolysis8
PRT9-Proteolysis9
PUBS- Plant U-box
RBX1- RING-Box 1
RGS- regulator of G protein signaling
RING E3s/ U-box -Really Interesting New Gene
ROS- reactive oxygen species

RP-regulatory particles

RT-PCR-Reverse transcriptase PCR

R-transferase- arginyl-tRNA-protein transferase

RV86-5-ubiquitin variant (ubK48R) expressing line in Col-0 background

SA-salicylic acid

SNP-Single nucleotide polymorphism

spl11- *Spotted leaf11*

SSC1- Sister Chromatin Cohesion Protein 1

Sud2-suppressor of ubiquitin variant induced cell death

TPR- tetratricopeptide repeat

ub- Ubiquitin

UBA1-ubiquitin activating enzyme 1

UBC- ubiquitin conjugation domain

UbdHFR- ubiquitin dihydroxyfolate reductase

ubK48R-ubiquitin Lys-48 replaced by Arg

UEVs- Ubiquitin-conjugating E2 enzyme variant

UPS- ubiquitin 26S proteasome pathway

6 APPENDIX

Appendix 1 Downregulated genes in induced compared to uninduced condition in *RV86-5*.

	AGI	FC (down-regulation)	Short description
1	AT3G19850	7,39	light responsive endomembrane NPH3 BTB/POZ domains
2	AT1G79110	7,21	RING Zn finger best match S-ribonuclease S binding
3	AT5G47610	4,28	RING protein
4	AT1G76410	3,75	RING U-box superfamily protein
5	AT5G64330	3,69	non-phototropic hypocotyl 3 NPH3 blue light sensor BTB/POZ like fold
6	AT4G04940	3,51	WD40 repeat protein component of CUL4 ligase complex rRNA processing
7	AT3G10530	3,31	WD40 repeat protein component of CUL4 ligase complex
8	AT1G23390	2,98	Kelch repeat F-box protein
9	AT4G28450	2,88	WD40 repeat DWD DDB1 binding protein Cul4 ligase ?
10	AT1G15440	2,87	periodic tryptophan protein 2 PWP2 component of CUL4 complex
11	AT2G34260	2,85	WD-40 repeat domain Cul4 ligase subunit
12	AT1G71850	2,84	ub carboxyl-terminal hydrolase family protein
13	AT5G22920	2,85	RING type Zn finger protein
14	AT5G14050	2,73	WD40 repeat protein of CUL4 ligase complex
15	AT3G58520	2,68	ubiquitin carboxyl terminal hydrolase family protein
16	AT1G49230	2,66	RING Ubox superfamily
17	AT2G44130	2,62	Kelch repeat F-box protein
18	AT2G18290	2,59	anaphase promoting complex subunit APC10
19	AT3G61060	2,61	phloem protein 2-A13 PP2-A13 wound responsive F box
20	AT4G05410	2,58	yaozhe YAO WD40 repeat protein ?APC/Cyclosome CUL4 complex component ? nucleolar ?
21	AT4G37610	2,55	BTB and TAZ domain protein 5 BT5 cold, chitin, SA, auxin induced
22	AT5G15550	2,48	TPR repeat domain CUL4 subunit ? heterotrimeric G protein ?
23	AT4G13100	2,39	RING U box superfamily calmodulin binding
24	AT1G10230	2,39	SKP1-like 18 BTB/POZ domain ubiquitin ligase subunit
25	AT5G02760	10,33	PP2C phosphatase
26	AT1G14700	7,12	purple acid phosphatase 3 PAP3 S/T vacuolar !
27	AT3G46280	5,74	protein kinase related endomembrane
28	AT1G16260	4,84	wall-associated protein kinase EGF-like Ca binding
29	AT3G22750	4,69	protein kinase of plasma membrane
30	AT2G45340	4,38	LRR receptor kinase S/T
31	AT5G48540	4,32	receptor-like protein kinase related response to kallirein
32	AT1G51800	4,08	LRR protein kinase of endomembrane system
33	AT5G63410	3,88	endomembrane LRR S/T protein kinase
34	AT2G16430	3,86	purple acid phosphatase 10 gene PAP10 of cell wall
35	AT5G11410	3,87	protein kinase
36	AT1G69730	3,40	wall-associated protein kinase EGF-like Ca binding
37	AT4G11460	3,31	cys-rich receptor like kinase 30 CRK30 lectin domain endomembrane location
38	AT2G23200	3,12	protein kinase
39	AT4G08850	2,83	LRR containing receptor like S/T kinase
40	AT1G51850	2,91	LRR containing receptor like S/T kinase
41	AT2G17820	2,80	AHK1 histidine kinase osmotic stress sensor
42	AT5G62710	2,77	LRR S/T protein kinase endomembrane system
43	AT5G51560	2,81	LRR S/T protein kinase endomembrane system
44	AT5G45800	2,74	maternal effect embryo arrest 62 MEE62 LRR S/T protein kinase
45	AT4G23130	2,74	cys-rich receptor like proein kinase 6 RLK6 SA induced
46	AT1G72180	2,63	LRR containing receptor like S/T kinase
47	AT3G08920	2,62	rhodanese / cell cycle control phosphatase superfamily of chloroplast
48	AT4G14930	2,59	survival protein SurE like phosphatase nucleotidase unknown function
49	AT1G51890	2,59	LRR S/T protein kinase endomembrane system
50	AT5G63140	2,58	purple acid phosphatase 29 PAP29 S/T phosphatase activity endomembrane
51	AT2G26330	2,53	quantitaitve resistance to plectosphaerella 1 QRP1 erecta ER LRR membrane protein kinase
52	AT1G51790	2,50	LRR S/T kinase
53	AT5G01540	2,45	lectin receptor kinase a4.1 LecRKA4.1 neg regulation of ABA response in seed germination S/T protein kinase
54	AT1G67820	2,41	S/T phosphatase PP2C family

55	AT1G07570	2,39	APK1A protein kinase
56	AT1G33260	2,39	S/T protein kinase
57	AT5G39210	16,47	chlororespiratory reduction 7 CRR7 NAD(P)H dehydrogenase complex chloroplast
58	AT5G13730	9,67	sigma factor 4 SIG4 chloroplast
59	AT1G70760	9,64	chlororespiratory reduction 23 CRR23 NAD(P)H dehydrogenase complex (plastoquinone) chloroplast
60	AT2G28900	9,26	outer plastid envelope protein 16-L OEP16-1 import of protochlorophyllide oxidoreductase A JA wounding cold responsive
61	AT1G10522	9,13	unknown chloroplast protein
62	AT3G13470	9,00	chloroplast chaperone chaperonin cpn60/TCP-1 family
63	AT4G39210	8,71	glucose-1-phosphate adenyltransferase activity ADP-glucose pyrophosphorylase large subunit APL3 starch biosynthesis
64	AT3G15357	8,51	unknown chloroplast protein
65	AT4G17090	8,47	beta amylase 8 BMX8 of chloroplast starch catabolism ? salt cold responsive
66	AT3G16250	8,06	NAD(P)H dehydrogenase complex (plastoquinone) chloroplast
67	AT5G58260	6,75	chloroplast NAD(P)H plastoquinone dehydrogenase complex subunit
68	AT5G36120	6,17	chloroplast Cyt b6f complex assembly
69	AT1G69200	5,97	fructokinase-like 2 FKL2 carbohydrate kinase fructose biosynthesis chloroplast
70	AT3G16470	5,67	wound salt cold stress JA responsive jasmonate responsive 1 JR1 mannose binding lectin superfamily vacuole ? chloroplast ?
71	AT1G19150	5,52	chlorophyll A/B binding protein LHCA6
72	AT2G04030	5,34	HSP90.5 HSP88.1 chloroplast-targeted
73	AT2G29180	5,21	unknown chloroplast protein
74	AT5G49910	5,02	heat shock protein 70 HSC70-7 chloroplast cpHsc70-2 (responds to Cd)
75	AT5G55220	5,02	chloroplast peptidyl prolyl cis trans isomerase
76	AT5G08610	4,94	chloroplast ATP-dependent helicase DEAD/DEAH box
77	AT3G52720	5,00	alpha carbonic anhydrase 1 ACA1 chloroplast stroma
78	AT5G19750	4,86	chloroplast envelope peroxisomal membrane protein
79	AT5G44190	4,95	Golden2-like 2 GLK2 TF regulator of chloroplast development (photosynthetic apparatus) homeodomain
80	AT3G17170	4,75	ribosomal protein S6 / EF1B family protein chloroplast ribosome regulator of fatty acid composition 3 RFC3
81	AT2G41950	4,76	unknown chloroplast protein
82	AT5G02710	4,73	unknown chloroplast protein
83	AT1G49975	4,69	chloroplast photosystem I component
84	AT1G14150	4,69	psbQ-like 2 PQL2 NADPH dehydrogenase complex thylakoid lumen Chloroplast
85	AT4G37470	4,56	alpha beta hydrolase fold superfamily de-etiolation (chloroplast-related)
86	AT1G26230	4,56	chloroplast chaperone chaperonin cpn60/TCP-1 family
87	AT1G03630	4,56	protochlorophyllide reductase chloroplast
88	AT2G02740	4,44	whirly 3 WHY3 putative chloroplast TF
89	AT4G34290	4,43	SWIB/MDM2 domain superfamily protein chloroplast
90	AT1G48570	4,41	chloroplast Zn finger (Ran binding) protein
91	AT1G14345	4,43	chloroplast oxidoreductase
92	AT2G42690	4,36	triglyceride lipase of chloroplast UV-B response
93	AT2G17972	4,23	chloroplast thylakoid membrane protein unknown
94	AT3G06730	4,26	thioredoxin z of chloroplast
95	AT2G01590	4,29	chlororespiratory reduction 3 CCR3 likely component of NADPH dehydrogenase complex chloroplast
96	AT1G06190	4,11	chloroplast Rho termination factor phosphorylative transmembrane transport
97	AT2G10940	4,09	chloroplast thylakoid lipid transport protein
98	AT4G28660	4,00	chloroplast PSII reaction center protein PSB28
99	AT2G04530	3,98	tRNase z 2 TRZ2 of chloroplast tRNA processing
100	AT3G26740	4,09	CCR-like gene CCL RNA stability determinant of chloroplast
101	AT3G54090	3,95	fructokinase-like gene FLN1 of chloroplast potential target of thioredoxin z
102	AT5G46580	3,95	pentatricopeptide (PPR) containing protein of chloroplast
103	AT2G20670	3,81	chloroplast unknown
104	AT1G34380	3,87	chloroplast 5' 3' exonuclease DNA binding
105	AT3G12930	3,88	Lojap-related protein of chloroplast
106	AT1G13930	3,89	salt stress Cd induced unknown chloroplast envelope ?
107	AT3G03630	3,89	O-acetyl serine (thiol) lyase of chloroplast Cys biosynthesis
108	AT5G55740	3,90	chlororespiratory reduction 21 CCR21 PPR domain containing protein chloroplast
109	AT3G01440	3,94	PsbQ-like 1 PQL1 component of chloroplast nitrite reductase complex or PSII or NADPH reductase
110	AT5G14660	3,80	peptide deformylase 1B PDF1B of chloroplast
111	AT1G55490	3,78	lesion initiation 1 LEN1 chaperonin 60 beta CPN60B of chloroplast mutants have SAR induced, accelerate cell death to heat stress
112	AT1G15390	3,74	peptide deformylase 1A PDF1A of chloroplast and / or mitochondrion
113	AT3G27690	3,76	light harvesting chlorophyll B binding protein 2.3 LHCB2.3 or LHCB2.4 of chloroplast PSII

114	AT4G09350	3,77	chloroplast DnaJ Hsp40 type chaperone
115	AT2G21330	3,70	fructose biphosphate aldolase FBA1 of chloroplast or/and 5 other locations including apoplast - there are chloroplast and cytosolic isoforms encoded by the same gene Cd responsive pentose phosphate shunt
116	AT1G55370	3,72	NDH-dependent cyclic electron flow 5 NDF5 chloroplast carbohydrate binding
117	AT1G22630	3,75	unknown chloroplast protein
118	AT5G39790	3,64	starch binding protein chloroplast AMP activated protein kinase related
119	AT3G58990	3,67	isopropylmalate isomerase 1 IPMP1 of chloroplast leu biosynthesis
120	AT1G62750	3,53	snowy cotyledon 1 SCO1 GTP binding chloroplast located
121	AT3G14900	3,52	unknown chloroplast protein
122	AT1G75460	3,65	LON protease of chloroplast
123	AT1G62780	3,55	unknown chloroplast protein
124	AT4G36580	3,45	chloroplast AAA type ATPase Zn binding
125	AT3G18420	3,44	prenyltransferase superfamily chloroplast protein
126	AT3G03920	3,43	chloroplast pseudouridine synthase
127	AT4G17600	3,42	light-harvesting like LIL3:1 chloroplast transcription factor
128	AT4G25990	3,41	chloroplast (import apparatus)
129	AT2G23670	3,39	chloroplast unknown thylakoid component
130	AT4G16390	3,36	pentatricopeptide (PPR) containing protein of chloroplast
131	AT2G28605	3,36	Ca binding OEC23 like protein of chloroplast
132	AT1G68590	3,34	chloroplast ribosomal protein
133	AT5G53580	3,32	aldo keto reductase (NAD(P)H) of chloroplast
134	AT1G34000	3,32	one helix protein 2 OHP2 chloroplast thylakoid
135	AT4G19100	3,31	unknown chloroplast protein
136	AT5G02050	3,32	unknown chloroplast mitochondrial matrix protein
137	AT5G28500	3,31	unknown chloroplast stroma
138	AT5G45930	3,29	Mg chelatase I2 of chloroplast stroma
139	AT4G25050	3,27	acyl carrier protein 4 ACP4 of chloroplast stroma
140	AT4G34190	3,27	stress enhanced protein 1 SEP1 chloroplast thylakoid
141	AT5G63420	3,26	chloroplast hydrolase RNA metabolisin metallo beta lactamase
142	AT3G06980	3,22	helicase of chloroplast unknown function
143	AT4G09040	3,20	chloroplast RNA binding protein (RRM domain)
144	AT5G64580	3,20	chloroplast metalloendopeptidase AAA type ATPase
145	AT3G47650	3,19	DnaJ/Hsp40 cys-rich domain superfamily protein chaperone of chloroplast
146	AT5G54180	3,19	chloroplast related to mt transcription termination factor
147	AT3G62910	3,16	albino and pale green 3 APG3 chloroplast translation release factor
148	AT1G12900	3,15	glaceraldehyde diphosphate dehydrogenase A subunit 2 GAPA-2 of chloroplast, apoplast, membrane
149	AT1G15510	3,17	early chloroplast biogenesis 2 ECB2 vanilla cream 1 VAC1 pentatricopeptide repeat protein necessary for accD RNA editing in chloroplast
150	AT3G12340	3,11	peptidyl prolyl cis trans isomerase of chloroplast thylakoid lumen
151	AT1G56050	3,12	chloroplast GTP binding unknown function
152	AT4G26370	3,12	chloroplast NusB-like protein (antitermination)
153	AT3G55330	3,10	PsbP like protein 1 PPL1 pf chloroplast Ca binding
154	AT4G33010	3,12	glycine decarboxylase P-protein 1 located in Chloroplast, mitochondrion, cytoplasm
155	AT1G09390	3,11	GDSL-like lipase/Acyl hydrolase of chloroplast stroma
156	AT1G73530	3,09	RRM domain RNA binding protein of chloroplast
157	AT1G01080	3,03	chloroplast RNA binding protein (RRM domain)
158	AT4G11175	3,02	translation initiation factor of chloroplast
159	AT4G24770	3,05	RNA binding protein RBP31 of chloroplast
160	AT4G28210	3,03	chloroplast embryo-defective 1923 emb1923 unknown function
161	AT5G62140	3,02	chloroplast unknown function
162	AT5G06290	3,01	2-Cys peroxyredoxin B of chloroplast
163	AT1G32520	3,01	unknown chloroplast protein
164	AT4G15550	3,04	UDP-Glucose indole 3-acetate beta glucosyl transferase IAGLU of chloroplast
165	AT5G57180	3,02	chloroplast import apparatus 2 CIA2 transcription regulator of import apparatus
166	AT5G18570	3,00	chloroplast GTPase Embryo defective 269 homolog of bacterial Ogb
167	AT5G63310	2,99	nucleoside diphosphate kinase NDPK2 NDPK1A of chloroplast ox UV stress light signal transduction
168	AT3G48500	2,98	pigment defective 312 PDE312 plastid transcriptionally active 10 PTAC10 of chloroplast OB fold like
169	AT5G22640	2,97	EMB1211 chloroplast thylakoid membrane protein
170	AT3G01500	3,00	SA binding protein 3 SABP3 beta carbonic anhydrase 1 CA1 chloroplast envelope
171	AT5G24165	2,94	unknown chloroplast protein
172	AT1G51100	2,93	unknown chloroplast protein (stroma)

173	AT5G48470	2,94	unknown chloroplast protein
174	AT1G11700	2,91	unknown chloroplast protein
175	AT5G52970	2,91	unknown chloroplast protein (thylakoid lumen)
176	AT3G20440	2,90	embryo defective EMB2729 putative glycoside hydrolase of chloroplast
177	AT2G31840	2,92	thioredoxin superfamily of chloroplast
178	AT4G29060	2,90	chloroplast translation elongation factor EF1B fold EMB2726 Cd responsive
179	AT3G25480	2,88	unknown chloroplast protein
180	AT3G04760	2,87	pentatricopeptide repeat (PPR) protein of chloroplast
181	AT2G28000	2,87	chaperonin 60 alpha CPN60A chloroplast rubisco folding
182	AT2G33180	2,86	unknown chloroplast protein
183	AT2G36145	2,83	unknown chloroplast protein
184	AT3G47070	2,83	chloroplast thylakoid protein
185	AT3G23940	2,83	dihydroxy acid / 6-phosphogluconate dehydratase family of chloroplast
186	AT3G02150	2,82	plastid transcription factor 1 PTF1 teosinte branched 1 of chloroplast
187	AT3G18680	2,83	uridylate (UMP) kinase of chloroplast
188	AT2G44640	2,83	unknown chloroplast protein (also mitochondria, plasma membrane ?)
189	AT1G59990	2,81	DEAD/DEAH RNA helicase of chloroplast
190	AT3G23700	2,83	ribosomal protein S1 of chloroplast
191	AT3G51510	2,78	chloroplast thylakoid protein
192	AT3G06950	2,78	pseudouridine synthase of chloroplast
193	AT2G36990	2,78	sigma factor 6 SIG6 of chloroplast sigma70 family
194	AT3G54050	2,77	high cyclic electron flow 1 HCEF1 fructose bisphosphate 1 phosphatase chloroplast stroma perhaps apoplast cold response
195	AT5G53490	2,77	thylakoid lumen protein chloroplast
196	AT1G79560	2,76	EMB1047, EMB156 FtsH protease 12 FTSH12 of chloroplast
197	AT1G78930	2,75	unknown chloroplast protein
198	AT3G48420	2,76	haloacid dehalogenase like hydrolase of chloroplast
199	AT1G74850	2,74	plastid transcriptionally active 2 PTAC2 TF of chloroplast
200	AT2G14880	2,73	SWIB/MDM2 superfamily chloroplast protein
201	AT5G62440	2,73	domino1 defective chloroplasts and leaves nuclear unknown function
202	AT3G63160	2,73	chloroplast thylakoid protein
203	AT4G34740	2,72	glutamine 5 phosphoribosylpyrophosphate amido transferase 2 ASE2 chloroplast import apparatus 1 CIA1 chloroplast stroma
204	AT1G42550	2,72	plastid movement impaired 1 PMI1 plasma membrane protein (chloroplast movement)
205	AT1G44575	2,71	nonphotochemical quenching 4 NPQ4 PS II component chloroplast
206	AT2G35370	2,71	glycine decarboxylase complex H protein chloroplast (mitochondrion)
207	AT2G35500	2,70	shikimate kinase like 2 aromatic amino acid biosynthesis chloroplast
208	AT1G28530	2,68	unknown chloroplast protein
209	AT5G58770	2,68	dehydrodolichyl diphosphate synthase of chloroplast
210	AT5G42070	2,69	chloroplast thylakoid membrane protein
211	AT4G30720	2,67	oxidoreductase / electron carrier of chloroplast stroma
212	AT3G15850	2,68	fatty acid desaturase 5 FAD5 of chloroplast
213	AT3G47860	2,67	lipocalin of chloroplast thylakoid lumen
214	AT2G36830	2,67	tonoplast intrinsic protein 1;1 TIP1;1 gamma-TIP1 water channel of vacuole and or chloroplast
215	AT1G32470	2,66	glycine decarboxylase of chloroplast mitochondrion ?
216	AT4G34730	2,67	ribosome binding factor A family rRNA processing of chloroplast
217	AT3G53460	2,66	RNA-binding protein29 of chloroplast
218	AT5G18660	2,67	3,8 divinyl protochlorophyllide 8 vinyl reductase of chloroplast
219	AT5G03940	2,67	signal recognition particle 54 kDa subunit of chloroplast
220	AT2G15000	2,65	unknown chloroplast protein
221	AT3G06790	2,64	role in chloroplast development subtilisin related peptidase
222	AT1G67080	2,65	ABA Xanthophyll biosynthesis (abscisic acid deficient 4 ABA4) PSII photoprotection chloroplast
223	AT1G57770	2,64	oxidoreductase amine oxidase of chloroplast
224	AT5G11450	2,64	chloroplast thylakoid protein PsbP family
225	AT3G12080	2,64	EMB2738 GTP binding HSR1 related chloroplast protein
226	AT2G21530	2,64	SMAD/FHA domain chloroplast thylakoid membrane
227	AT5G23310	2,62	Fe superoxide dismutase FSD3 of chloroplast
228	AT5G51460	2,61	trehalose 6 phosphate phosphatase of chloroplast
229	AT1G31800	2,61	Cytochrome P450 CYP97A3 Lutein deficient 5 carotene beta ring hydroxylase of chloroplast
230	AT3G24430	2,61	high chlorophyll fluorescence 101 HCF101 chloroplast protein
231	AT2G27775	2,58	unknown chloroplast protein
232	AT3G05410	2,59	chloroplast PpsbP family membrane protein oxygen evolving complex Ca binding

233	AT4G38100	2,57	unknown chloroplast thylakoid membrane protein
234	AT5G05740	2,59	ethylene-dependent gravitropism-deficient and yellow-green like 2 EGY2 S2P like endopeptidase of chloroplast thylakoid membrane
235	AT1G50450	2,58	Saccharopine dehydrogenase of chloroplast
236	AT4G20130	2,58	plastid transcriptionally active 14 PTAC14 of chloroplast
237	AT5G16715	2,57	EMB2247 of chloroplast Val tRNA ligase
238	AT3G21110	2,55	phosphoribosylaminoimidazolesuccinocarboxamide synthase activity PUR7 purine biosynthesis auxin stimulated chloroplast
239	AT1G67660	2,55	DNA binding restriction endonuclease type II like chloroplast
240	AT5G09760	2,54	chloroplast pectin methylesterase inhibitor
241	AT3G20680	2,54	unknown chloroplast protein
242	AT1G71500	2,52	chloroplast thylakoid membrane protein
243	AT1G79510	2,50	unknown chloroplast protein
244	AT3G20330	2,51	pyrimidine B of chloroplast aspartate carbamoyl transferase pyrimidine ribonucleotide de novo synthesis
245	AT3G55010	2,50	phosphoribosylformylglycinamide cyclo-ligase activity PUR5 aminoimidazole ribonucleotide synthase AIR of chloroplast, mitochondrion
246	AT5G45680	2,49	FK506 binding protein 13 FKBP13 peptidyl prolyl cis trans isomerasethylakoid lumen of chloroplast
247	AT3G48110	2,50	glycine tRNA synthetase embryo-defective-development 1 EDD1 of chloroplast
248	AT5G47190	2,48	chloroplast ribosomal protein L19 family
249	AT4G30620	2,48	chloroplast envelope protein YbyB family
250	AT3G61080	2,47	fructosamine ketosamine 3 kinase family of mitochondria and chloroplasts
251	AT1G09340	2,48	chloroplast stem loop binding of 41 kDa CSP41B CRB RNA binding protein
252	AT3G49240	2,47	EMB1796 mitochondrial chloroplast protein PPR superfamily
253	AT1G13270	2,47	methionine amino peptidase 1B now MAP1C of mitochondria chloroplasts
254	AT2G38270	2,45	CAX interacting protein 2 CXIP2 glutaredoxin of chloroplast
255	AT3G51870	2,45	chloroplast mitochondria membrane transporter (substrate carrier)
256	AT5G51110	2,45	4-alpha-hydroxytetrahydrobiopterin dehydratase activity of chloroplast cofactor biosynthesis
257	AT4G14890	2,46	2 Fe 2 S cluster component chloroplast
258	AT3G04550	2,45	unknown chloroplast protein
259	AT4G09730	2,44	RNA helicase RH9 DEAD box ribosome biogenesis of chloroplast
260	AT3G59840	2,43	unknown chloroplast protein
261	AT5G51720	2,43	2Fe 2S cluster component Chloroplast
262	AT3G44020	2,42	chloroplast thylakoid lumen protein
263	AT5G61440	2,41	atypical Cys His rich thioredoxin 5 ACHT5 of chloroplast
264	AT5G42310	2,42	PPR domain protein of chloroplast
265	AT3G61770	2,43	haloperoxidase related protein of chloroplast
266	AT1G23740	2,44	Zn binding dehydrogenase of chloroplast
267	AT5G55580	2,41	similar to mitochondrial transcription termination factor of chloroplast ?
268	AT2G17240	2,40	unknown chloroplast protein
269	AT1G70200	2,41	RRM domain protein of chloroplast
270	AT1G67120	2,41	chloroplast envelope protein AAA+ ATPase domain sigma factor domain
271	AT3G17930	2,39	chloroplast thylakoid membrane protein
272	AT2G48070	2,38	chloroplast protein RPH1 resistant to Phytophthora brassicae 1 mutants are susceptible pos regulation of hydrogen peroxide
273	AT4G00370	2,39	inorganic phosphate transporter PHOT4;4 of chloroplast
274	AT1G02280	2,39	plastid protein import 1 TOC33 chloroplast protein
275	AT3G53130	2,38	zeinoxanthin epsilon hydroxylase activity cytochrome P450 CYP97C1 lutein-deficient 1 LUT1 carotenoid biosynthesis in chloroplast
276	AT1G70890	2,37	MLP like protein 43 MLP43 of chloroplast
277	AT1G80270	2,39	PPR 596 (pentatricopeptide repeat) protein of chloroplast
278	AT1G07320	2,37	ribosomal protein L4 of chloroplast
279	AT1G36390	2,38	GrpE family of co-chaperones of chloroplast
280	AT3G04260	2,37	plastid transcriptionally active 3 PTAC3 TF of chloroplast
281	AT2G21640	6,09	response to ox stress mitochondrial unknown
282	AT4G25200	5,40	mt heat shock protein response to heat, Cd HSP23.6
283	AT1G64220	4,36	mitochondrial outer membrane transporter TOM7-2
284	AT2G24120	3,78	scabra 3 SCA3 DNA-dep RNA pol of mitochondria
285	AT3G30775	3,72	proline oxidase early response to dehydration 5 ERD5 inner mitochondrial membrane ox stress responsive
286	AT4G02990	3,54	mitochondrial transcription termination factor related chloroplast ?
287	AT5G61030	3,18	glycin-rich RNA binding protein 3 GR-RBP3 of mitochondria
288	AT5G14580	3,16	mitochondrial 3' 5' exonuclease polyribonucleotide nucleotidyltransferase mRNA catabolism
289	AT3G50930	3,06	mitochondrial ATPase Cytochrome BC1 synthesis
290	AT1G23100	2,95	GroES like mitochondria

291	AT3G23990	2,89	HSP60 3B mitochondrial chaperone Cd responsive
292	AT1G04640	2,71	lipoyltransferase 2 LIP2 mitochondria
293	AT5G47630	2,68	mitochondrial acyl carrier protein 3 mtACP3 fatty acid biosynthesis
294	AT1G73260	2,66	kunitz trypsin inhibitor 1 ATK11 SA induced bacterial defense modulates cell death mitochondrial protein
295	AT1G50400	2,62	voltage gated anion channel mitochondrial outer membrane porin family
296	AT1G15870	2,56	mitochondrial matrix protein
297	AT5G05990	2,55	mitochondrial matrix protein
298	AT3G07770	2,53	Hsp89.1 of mitochondria
299	AT5G39840	2,45	mitochondrial RNA helicase
300	AT4G23290	2,43	cys-rich receptor like protein kinase 21 RLK21 mitochondrial ?
301	AT5G43150	2,40	unknown mitochondrial protein
302	AT5G13930	13,87	chalcone synthase TT4 response to UV, ox stress, JA
303	AT5G08640	11,39	flavonol synthase 1 FLS1
304	AT1G14280	5,04	phytochrome kinase substrate 2 PKS2 hypocotyl phototropism complex with Phot1 Phot2 NPH3
305	AT3G51240	4,94	Flavone 3 hydroxylase TT6 UV-B response flavin biosynthesis
306	AT1G06000	4,68	flavonol 7- O rhamnosyl transferase
307	AT5G58140	3,54	nonphototropic hypocotyl1 like NPL1 phototropin 2 PHOT2 blue light receptor
308	AT3G55120	3,48	chalkone flavanone isomerase TT5 response to UV
309	AT5G05270	2,71	chalkone flavone isomerase family protein intramolecular lyase response to karrikin
310	AT3G45780	2,71	nonphototropic hypocotyl 1 NPH1 PHOT1 phototropin 1
311	AT4G34138	2,63	UDP-glucosyl transferase 73B1 Quercetin 7/3 O glucosyl transferase ABA glucosyl transferase
312	AT3G46660	2,41	quercetin 7-O-glucosyltransferase
313	AT3G15570	2,41	phototropic response NPH3 family protein light signaling
314	AT1G65060	2,39	4-coumarate CoA ligase 3 4CL3 UV stress
315	AT4G36540	10,19	BR enhanced expression 2 BEE2 HLH TF
316	AT2G18300	9,31	bHLH TF response to cytokinin
317	AT3G61630	6,40	cytokinin response factor 6 CRF6 ERF/AP2 TF
318	AT2G31380	6,05	B-box Zn finger protein COP1 interacting TF endomembrane-localized ? salt tolerance homologous gene (STH)
319	AT5G46710	5,61	PLATZ transcription factor family
320	AT3G56400	4,68	WRKY70 activator of SA defenses, suppressor of JA responses
321	AT2G28510	3,96	Dof TF
322	AT4G32980	3,85	ATH1 homeob ox TF
323	AT1G73830	3,70	BR enhanced expression 3 BEE3 TF w/ HLH domain
324	AT5G25190	3,69	b6 subfamily of ERF/AP2 family transcription factors pathogenesis-related ? TF
325	AT5G17300	3,72	reveille1 RVE1 myb TF regulates auxin levels in circadian dependance
326	AT1G22590	3,48	agamous-like 87 AGL87 TF
327	AT2G21320	3,40	B -box Zn finger TF endomembrane system ?
328	AT4G00950	3,39	maternal effect embryo arrest 47 MEE47 unknown function TF ?
329	AT5G60890	3,40	altered tryptophan regulation 1 ATR1 MYB34 TF of tryptophan biosynthesis (auxin and glucosinolates !!)
330	AT2G36080	3,32	AP2/B3 plant specific TF repressor response to karrikin
331	AT3G15680	3,31	Ran BP2/ZNF Zn finger like protein unknown function
332	AT5G06550	3,29	F- box,Jumonji TF / aspartyl beta hydroxylase domain surface receptor downstream signaling
333	AT4G01250	3,28	WRKY22 TF senescence chitin induced ?
334	AT3G17609	3,26	HYH HY5 homolog TF bZIP light UV responsive
335	AT1G70700	3,28	JAZ9 JA response gene
336	AT1G79700	3,26	TF ERF like domain
337	AT2G46510	3,22	bHLH TF ABA inducible
338	AT5G59820	3,14	response to high light 41 RHL41 putative ZnF TF high light and ox stress response
339	AT2G31070	3,16	TCP domain TF TCP10 leaf morphogenesis
340	AT1G35560	3,13	TCP family transcription factor TF
341	AT4G36930	3,10	spatula SPT bHLH TF
342	AT3G25940	3,06	TFIIB Zn binding protein transcript elongation
343	AT5G37260	3,06	reveille 2 RVE2 Myb TF aka circadian 1 CIR1 salt ABA ethylene auxin JA GA responsive
344	AT2G02080	2,97	Zn finger transcription factor TF
345	AT4G29080	2,96	IAA 27, phytochrome-associated protein 2 PAP2 auxin-associated TF
346	AT4G00050	2,97	bHLH TF unfertilized embryo sac 10 UNE10
347	AT1G43160	3,00	subfamily B-4 of ERF/AP2 TF family response to cold salt osmotic stress ABA SA, JA
348	AT1G49010	2,95	Myb double homeodomain protein TF ? salt stress auxin induced Cd induced also JA SA GA responsive
349	AT5G60850	2,87	OBP binding protein 4 Zf DOF TF
350	AT3G53310	2,84	AP2/B3 like TF
351	AT3G15500	2,86	ANAC55 ATNAC3 NAC domain containing TF response to drought ABA JA

352	AT5G15210	2,77	Zinc finger homeodomain 3 ZFHD3 homeobox protein 30 HB30 gene TF
353	AT5G52380	2,76	vascular-related NAC domain 6 VND6 TF?
354	AT3G16770	2,76	ethylene response factor AP2 3 related B2 subfamily of ethylene response factor family TF suppressor of Bax-induced cell death in tobacco OX in tobacco H2O2 and heat stress resistance
355	AT5G39610	2,77	NAC2 NAC6 ORE1 oesara 1 senescence related TF salt ethylene upregulated
356	AT1G76590	2,75	PLATZ transcription factor family TF
357	AT4G17500	2,77	ERF1 ethylene response element binding factor 1 TF
358	AT2G03710	2,66	agamous like 3 MADS box protein TF
359	AT4G16780	2,64	homeobox protein 2 HB-2 cytokinin responsive TF
360	AT1G76110	2,63	HMG box protein w/ arid/iright Dna binding domain TF
361	AT1G72440	2,58	embryo sac development arrest 25 EDA25 slow walker 2 SWA2 CCAAT binding domain TF ? cell cycle control nuclear fusion
362	AT3G25890	2,54	ERF/AP2 TF subfamily B-6
363	AT5G05790	2,51	TF homeodomain superfamily Myb domain
364	AT1G04250	2,49	IAA17 AXR3 TF auxin response
365	AT1G04240	2,44	short hypocotyl 2 SHY2 IAA3 TF
366	AT3G55980	2,47	salt inducible Zn finger 1 gene SZF1 TF salt chitin inducible
367	AT5G15830	2,45	bZIP3 TF transcription factor
368	AT3G46130	2,41	Myb48 (Myb111) TF SA induced
369	AT5G64570	17,51	xylan 1,4-beta-xylosidase activity xylan catabolism
370	AT5G65730	15,46	xyloglucan endotransglucosylase/hydrolase 6 xTH6 water stress responsive
371	AT3G15720	12,18	polygalacturonase cell wall
372	AT5G44130	8,12	fasciclin arabingalactan protein 13 precursor cell wall
373	AT3G44990	7,64	xyloglucan:xyloglucosyl transferase activity xyloglucan endotransglucosylase/hydrolase 31 XTH31apoplatic
374	AT4G37800	6,16	xyloglucan:xyloglucosyl transferase activity xyloglucyn endotransglycosylase/hydrolase 7 XLH7 glucan metabolism
375	AT1G09750	5,37	aspartic endopeptidase of apoplast
376	AT5G64620	5,14	cell wall/vacuolar inhibitor of fructosidase 2 C/VIF2 pectinesterase inhibitor endomembrane
377	AT5G03760	4,86	cellulose synthase like 9A CSLA9 beta mannan synthase enzyme required for Agrobacterium transformation
378	AT2G05540	4,95	Gly-rich protein endomembrane
379	AT4G01080	4,80	trichome birefringence like 26 sek. cell wall cellulose deposition via pectin esterification state
380	AT4G03210	4,18	xyloglucan:xyloglucosyl transferase activity XTH9 (cell wall loosening) endotransglycosylase
381	AT3G23730	4,06	xyloglucan:xyloglucosyl transferase activity XTH16 (cell wall loosening) endotransglycosylase
382	AT2G34070	3,74	trichome birefringence-like 37 TBL37 sec. cellulose deposition influenced via pectin esterification
383	AT4G30280	3,80	xyloglucan endotransglucosylase activity XTH18
384	AT5G11420	3,70	unknown cell wall component
385	AT1G53070	3,51	lectin cell wall component unknown function
386	AT5G51550	3,45	exordium like 3 EXL3 cell wall component
387	AT5G62360	3,23	invertase/pectin methylesterase inhibitor superfamily
388	AT3G54400	3,23	aspartic type endopeptidase of cell wall
389	AT1G10550	3,19	xyloglucan:xyloglucosyl transferase activity XTH33 cell wall modification plasma membrane localized
390	AT4G30450	3,17	gly-rich protein
391	AT5G23870	3,02	pectinacetylerase domain plant cell wall
392	AT3G62820	2,96	plant invertase / pectin metylesterase inhibitor of endomembrane system ?
393	AT4G11190	2,92	lignan biosynthesis defense response
394	AT1G64160	2,99	lignan biosynthesis defense response
395	AT3G54260	2,72	trichome birefringence like 36 TBL36 sek. cell wall deposition
396	AT4G25260	2,47	invertase / pectin methylesterase inhibitor superfamily of endomembrane shade avoidance response
397	AT3G53190	2,46	pectate lyase superfamily membrane anchored
398	AT5G44400	2,42	FAD oxidoreductase of cell wall
399	AT5G49360	2,41	alpha-N-arabinofuranosidase activity beta xylosidase 1 sek cell wall thickening

Appendix 2 Upregulated genes in induced compared to uninduced condition in *RV86-5*.

	AGI	FC (up-regulation)	Short description
1	AT4G36410	10,76	(ub ligase) UBC17
2	AT3G02070	4,85	Cys peptidase otubain domain ?
3	AT4G31820	4,32	enhancer of pinoid BTB/POZ domain protein PIN localization
4	AT2G41370	4,14	Blade on petiole 2 BOP2 BTB/POZ domain Ankyrin protein flower morphogenesis
5	AT3G08700	4,07	UBC 12
6	AT4G28890	3,67	RIING/U-box superfamily C3HC4 RING
7	AT1G72220	3,62	RIING/U-box superfamily C3HC4 RING
8	AT4G30940	3,27	voltage-gated K transporter BTB/POZ, WD40 domains
9	AT3G09760	3,19	RING U-box Zn finger C3HC4
10	AT4G35480	3,18	RING H2 finger protein RHA3B
11	AT5G66620	2,80	DAR6 DA1-related protein 6 ubiquitin interacting motif
12	AT1G63850	2,68	BTB /POZ domain protein
13	AT3G47910	2,67	ub carboxyl-terminal hydrolase
14	AT1G51550	2,64	Kelch F-box protein
15	AT1G80960	2,65	F-box LRR protein
16	AT1G18910	2,55	Znf: CHY type CTCHY type RING type
17	AT2G47700	2,48	red and far red insensitive 2 RFI2 RING ligase
18	AT5G06600	2,46	ub protease UBP12
19	AT5G59550	2,44	RING finger protein
20	AT1G24440	2,40	RING finger protein
21	AT5G64920	2,39	COP-interacting protein 8 CIP8 ub ligase
22	AT3G57130	2,35	Blade on petiole 1 BPO1 ankyrin repeat protein BTB/POZ domain flower development
23	AT3G58040	2,30	SINAT2 RING ligase TRAF-like domain
24	AT4G39140	2,28	RING finger protein
25	AT3G05200	2,28	RING H2 Zn finger protein ALT2 ligase chitin-induced
26	AT4G30520	18,94	LRR family S/T kinase
27	AT3G01840	9,23	endomembrane S/T/Y protein kinase
28	AT5G04470	6,03	Cyclin-dep. kinase inhibitor (endoreduplication, mitosis) SIAMESE SIM
29	AT5G07150	5,64	protein kinase membrane, LRR containing
30	AT2G32800	4,70	membrane S/T kinase
31	AT4G37870	4,51	phosphoenolpyruvate cyrboxykinase resp. to Cd, incompatible fungus
32	AT5G57050	4,49	ABA, osmotic stress responsive phosphatase PP2C ABI2
33	AT1G76040	4,33	Calcium/Calmodulin-dep protein kinase CPK29
34	AT5G59220	4,38	ABA-dep phosphatase ABA-induced PP2C gene 1
35	AT5G11020	4,32	S/T kinase
36	AT1G48480	4,19	receptor-like kinase 1 (RKL1)
37	AT5G58150	4,17	S/T kinase LRR transmembrane kinase
38	AT4G38470	4,03	S/T/Y kinase ACT type
39	AT1G45160	4,04	protein kinase AGC kinase
40	AT2G17520	4,01	transmembrane ribonuclease / protein kinase IRE1A
41	AT1G60440	3,91	pantothenate kinase (CoA biosynthesis)
42	AT5G24810	3,74	ER protein ABC1 related kinase-like domain
43	AT5G56790	3,64	protein kinase
44	AT2G25090	3,47	SNF1-related protein kinase 3.18, CIPK 16
45	AT5G01820	3,46	SNF1-related protein kinase 3.15, CIPK 14
46	AT4G02420	3,42	Con-A like lectin protein kinase
47	AT2G26290	3,39	root-specific kinase 1 ARSK1 water response ?
48	AT3G60440	3,35	phosphglycerate mutase family Histidine phosphatase family
49	AT2G20900	3,30	diacylglycerol kinase DGK5 prot kinase C activation
50	AT5G16590	3,29	LRR1 LRR containing transmembrane rec. kinase symbiotic funguns induced
51	AT3G11870	3,26	IRE1-like protein kinase ribonuclease
52	AT3G13380	3,27	BR1-like 3 BRL3 LRR transmembrane kinase
53	AT2G18170	3,12	MPK7 map kinase
54	AT2G30040	3,08	MAPKKK14 protein kinase
55	AT5G41990	3,02	WNK8 with no lysine kinase 8 phosphorylates vacuolar proton ATPase subunit
56	AT4G28490	3,02	HAESA receptor-like kinase LRR repeat
57	AT4G35500	2,98	S/T kinase

58	AT2G22560	2,97	kinase interacting protein (KIP1) family protein
59	AT1G71830	3,03	LRR S/T kinase SERK1
60	AT4G13000	2,96	AGC (cAMP, cGMP Prot kinase C) family protein kinase
61	AT3G22420	2,92	WNK2 with no lysine kinase 2
62	AT5G42750	2,86	BR11 kinase inhibitor BKI1
63	AT3G24660	2,82	transmembrane kinase like 1 TMKL1 LRR containing
64	AT5G65530	2,82	protein kinase
65	AT4G38230	2,80	CPK26 Ca-dependent protein kinase calmodulin-dependent EF hand
66	AT5G64450	2,73	Histidin phosphatase phosphoglycerate mutase family unknown function
67	AT4G39270	2,63	LRR S/T kinase
68	AT1G05000	2,68	protein tyrosin phosphatase
69	AT1G18350	2,63	map kinase kinase 7 MKK7 BUD1 auxin transport SA signaling SAR regulation
70	AT1G03920	2,53	protein kinase (cGMP dependent ?)
71	AT1G79630	2,50	protein phosphatase PP2C family
72	AT4G24400	2,50	protein kinase CIPK8 Snf related protein kinase 3.13 calcineurinB like (CLB) interacting regulates early nitrate response
73	AT1G48260	2,48	SnRK3.21CIPK17 protein kinase
74	AT5G10720	2,48	cytokinin independent 2 CKI2 AHK5 histidine kinase response to H2O2 ABA neg regulation
75	AT5G07180	2,42	protein kinase erecta-like 2 ERL2 LRR receptor like kinase
76	AT3G09780	2,37	crinkly4-related 1 CCR1 protein kinase regulator of chromatin condensation domain
77	AT2G39360	2,37	receptor-like protein kinase
78	AT4G40010	2,45	SnRK2.7 protein kinase osmotic stress activated
79	AT1G74740	2,33	Ca dependent protein kinase1A CDPK1A, CPK30
80	AT1G75820	2,34	Clavata 1 CLV1 receptor kinase LRR
81	AT4G32000	2,33	protein kinase endomembrane system
82	AT1G69960	2,32	protein phosphatase PP2A auxin transport regulation
83	AT1G47380	2,29	phosphatase 2C family member
84	AT5G03140	2,33	Con A like lectin protein kinase
85	AT1G68690	2,27	S/T protein kinase
86	AT3G54030	8,28	membrane-associated protein kinase N-myristoylated
87	AT1G68450	25,56	small unknown plant-specific evtl chloroplast
88	AT1G26130	13,69	haloacid dehalogenase-like hydrolyse family (chloroplast envelope ion transporter phospholipid transport
89	AT1G58300	11,25	heme oxygenase 4 chloroplast (decyclizing)
90	AT4G35985	11,25	senescence/dehydration associated chloroplast
91	AT5G59080	9,97	chloroplast ox stress response
92	AT4G32810	7,41	Chloroplast protein Oxidoreductase in Xynthophyll catabolism Carotenoid cleaving oxygenase 8 CCD8 MAX4
93	AT5G52810	6,13	chloroplast protein NAD(P) binding Rossmann fold Arg to Glu metabolism
94	AT5G67520	5,49	chloroplast kinase APK4 sulfate activation
95	AT1G08550	4,71	violaxanthin de-epoxidase chloroplastic thylakoid lumen
96	AT5G01790	4,50	unknown chloroplast
97	AT5G43745	4,39	chl envelope protein
98	AT5G24300	4,30	glycosyl transferase starch degradation Starch Synthase 1 Suppressor of Salicylic acid insensitivity 1
99	AT1G17745	4,14	phosphoglycerate dehydrogenase chl Ser biosynthesis
100	AT2G07732	4,09	rubisco large chain
101	ATMG00280	4,09	rubisco (large chain ?)
102	AT1G02470	3,72	unknown chloroplast
103	AT1G50020	3,57	thylakoid membrane proten unknown function
104	AT5G62220	3,50	glycosyl transferase chloroplast membrane
105	AT1G13990	3,47	unknown chloroplast
106	AT2G07713	3,23	unknown chloroplast
107	AT1G32080	3,17	chloroplast membrane
108	AT5G59400	3,02	unknown chloroplast
109	AT3G16950	2,84	lipoamide dehydrogenase 1 LPD1 chloroplast
110	AT2G36810	2,78	Armadillo repeat chloroplast protein
111	AT2G15570	2,75	thioredoxin M-type 3 chloroplast
112	AT1G69680	2,73	maybe PSII protein (chloroplast)
113	AT1G49840	2,64	unknown chloroplast
114	AT1G08250	2,68	ADT6 arogenate dehydratase chloroplast Phe biosynthesis
115	AT1G18640	2,63	3-phosphoserine phosphatase Ser biosynthesis Chloroplast
116	AT1G16540	2,63	chloroplast molybdenum cofactor sulfuryase ABA deficient 3, Sirtinol resistant 3, altered chloroplast import 2

117	AT3G13070	2,60	CBS domain transporter-associated domain containing chloroplast unkown function
118	AT2G41220	2,61	glutmate synthase GLU2 ferredoxin-dependent chloroplast
119	AT1G15410	2,57	chloroplast racemase epimerase apartate glutamate racemase family
120	AT1G16720	2,54	translation of PsbA, mutant defect in PS II assembly chloroplast
121	AT2G04400	2,53	indole 3 glycerol phosphate synthase Trp biosynthesis chloroplast
122	AT5G55700	2,50	beta amylase 4 BAM4 chloroplast starch degradation
123	AT1G54130	2,42	GTP diphosphokinase RelA SpoT homolog 3 RSH3 chloroplast
124	AT1G76140	2,38	Ser type endopeptidase of chloroplast
125	AT2G41180	2,42	chloroplast unknown function VQ motif
126	AT1G06430	2,32	FTSH8 protease FtsH type chloroplast thylakoid
127	AT5G43780	2,35	APS4 sulfate adenlylyl transferase Chloroplast Mitochondrium
128	AT3G18295	2,23	unknown chloroplast
129	AT3G11650	2,25	NDR1/HIN1-like 2 NHL2 induced by cucumber mosaic virus chloroplast protein
130	ATMG00980	8,29	large mitochondrial ribosomal subunit protein
131	ATMG01200	7,68	mt membrane protein
132	ATMG01220	6,91	unknown mitochondrial
133	ATMG00080	6,39	mt ribosomal protein
134	ATMG00090	6,39	mt ribosomal protein
135	AT2G20500	6,44	unknown mt protein
136	AT4G03340	6,12	acetylglucosaminyl transferase mt membrane
137	AT2G07727	5,42	repsiratory electron transport electron carrier Cyt b/b6
138	ATMG00220	5,42	mt respiratory chain complex III
139	AT4G05030	5,03	mt copper transport
140	AT4G14695	4,84	unknown mitochondrial
141	AT2G07687	4,82	mt electron transport Cyt c oxidase
142	ATMG00730	4,82	mt complex IV subunit
143	AT2G07751	4,81	NADH dehydrogenase (ubiquinone)
144	ATMG00990	4,81	mt complex I subunit (NADH dehydrogenase)
145	ATMG01270	4,78	mt ribosomal small subunit protein
146	ATMG00560	4,75	mt ribosomal protein
147	AT2G07698	4,51	mt or vacuolar ATP synthase F1 ATPase component H+ transport
148	ATMG01190	4,51	mt proton transporting ATP synthase complex
149	AT1G63300	4,49	unknown mt protein
150	ATMG00210	4,51	mt large rib subunit protein
151	ATMG01000	4,44	unknown mt protein
152	ATMG00490	4,16	unknown mt protein
153	ATMG00500	4,16	unknown mt protein
154	ATMG00920	4,11	unknown mt protein
155	ATMG01010	3,68	unknown mitochondrial
156	ATMG00290	3,53	mt rib protein S4
157	ATMG01110	3,45	unknown mt protein
158	AT2G07768	3,48	mt complex IV assembly
159	ATMG00270	3,40	complex I NAD6 mt
160	ATMG01060	3,37	unknown mitochondrial
161	ATMG00180	3,32	mt membrane protein
162	ATMG01100	3,23	unknown mitochondrial
163	ATMG00970	3,24	unknown mitochondrial
164	ATMG00540	3,23	unknown mitochondrial
165	ATMG00310	3,18	unknown mitochondrial
166	AT3G01820	3,15	mt protein nucleotide kinase ?
167	ATMG00960	3,14	mt protein
168	AT2G39690	3,10	mt protein
169	AT1G07180	3,03	NADH dehydrogenase mt protein
170	ATMG01020	3,08	mt protein
171	AT5G46180	2,99	mt ornithine delta-amino transaminase (Arg catabolism hyperosmotic shock, salt stress upregulated)
172	ATMG00570	3,00	unknown mitochondrial
173	ATMG01180	2,91	unknown mitochondrial
174	AT2G07667	2,91	unknown mitochondrial
175	AT2G25530	2,82	AFG1-like ATPase protein unknown function (AFG1: mt chaperone for cyt c oxidase complex)
176	AT2G07777	2,83	unknown mt ATP synthase 9 similarity
177	ATMG01090	2,83	unknown mitochondrial

178	ATMG00513	2,78	complex 1 component mt
179	ATMG00260	2,76	unknown mitochondrial
180	ATMG00740	2,80	unknown mitochondrial
181	ATMG00940	2,73	unknown mitochondrial
182	ATMG00580	2,57	mt complex I component NADH dehydrogenase subunit 4
183	ATMG00520	2,60	mitochondrial; splicing factor ?
184	ATMG00820	2,53	unknown mitochondrial
185	ATMG00850	2,51	unknown mitochondrial
186	ATMG00300	2,47	unknown mitochondrial
187	ATMG00470	2,50	unknown mitochondrial
188	ATMG00400	2,48	unknown mitochondrial
189	AT2G38580	2,40	mitochondrial ATP synthase D chain related cytoplasmic ?
190	AT1G64960	2,34	mitochondrial armadillo repeat protein
191	AT1G52710	2,32	mitochondrial Cyt c oxidase ? rubredoxin-like superfamily
192	ATMG01260	2,29	unknown mitochondrial
193	AT5G61160	3,44	agmatine N4-coumaroyltransferase activity; anthocyan 5-aromatic acyl transferase (AACT1)
194	AT1G09570	2,61	far red elongated1 FRE1 far red elongated hypocotyl 2 FHY2 PhyA dependent regulator nuclear
195	AT1G07250	2,37	(quercetin 3, 7, 3'-) UDP glucosyl transferase UGT71C4
196	AT4G34050	2,27	caffeoyl-CoA O-methyltransferase activity coumarin biosynthesis Cd responsive
197	AT5G07500	30,47	PEI1 Zn finger transcr factor embryo-specific
198	AT4G32280	20,86	IAA29, TF response to red light, auxin
199	AT1G10585	20,43	bHLH TF
200	AT5G10120	19,68	ethylene insensitive 3 family TF responds to karrikin
201	AT2G40970	11,87	MYBC1 homeodomain TF
202	AT2G41070	10,15	senescence-specific TF ATBZIP12, DPBF4, EEL
203	AT2G46990	9,55	auxin induced TF IAA20
204	AT1G49830	9,16	TF bHLH
205	AT5G43410	8,81	TF AP2 type B-3 subfamily (18 AP2 members in family)
206	AT1G06160	8,32	TF ethylene stimulated JA responsive AP2 type B-3 (Ethylene response factor type ERF)
207	AT1G56010	7,93	TF auxin signaling NAC1 (NAC domain)
208	AT3G10330	7,87	Cyclin type Pol II transcription factor TFIIB type Zn finger, translation initiation
209	AT5G13220	7,66	JA and wounding induced protein JAZ 10 CCT domain
210	AT3G50410	7,20	Dof TF OBF binding protein 1 auxin-responsive SA responsive cell cycle regulation
211	AT1G73805	7,22	Calmodulin-binding like (TF ?)
212	AT2G46410	6,90	TF JA responsive SA responsive R3 type Myb, moves from atrichoblasts to trichoblasts (hair formation)
213	AT1G05710	6,79	ethylene responsive TF bHLH
214	AT4G36710	6,67	TF GRAS family shoot development
215	AT4G00940	6,24	TF Dof type
216	AT5G26660	5,39	MYB86 sinapate ester biosynthesis TF
217	AT1G74080	5,22	R2R3 type Myb MYB122 TF
218	AT3G50260	5,15	stress cell death regulation JA ethylene regulated ERF/AP2 TF
219	AT4G27240	4,83	Zn binding TF C2H type
220	AT3G10500	4,60	TF NAC domain
221	AT1G54040	4,56	JA inducible senescence defense epithiospecifying senescence regulator ESP, ESR WRKY53 interactor
222	AT5G53210	4,56	TF bHLH Speechless SPCH
223	AT5G50915	4,47	GA responsive TF bHLH
224	AT3G60530	4,38	GATA4 Zn finger TF
225	AT2G31230	4,34	TF ethylene pathway ERF15 (B3 ERF/AP2 family)
226	AT5G61620	4,36	Myb type CCHC Zn finger TF
227	AT1G54330	4,19	NAC domain TF NAC020
228	AT3G04030	4,15	TF homeodomain Myb type
229	AT1G01720	4,00	wound ABA induced TF NAC domain
230	AT4G31800	3,94	defense TF WRKY18
231	AT3G61950	3,68	TF bHLH
232	AT3G50700	3,57	IDD22 (maize indeteminte 2 homolog) TF C2H2 Znf Chloroplast development ?
233	AT1G69810	3,47	WRKY36 TF
234	AT5G10970	3,42	TF Znf C2H2 type
235	AT1G62700	3,41	TF vascular-specific NAC domain
236	AT1G18100	3,64	ABA response phosphatidylethanolamine binding FT family Mother of FT and TFL1 MFT
237	AT3G01470	3,31	salt blue light responsive TF homeobox HD-ZIP1 ATHB1 leaf development
238	AT2G14210	3,24	AGL4 4 TF MADS box nutrient response side root format.

239	AT1G20900	3,18	leaf senescence ORE7 AT hook TF inhib of hypocotyl growth in light
240	AT2G47260	3,13	TF WRKY 23
241	AT3G10590	3,19	Myb domain TF
242	AT2G17150	3,12	RWP-RK family TF plant regulator
243	AT3G47600	3,10	MYB94 TF
244	AT4G35550	2,99	WOX 13 Wuschel-related homeobox family protein TF
245	AT5G52830	2,89	GA, NO mediated signal transduction TF WRKY27
246	AT5G18000	2,91	VERDANDI VDD TF MADS box ovule sac development
247	AT1G53170	2,85	ethylene response factor ERF8 TF AP2 domain
248	AT4G23750	2,87	target of monoperos 3 TMO3 ERF/AP2 TF B-2 family cytokinin response factor 2 CRF2
249	AT5G60140	2,81	AP2 B3 type TF
250	AT5G07680	2,73	TF NAC domain containing protein 80 NAC080
251	AT4G39780	2,74	DREB subfamily A-6 ERF/AP2 TF
252	AT3G60580	2,63	TF C2H2 Zn finger
253	AT1G17460	2,61	TRF-like 3 TF TRFL3 myb type TF
254	AT1G34650	2,60	homeodomain glabrous 10 HDG10 TF HD-ZIP class IV
255	AT5G60200	2,55	TF Dof type target of monoperos 6 TMO6
256	AT1G69310	2,53	WRKY57 TF
257	AT4G18020	2,57	PRR2 interacts w/Ca sensor Myb domain homeo domain TF?
258	AT3G01970	2,50	WRKY45 senescence TF
259	AT2G02450	2,52	long vegetative phase 1 LOV1 NAC domain TF NAC035
260	AT2G45120	2,42	C2H2 Zn finger DNA binding TF
261	AT4G36730	2,40	G-box binding factor 1 GBF1 bZIP TF H2O2 regulation cell aging
262	AT4G00730	2,38	homeodomain TF anthocyanless 2 ANL2 Glabra2 type
263	AT4G30410	2,40	TF sequence-specific DNA binding protein
264	AT1G53160	2,36	squamosa promoter binding protein 4 SPL4 TF
265	AT1G14580	2,37	TF C2H2 Zn finger
266	AT5G46590	2,37	Nac domain containing protein 96 NAC096 TF
267	AT2G02820	2,35	TF Myb domain protein 88 MYB88 stomata development
268	AT5G62940	2,33	TF DOF5.6 cambium development vascular tissue dev.
269	AT5G61430	2,35	NAC100 TF
270	AT2G31730	2,27	bHLH TF ethylene, GA response
271	AT3G57600	2,25	DREB subfamily A-2 of ERF/AP2 TFs
272	AT5G66350	2,29	Short internodes SHI Znf TF GA responsive
273	AT4G24020	2,25	NIN like protein 7 RWP RK family TF drought responsive Nitrate regulation
274	AT4G08150	2,37	TF Brevipedicellus 1 BP1 KNAT1 knotted 1 homolog homeobox
275	AT1G69500	22,93	(CYP704B1 Cyt P450) alkane monooxygenase activity sporopollenine formation
276	AT1G20490	16,14	4-coumarate CoA ligase
277	AT5G04970	14,64	invertase / pectin esterase inhibitor superfamily
278	AT3G10340	13,68	PAL4 (Phenylalanin ammonia lyase)
279	AT1G76470	13,16	lignin biosynthesis NADP binding; Cinnamoyl Co A reductase
280	AT1G67750	12,00	pectate lyase mambrane protein
281	AT1G80240	8,99	unknown cell-wall related
282	AT4G30290	8,60	xyloglucan endotransglucosylase
283	AT1G20160	7,43	apoplastic serine endopeptidase
284	AT4G30380	6,27	unknown hypoxia responsive natriuretic peptide / expansin precursor
285	AT5G23130	6,06	cell wall catabolism LysM domain containing peptidoglycan binding
286	AT5G22500	5,84	long chain fatty acyl CoA reductase (alcohol forming) suberin biosynthesis wound-responsive salt stress resp.
287	AT1G11545	5,47	xyloglucan:xyloglucosyl transferase
288	AT5G04310	5,27	pectin lyase membrane-anchored
289	AT1G20480	4,91	4-coumarate CoA ligase
290	AT4G36220	4,59	ferulate 5 hydroxylase lignin biosynthesis UVB inducible
291	AT4G19420	4,00	carboxyl ester hydrolase Pectinacetyl esterase family
292	AT4G13340	3,92	cell wall component
293	AT1G48100	3,81	polygalacturonase
294	AT3G44550	3,93	long chain fatty acyl CoA reductase FAR5 (suberin biosynth ? salt stress)
295	AT1G53830	3,47	pectinesterase PME2
296	AT3G16530	3,41	apoplastic cell wall lectin chitin responsive
297	AT4G11210	3,35	lignan biosynthesis defense-induced
298	AT3G62360	3,18	carbohydrate binding-like fold cell wall / ER
299	AT5G18470	3,15	Curculin-like lectin plant cell wall component

300	AT5G06330	3,14	LEA type hydroxyprolin rich glycoprotein (cell wall)
301	AT5G63180	3,02	pectate lyase
302	AT5G06860	3,03	Polygalacturonase inhibiting protein 1 PGIP1
303	AT5G62350	3,01	pectinesterase inhibitor
304	AT2G31990	3,00	exostosin family protein
305	AT1G21310	2,94	EXT3 extensin 3 cell wall component
306	AT1G02810	3,15	pectinesterase
307	AT4G34980	2,83	subtilisin-type Ser protease 2 SLP2 of middle lamella cell wall remodeling
308	AT4G26490	2,77	hydroxyprolin-rich glycoprotein family (LEA)
309	AT1G70710	2,73	endo 1,4 beta glucanase GH9B1 cell wall modification
310	AT2G19170	2,67	subtilisin type Ser endopeptidase middle lamella localized
311	AT4G39350	2,54	cellulase synthase A1 CESA2
312	AT3G50740	2,43	UDP glucose:coniferyl alcohol glucosyl transferase lignin metabolism
313	AT3G27400	2,40	pectin-lyase like superfamily unknown endomembrane
314	AT4G23820	2,40	polygalacturonase endomembrane pectin lyase superfamily
315	AT2G37585	2,32	acetylglucosaminyltransferase activity carbohydrate biosynthesis, membrane localized
316	AT1G14890	2,30	pectin methylesterase inhibitor unknown function
317	AT4G39330	2,31	cinnamyl alcohol dehydrogenase 9 CAD9
318	AT3G24670	2,36	Pectin lyase like superfamily
319	AT5G14700	2,30	cinnamoyl-CoA reductase activity lignin biosynthesis
320	AT3G16920	2,28	chitinase like protein 2 CTL2 lignin biosynthesis ?

Appendix 3 Upregulated genes in induced *RV86-5* compared to induced *sud2*.

	AGI	FC (upregulation)	Annotation
1	AT3G19980	3,30	S/T phosphatase EMB2736
2	AT5G65080	2,75	AGL68 regulates vernalization
3	AT3G46230	2,57	HSP17.4
4	AT1G43590	2,21	transposon pseudogene
5	AT5G48350	2,25	unknown
6	AT4G30380	2,70	Expansin
7	AT1G15400	2,03	unknown
8	AT5G20790	2,35	unknown endomembrane protein
9	AT3G29970	2,29	unknown
10	AT5G05250	2,06	unknown
11	AT2G29500	2,17	HSP20 family, responds to ox stress
12	AT1G73120	2,11	unknown
13	AT4G31760	1,91	peroxidase (haem)
14	AT2G11810	2,10	monogalactosyldiacylglycerol synthase
15	AT1G43800	2,34	stearoyl-acyl carrier protein desaturase
16	AT5G26130	1,77	PR1 superfamily, Cys-rich secretory protein
17	AT1G62290	1,80	Aspartyl protease Saposin family
18	AT1G70830	1,87	MLP-like 28 defense related
19	AT1G47400	2,17	unknown very small protein
20	AT5G21150	1,78	AGO9
21	AT3G05150	1,89	sugar hydrogen symport
22	AT5G39520	1,70	unknown function medium size protein
23	AT4G16370	1,73	oligopeptide transporter OPT3
24	AT5G44580	1,71	unknown small conserved membrane assoc.
25	AT4G29370	1,78	galactose oxidase Kelch repeat superfamily
26	AT4G32630	2,04	ARF GTPase activator, Znf containing
27	AT5G10040	2,21	unknown
28	AT5G15360	1,66	unknown
29	AT1G74010	1,66	Strictosidine synthase, Ca dep phosphotriesterase
30	AT3G23450	1,89	unknown intron-less
31	AT4G08150	1,73	brevipedicellus BP1 knotted-like
32	AT2G18420	1,66	GA responsive GASA/GAST/Snakin family
33	AT5G57190	1,88	phosphatidyl-Ser decarboxylase
34	AT3G43920	1,72	Dicer like 3
35	AT1G19330	1,97	unknown
36	AT4G31020	1,90	phospholipase / carboxylesterase superfamily
37	AT3G43190	1,76	sucrose synthase SUS4
38	AT3G04000	1,58	NAD(P) binding (oxidoreductase)

7 REFERENCES

- ABRAMOVITCH, R. B., R. JANJUSEVIC, C. E. STEBBINS and G. B. MARTIN, 2006 Type III effector AvrPtoB requires intrinsic E3 ubiquitin ligase activity to suppress plant cell death and immunity. *Proc Natl Acad Sci U S A* **103**: 2851-2856.
- ABRAMOVITCH, R. B., and G. B. MARTIN, 2005 AvrPtoB: a bacterial type III effector that both elicits and suppresses programmed cell death associated with plant immunity. *FEMS Microbiol Lett* **245**: 1-8.
- ANDERSEN, P., B. B. KRAGELUND, A. N. OLSEN, F. H. LARSEN, N. H. CHUA *et al.*, 2004 Structure and biochemical function of a prototypical Arabidopsis U-box domain. *J Biol Chem* **279**: 40053-40061.
- ARAVIND, L., and E. V. KOONIN, 2000 The U box is a modified RING finger - a common domain in ubiquitination. *Curr Biol* **10**: R132-134.
- AZEVEDO, C., M. J. SANTOS-ROSA and K. SHIRASU, 2001 The U-box protein family in plants. *Trends Plant Sci* **6**: 354-358.
- BACH, I., and H. P. OSTENDORFF, 2003 Orchestrating nuclear functions: ubiquitin sets the rhythm. *Trends Biochem Sci* **28**: 189-195.
- BACHMAIR, A., D. FINLEY and A. VARSHAVSKY, 1986 In vivo half-life of a protein is a function of its amino-terminal residue. *Science* **234**: 179-186.
- BACHMAIR, A., M. NOVATCHKOVA, T. POTUSCHAK and F. EISENHABER, 2001 Ubiquitylation in plants: a post-genomic look at a post-translational modification. *Trends Plant Sci* **6**: 463-470.
- BACHMAIR, A., and A. VARSHAVSKY, 1989 The degradation signal in a short-lived protein. *Cell* **56**: 1019-1032.
- BALZI, E., M. CHODER, W. N. CHEN, A. VARSHAVSKY and A. GOFFEAU, 1990 Cloning and functional analysis of the arginyl-tRNA-protein transferase gene ATE1 of *Saccharomyces cerevisiae*. *J Biol Chem* **265**: 7464-7471.
- BARI, R., B. DATT PANT, M. STITT and W. R. SCHEIBLE, 2006 PHO2, microRNA399, and PHR1 define a phosphate-signaling pathway in plants. *Plant Physiol* **141**: 988-999.
- BARTEL, B., I. WUNNING and A. VARSHAVSKY, 1990 The recognition component of the N-end rule pathway. *EMBO J* **9**: 3179-3189.
- BATES, P. W., and R. D. VIERSTRA, 1999 UPL1 and 2, two 405 kDa ubiquitin-protein ligases from *Arabidopsis thaliana* related to the HECT-domain protein family. *Plant J* **20**: 183-195.
- BLILOU, I., F. FRUGIER, S. FOLMER, O. SERRALBO, V. WILLEMSSEN *et al.*, 2002 The Arabidopsis HOBBIT gene encodes a CDC27 homolog that links the plant cell cycle to progression of cell differentiation. *Genes Dev* **16**: 2566-2575.
- BOOK, A. J., J. SMALLE, K. H. LEE, P. YANG, J. M. WALKER *et al.*, 2009 The RPN5 subunit of the 26s proteasome is essential for gametogenesis, sporophyte development, and complex assembly in Arabidopsis. *Plant Cell* **21**: 460-478.
- BROOMFIELD, S., B. L. CHOW and W. XIAO, 1998 MMS2, encoding a ubiquitin-conjugating-enzyme-like protein, is a member of the yeast error-free postreplication repair pathway. *Proc Natl Acad Sci U S A* **95**: 5678-5683.

- BURKE, T. J., J. CALLIS and R. D. VIERSTRA, 1988 Characterization of a polyubiquitin gene from *Arabidopsis thaliana*. *Mol Gen Genet* **213**: 435-443.
- CADWELL, K., and L. COSCOY, 2005 Ubiquitination on nonlysine residues by a viral E3 ubiquitin ligase. *Science* **309**: 127-130.
- CALLIS, J., T. CARPENTER, C. W. SUN and R. D. VIERSTRA, 1995 Structure and evolution of genes encoding polyubiquitin and ubiquitin-like proteins in *Arabidopsis thaliana* ecotype Columbia. *Genetics* **139**: 921-939.
- CARRANO, A. C., E. EYTAN, A. HERSHKO and M. PAGANO, 1999 SKP2 is required for ubiquitin-mediated degradation of the CDK inhibitor p27. *Nat Cell Biol* **1**: 193-199.
- CHAE, H. S., F. FAURE and J. J. KIEBER, 2003 The *eto1*, *eto2*, and *eto3* mutations and cytokinin treatment increase ethylene biosynthesis in *Arabidopsis* by increasing the stability of ACS protein. *Plant Cell* **15**: 545-559.
- CHAU, V., J. W. TOBIAS, A. BACHMAIR, D. MARRIOTT, D. J. ECKER *et al.*, 1989 A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* **243**: 1576-1583.
- CHEN, Y. F., N. ETHERIDGE and G. E. SCHALLER, 2005 Ethylene signal transduction. *Ann Bot* **95**: 901-915.
- CHO, S. K., M. Y. RYU, C. SONG, J. M. KWAK and W. T. KIM, 2008 *Arabidopsis* PUB22 and PUB23 are homologous U-Box E3 ubiquitin ligases that play combinatory roles in response to drought stress. *Plant Cell* **20**: 1899-1914.
- CIECHANOVER, A., S. FERBER, D. GANOTH, S. ELIAS, A. HERSHKO *et al.*, 1988 Purification and characterization of arginyl-tRNA-protein transferase from rabbit reticulocytes. Its involvement in post-translational modification and degradation of acidic NH₂ termini substrates of the ubiquitin pathway. *J Biol Chem* **263**: 11155-11167.
- COOK, W. J., L. C. JEFFREY, M. L. SULLIVAN and R. D. VIERSTRA, 1992 Three-dimensional structure of a ubiquitin-conjugating enzyme (E2). *J Biol Chem* **267**: 15116-15121.
- COOK, W. J., L. C. JEFFREY, Y. XU and V. CHAU, 1993 Tertiary structures of class I ubiquitin-conjugating enzymes are highly conserved: crystal structure of yeast Ubc4. *Biochemistry* **32**: 13809-13817.
- DAI, Q., C. ZHANG, Y. WU, H. MCDONOUGH, R. A. WHALEY *et al.*, 2003 CHIP activates HSF1 and confers protection against apoptosis and cellular stress. *EMBO J* **22**: 5446-5458.
- DE TORRES, M., J. W. MANSFIELD, N. GRABOV, I. R. BROWN, H. AMMOUNEH *et al.*, 2006 *Pseudomonas syringae* effector AvrPtoB suppresses basal defence in *Arabidopsis*. *Plant J* **47**: 368-382.
- DITZEL, M., R. WILSON, T. TENEV, A. ZACHARIOU, A. PAUL *et al.*, 2003 Degradation of DIAP1 by the N-end rule pathway is essential for regulating apoptosis. *Nat Cell Biol* **5**: 467-473.
- DOELLING, J. H., N. YAN, J. KUREPA, J. WALKER and R. D. VIERSTRA, 2001 The ubiquitin-specific protease UBP14 is essential for early embryo development in *Arabidopsis thaliana*. *Plant J* **27**: 393-405.
- DOHERTY, F. J., S. DAWSON and R. J. MAYER, 2002 The ubiquitin-proteasome pathway of intracellular proteolysis. *Essays Biochem* **38**: 51-63.

- DOWNES, B. P., R. M. STUPAR, D. J. GINGERICH and R. D. VIERSTRA, 2003 The HECT ubiquitin-protein ligase (UPL) family in Arabidopsis: UPL3 has a specific role in trichome development. *Plant J* **35**: 729-742.
- DREHER, K., and J. CALLIS, 2007 Ubiquitin, hormones and biotic stress in plants. *Ann Bot* **99**: 787-822.
- DUNCAN, L. M., S. PIPER, R. B. DODD, M. K. SAVILLE, C. M. SANDERSON *et al.*, 2006 Lysine-63-linked ubiquitination is required for endolysosomal degradation of class I molecules. *EMBO J* **25**: 1635-1645.
- EWAN, R., R. PANGESTUTI, S. THORNER, A. CRAIG, C. CARR *et al.*, 2011 Deubiquitinating enzymes AtUBP12 and AtUBP13 and their tobacco homologue NtUBP12 are negative regulators of plant immunity. *New Phytol* **191**: 92-106.
- FINLEY, D., 2009 Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu Rev Biochem* **78**: 477-513.
- FINLEY, D., S. SADIS, B. P. MONIA, P. BOUCHER, D. J. ECKER *et al.*, 1994 Inhibition of proteolysis and cell cycle progression in a multiubiquitination-deficient yeast mutant. *Mol Cell Biol* **14**: 5501-5509.
- FROTTIN, F., A. MARTINEZ, P. PEYNOT, S. MITRA, R. C. HOLZ *et al.*, 2006 The proteomics of N-terminal methionine cleavage. *Mol Cell Proteomics* **5**: 2336-2349.
- FU, H., P. A. GIROD, J. H. DOELLING, S. VAN NOCKER, M. HOCHSTRASSER *et al.*, 1999 Structure and functional analysis of the 26S proteasome subunits from plants. *Mol Biol Rep* **26**: 137-146.
- FU, H., S. SADIS, D. M. RUBIN, M. GLICKMAN, S. VAN NOCKER *et al.*, 1998 Multiubiquitin chain binding and protein degradation are mediated by distinct domains within the 26 S proteasome subunit Mcl1. *J Biol Chem* **273**: 1970-1981.
- GANOOTH, D., G. BORNSTEIN, T. K. KO, B. LARSEN, M. TYERS *et al.*, 2001 The cell-cycle regulatory protein Cks1 is required for SCF(Skp2)-mediated ubiquitylation of p27. *Nat Cell Biol* **3**: 321-324.
- GARZON, M., K. EIFLER, A. FAUST, H. SCHEEL, K. HOFMANN *et al.*, 2007 PRT6/At5g02310 encodes an Arabidopsis ubiquitin ligase of the N-end rule pathway with arginine specificity and is not the CER3 locus. *FEBS Lett* **581**: 3189-3196.
- GIEFFERS, C., P. DUBE, J. R. HARRIS, H. STARK and J. M. PETERS, 2001 Three-dimensional structure of the anaphase-promoting complex. *Mol Cell* **7**: 907-913.
- GINGERICH, D. J., J. M. GAGNE, D. W. SALTER, H. HELLMANN, M. ESTELLE *et al.*, 2005 Cullins 3a and 3b assemble with members of the broad complex/tramtrack/bric-a-brac (BTB) protein family to form essential ubiquitin-protein ligases (E3s) in Arabidopsis. *J Biol Chem* **280**: 18810-18821.
- GIROD, P. A., T. B. CARPENTER, S. VAN NOCKER, M. L. SULLIVAN and R. D. VIERSTRA, 1993 Homologs of the essential ubiquitin conjugating enzymes UBC1, 4, and 5 in yeast are encoded by a multigene family in *Arabidopsis thaliana*. *Plant J* **3**: 545-552.
- GLICKMAN, M. H., and A. CIECHANOVER, 2002 The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol Rev* **82**: 373-428.
- GONDA, D. K., A. BACHMAIR, I. WUNNING, J. W. TOBIAS, W. S. LANE *et al.*, 1989 Universality and structure of the N-end rule. *J Biol Chem* **264**: 16700-16712.

- GORITSCHNIG, S., Y. ZHANG and X. LI, 2007 The ubiquitin pathway is required for innate immunity in Arabidopsis. *Plant J* **49**: 540-551.
- GRAY, W. M., J. C. DEL POZO, L. WALKER, L. HOBBIE, E. RISSEEUW *et al.*, 1999 Identification of an SCF ubiquitin-ligase complex required for auxin response in *Arabidopsis thaliana*. *Genes Dev* **13**: 1678-1691.
- GRIGORYEV, S., A. E. STEWART, Y. T. KWON, S. M. ARFIN, R. A. BRADSHAW *et al.*, 1996 A mouse amidase specific for N-terminal asparagine. The gene, the enzyme, and their function in the N-end rule pathway. *J Biol Chem* **271**: 28521-28532.
- GROLL, M., L. DITZEL, J. LOWE, D. STOCK, M. BOCHTLER *et al.*, 1997 Structure of 20S proteasome from yeast at 2.4 Å resolution. *Nature* **386**: 463-471.
- HARDTKE, C. S., H. OKAMOTO, C. STOOP-MYER and X. W. DENG, 2002 Biochemical evidence for ubiquitin ligase activity of the Arabidopsis COP1 interacting protein 8 (CIP8). *Plant J* **30**: 385-394.
- HATAKEYAMA, S., M. YADA, M. MATSUMOTO, N. ISHIDA and K. I. NAKAYAMA, 2001 U box proteins as a new family of ubiquitin-protein ligases. *J Biol Chem* **276**: 33111-33120.
- HATFIELD, P. M., J. CALLIS and R. D. VIERSTRA, 1990 Cloning of ubiquitin activating enzyme from wheat and expression of a functional protein in *Escherichia coli*. *J Biol Chem* **265**: 15813-15817.
- HATFIELD, P. M., M. M. GOSINK, T. B. CARPENTER and R. D. VIERSTRA, 1997 The ubiquitin-activating enzyme (E1) gene family in *Arabidopsis thaliana*. *Plant J* **11**: 213-226.
- HATFIELD, P. M., and R. D. VIERSTRA, 1992 Multiple forms of ubiquitin-activating enzyme E1 from wheat. Identification of an essential cysteine by in vitro mutagenesis. *J Biol Chem* **267**: 14799-14803.
- HERSHKO, A., and A. CIECHANOVER, 1998 The ubiquitin system. *Annu Rev Biochem* **67**: 425-479.
- HERSHKO, A., H. HELLER, E. EYTAN and Y. REISS, 1986 The protein substrate binding site of the ubiquitin-protein ligase system. *J Biol Chem* **261**: 11992-11999.
- HICKE, L., 2001 Protein regulation by monoubiquitin. *Nat Rev Mol Cell Biol* **2**: 195-201.
- HICKE, L., and R. DUNN, 2003 Regulation of membrane protein transport by ubiquitin and ubiquitin-binding proteins. *Annu Rev Cell Dev Biol* **19**: 141-172.
- HIRAMATSU, Y., K. KITAGAWA, T. SUZUKI, C. UCHIDA, T. HATTORI *et al.*, 2006 Degradation of Tob1 mediated by SCFSkp2-dependent ubiquitination. *Cancer Res* **66**: 8477-8483.
- HOFMANN, R. M., and C. M. PICKART, 1999 Noncanonical MMS2-encoded ubiquitin-conjugating enzyme functions in assembly of novel polyubiquitin chains for DNA repair. *Cell* **96**: 645-653.
- HOFMANN, R. M., and C. M. PICKART, 2001 In vitro assembly and recognition of Lys-63 polyubiquitin chains. *J Biol Chem* **276**: 27936-27943.
- HOLMAN, T. J., P. D. JONES, L. RUSSELL, A. MEDHURST, S. UBEDA TOMAS *et al.*, 2009 The N-end rule pathway promotes seed germination and establishment through removal of ABA sensitivity in Arabidopsis. *Proc Natl Acad Sci U S A* **106**: 4549-4554.
- HONDA, R., H. TANAKA and H. YASUDA, 1997 Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53. *FEBS Lett* **420**: 25-27.
- HU, R. G., J. SHENG, X. QI, Z. XU, T. T. TAKAHASHI *et al.*, 2005 The N-end rule pathway as a nitric oxide sensor controlling the levels of multiple regulators. *Nature* **437**: 981-986.

- HUA, Z., and R. D. VIERSTRA, 2011 The cullin-RING ubiquitin-protein ligases. *Annu Rev Plant Biol* **62**: 299-334.
- HUIBREGTSE, J. M., M. SCHEFFNER, S. BEAUDENON and P. M. HOWLEY, 1995 A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. *Proc Natl Acad Sci U S A* **92**: 5249.
- HUSNJAK, K., S. ELSASSER, N. ZHANG, X. CHEN, L. RANGLES *et al.*, 2008 Proteasome subunit Rpn13 is a novel ubiquitin receptor. *Nature* **453**: 481-488.
- HWANG, C. S., A. SHEMORRY and A. VARSHAVSKY, 2010 N-terminal acetylation of cellular proteins creates specific degradation signals. *Science* **327**: 973-977.
- IKEDA, F., and I. DIKIC, 2008 Atypical ubiquitin chains: new molecular signals. 'Protein Modifications: Beyond the Usual Suspects' review series. *EMBO Rep* **9**: 536-542.
- INADA, S., M. OHGISHI, T. MAYAMA, K. OKADA and T. SAKAI, 2004 RPT2 is a signal transducer involved in phototropic response and stomatal opening by association with phototropin 1 in *Arabidopsis thaliana*. *Plant Cell* **16**: 887-896.
- JANJUSEVIC, R., R. B. ABRAMOVITCH, G. B. MARTIN and C. E. STEBBINS, 2006 A bacterial inhibitor of host programmed cell death defenses is an E3 ubiquitin ligase. *Science* **311**: 222-226.
- JENTSCH, S., 1992 The ubiquitin-conjugation system. *Annu Rev Genet* **26**: 179-207.
- JIANG, J., C. A. BALLINGER, Y. WU, Q. DAI, D. M. CYR *et al.*, 2001 CHIP is a U-box-dependent E3 ubiquitin ligase: identification of Hsc70 as a target for ubiquitylation. *J Biol Chem* **276**: 42938-42944.
- JOHNSTON, S. C., S. M. RIDDLE, R. E. COHEN and C. P. HILL, 1999 Structural basis for the specificity of ubiquitin C-terminal hydrolases. *EMBO J* **18**: 3877-3887.
- JOO, S., Y. LIU, A. LUETH and S. ZHANG, 2008 MAPK phosphorylation-induced stabilization of ACS6 protein is mediated by the non-catalytic C-terminal domain, which also contains the cis-determinant for rapid degradation by the 26S proteasome pathway. *Plant J* **54**: 129-140.
- KAMURA, T., T. HARA, S. KOTOSHIBA, M. YADA, N. ISHIDA *et al.*, 2003 Degradation of p57Kip2 mediated by SCFSkp2-dependent ubiquitylation. *Proc Natl Acad Sci U S A* **100**: 10231-10236.
- KAMURA, T., K. MAENAKA, S. KOTOSHIBA, M. MATSUMOTO, D. KOHDA *et al.*, 2004 VHL-box and SOCS-box domains determine binding specificity for Cul2-Rbx1 and Cul5-Rbx2 modules of ubiquitin ligases. *Genes Dev* **18**: 3055-3065.
- KANEI-ISHII, C., T. NOMURA, T. TAKAGI, N. WATANABE, K. I. NAKAYAMA *et al.*, 2008 Fbxw7 acts as an E3 ubiquitin ligase that targets c-Myb for nemo-like kinase (NLK)-induced degradation. *J Biol Chem* **283**: 30540-30548.
- KIRKPATRICK, D. S., N. A. HATHAWAY, J. HANNA, S. ELSASSER, J. RUSH *et al.*, 2006 Quantitative analysis of in vitro ubiquitinated cyclin B1 reveals complex chain topology. *Nat Cell Biol* **8**: 700-710.
- KITAGAWA, K., Y. HIRAMATSU, C. UCHIDA, T. ISOBE, T. HATTORI *et al.*, 2009 Fbw7 promotes ubiquitin-dependent degradation of c-Myb: involvement of GSK3-mediated phosphorylation of Thr-572 in mouse c-Myb. *Oncogene* **28**: 2393-2405.

- KOEGEL, M., T. HOPPE, S. SCHLENKER, H. D. ULRICH, T. U. MAYER *et al.*, 1999 A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. *Cell* **96**: 635-644.
- KOEPP, D. M., L. K. SCHAEFER, X. YE, K. KEYOMARSI, C. CHU *et al.*, 2001 Phosphorylation-dependent ubiquitination of cyclin E by the SCFFbw7 ubiquitin ligase. *Science* **294**: 173-177.
- KOSAREV, P., K. F. MAYER and C. S. HARDTKE, 2002 Evaluation and classification of RING-finger domains encoded by the Arabidopsis genome. *Genome Biol* **3**: RESEARCH0016.
- KOSTOVA, Z., and D. H. WOLF, 2003 For whom the bell tolls: protein quality control of the endoplasmic reticulum and the ubiquitin-proteasome connection. *EMBO J* **22**: 2309-2317.
- KRAFT, E., S. L. STONE, L. MA, N. SU, Y. GAO *et al.*, 2005 Genome analysis and functional characterization of the E2 and RING-type E3 ligase ubiquitination enzymes of Arabidopsis. *Plant Physiol* **139**: 1597-1611.
- LISSO, J., D. STEINHAUSER, T. ALTMANN, J. KOPKA and C. MUSSIG, 2005 Identification of brassinosteroid-related genes by means of transcript co-response analyses. *Nucleic Acids Res* **33**: 2685-2696.
- LU, Y., N. BEDARD, S. CHEVALIER and S. S. WING, 2011 Identification of distinctive patterns of USP19-mediated growth regulation in normal and malignant cells. *PLoS One* **6**: e15936.
- MANNICK, J. B., 2007 Regulation of apoptosis by protein S-nitrosylation. *Amino Acids* **32**: 523-526.
- MCCALL, C. M., J. HU and Y. XIONG, 2005 Recruiting substrates to cullin 4-dependent ubiquitin ligases by DDB1. *Cell Cycle* **4**: 27-29.
- MERCHANT, S., and L. BOGORAD, 1986 Rapid degradation of apoplastocyanin in Cu(II)-deficient cells of *Chlamydomonas reinhardtii*. *J Biol Chem* **261**: 15850-15853.
- MIAO, Y., and U. ZENTGRAF, 2010 A HECT E3 ubiquitin ligase negatively regulates Arabidopsis leaf senescence through degradation of the transcription factor WRKY53. *Plant J* **63**: 179-188.
- MICHAELY, P., and V. BENNETT, 1992 The ANK repeat: a ubiquitous motif involved in macromolecular recognition. *Trends Cell Biol* **2**: 127-129.
- MIWA, S., C. UCHIDA, K. KITAGAWA, T. HATTORI, T. ODA *et al.*, 2006 Mdm2-mediated pRB downregulation is involved in carcinogenesis in a p53-independent manner. *Biochem Biophys Res Commun* **340**: 54-61.
- MOGK, A., R. SCHMIDT and B. BUKAU, 2007 The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies. *Trends Cell Biol* **17**: 165-172.
- MORRIS, J. R., and E. SOLOMON, 2004 BRCA1 : BARD1 induces the formation of conjugated ubiquitin structures, dependent on K6 of ubiquitin, in cells during DNA replication and repair. *Hum Mol Genet* **13**: 807-817.
- MOTCHOULSKI, A., and E. LISCUM, 1999 Arabidopsis NPH3: A NPH1 photoreceptor-interacting protein essential for phototropism. *Science* **286**: 961-964.

- MUDGIL, Y., S. H. SHIU, S. L. STONE, J. N. SALT and D. R. GORING, 2004 A large complement of the predicted Arabidopsis ARM repeat proteins are members of the U-box E3 ubiquitin ligase family. *Plant Physiol* **134**: 59-66.
- MUKHOPADHYAY, D., and H. RIEZMAN, 2007 Proteasome-independent functions of ubiquitin in endocytosis and signaling. *Science* **315**: 201-205.
- NAKAYAMA, K., H. NAGAHAMA, Y. A. MINAMISHIMA, M. MATSUMOTO, I. NAKAMICHI *et al.*, 2000 Targeted disruption of Skp2 results in accumulation of cyclin E and p27(Kip1), polyploidy and centrosome overduplication. *EMBO J* **19**: 2069-2081.
- NISHIKAWA, H., S. OOKA, K. SATO, K. ARIMA, J. OKAMOTO *et al.*, 2004 Mass spectrometric and mutational analyses reveal Lys-6-linked polyubiquitin chains catalyzed by BRCA1-BARD1 ubiquitin ligase. *J Biol Chem* **279**: 3916-3924.
- PEDMALE, U. V., and E. LISCUM, 2007 Regulation of phototropic signaling in Arabidopsis via phosphorylation state changes in the phototropin 1-interacting protein NPH3. *J Biol Chem* **282**: 19992-20001.
- PENG, J., D. SCHWARTZ, J. E. ELIAS, C. C. THOREEN, D. CHENG *et al.*, 2003 A proteomics approach to understanding protein ubiquitination. *Nat Biotechnol* **21**: 921-926.
- PERAZZA, D., G. VACHON and M. HERZOG, 1998 Gibberellins promote trichome formation by Up-regulating GLABROUS1 in arabidopsis. *Plant Physiol* **117**: 375-383.
- PERSSON, S., H. WEI, J. MILNE, G. P. PAGE and C. R. SOMERVILLE, 2005 Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc Natl Acad Sci U S A* **102**: 8633-8638.
- PICKART, C. M., 2001 Mechanisms underlying ubiquitination. *Annu Rev Biochem* **70**: 503-533.
- PICKART, C. M., and D. FUSHMAN, 2004 Polyubiquitin chains: polymeric protein signals. *Curr Opin Chem Biol* **8**: 610-616.
- PINTARD, L., A. WILLEMS and M. PETER, 2004 Cullin-based ubiquitin ligases: Cul3-BTB complexes join the family. *EMBO J* **23**: 1681-1687.
- PINTARD, L., J. H. WILLIS, A. WILLEMS, J. L. JOHNSON, M. SRAYKO *et al.*, 2003 The BTB protein MEL-26 is a substrate-specific adaptor of the CUL-3 ubiquitin-ligase. *Nature* **425**: 311-316.
- POTUSCHAK, T., S. STARY, P. SCHLOGELHOFER, F. BECKER, V. NEJNSKAIA *et al.*, 1998 PRT1 of Arabidopsis thaliana encodes a component of the plant N-end rule pathway. *Proc Natl Acad Sci U S A* **95**: 7904-7908.
- QIAN, S. B., H. MCDONOUGH, F. BOELLMANN, D. M. CYR and C. PATTERSON, 2006 CHIP-mediated stress recovery by sequential ubiquitination of substrates and Hsp70. *Nature* **440**: 551-555.
- RAAB, S., G. DRECHSEL, M. ZAREPOUR, W. HARTUNG, T. KOSHIBA *et al.*, 2009 Identification of a novel E3 ubiquitin ligase that is required for suppression of premature senescence in Arabidopsis. *Plant J* **59**: 39-51.
- RAO, H., F. UHLMANN, K. NASMYTH and A. VARSHAVSKY, 2001 Degradation of a cohesin subunit by the N-end rule pathway is essential for chromosome stability. *Nature* **410**: 955-959.

- RENAUDIN, J. P., J. H. DOONAN, D. FREEMAN, J. HASHIMOTO, H. HIRT *et al.*, 1996 Plant cyclins: a unified nomenclature for plant A-, B- and D-type cyclins based on sequence organization. *Plant Mol Biol* **32**: 1003-1018.
- RIOS, G., A. LOSSOW, B. HERTEL, F. BREUER, S. SCHAEFER *et al.*, 2002 Rapid identification of Arabidopsis insertion mutants by non-radioactive detection of T-DNA tagged genes. *Plant J* **32**: 243-253.
- RISSEEUW, E. P., T. E. DASKALCHUK, T. W. BANKS, E. LIU, J. COTELESAGE *et al.*, 2003 Protein interaction analysis of SCF ubiquitin E3 ligase subunits from Arabidopsis. *Plant J* **34**: 753-767.
- SAKAI, T., T. WADA, S. ISHIGURO and K. OKADA, 2000 RPT2. A signal transducer of the phototropic response in Arabidopsis. *Plant Cell* **12**: 225-236.
- SAMUEL, M. A., Y. MUDGIL, J. N. SALT, F. DELMAS, S. RAMACHANDRAN *et al.*, 2008 Interactions between the S-domain receptor kinases and AtPUB-ARM E3 ubiquitin ligases suggest a conserved signaling pathway in Arabidopsis. *Plant Physiol* **147**: 2084-2095.
- SAMUEL, M. A., J. N. SALT, S. H. SHIU and D. R. GORING, 2006 Multifunctional arm repeat domains in plants. *Int Rev Cytol* **253**: 1-26.
- SANCHO, E., M. R. VILA, L. SANCHEZ-PULIDO, J. J. LOZANO, R. PACIUCCI *et al.*, 1998 Role of UEV-1, an inactive variant of the E2 ubiquitin-conjugating enzymes, in in vitro differentiation and cell cycle behavior of HT-29-M6 intestinal mucosecretory cells. *Mol Cell Biol* **18**: 576-589.
- SCHEFFNER, M., U. NUBER and J. M. HUIBREGTSE, 1995 Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade. *Nature* **373**: 81-83.
- SCHLOGELHOFER, P., M. GARZON, C. KERZENDORFER, V. NIZHYNKA and A. BACHMAIR, 2006 Expression of the ubiquitin variant ubR48 decreases proteolytic activity in Arabidopsis and induces cell death. *Planta* **223**: 684-697.
- SCHMIDT, M., J. HANNA, S. ELSASSER and D. FINLEY, 2005 Proteasome-associated proteins: regulation of a proteolytic machine. *Biol Chem* **386**: 725-737.
- SCHNELL, J. D., and L. HICKE, 2003 Non-traditional functions of ubiquitin and ubiquitin-binding proteins. *J Biol Chem* **278**: 35857-35860.
- SCHWARZ, S. E., J. L. ROSA and M. SCHEFFNER, 1998 Characterization of human hect domain family members and their interaction with UbcH5 and UbcH7. *J Biol Chem* **273**: 12148-12154.
- SEO, H. S., J. Y. YANG, M. ISHIKAWA, C. BOLLE, M. L. BALLESTEROS *et al.*, 2003 LAF1 ubiquitination by COP1 controls photomorphogenesis and is stimulated by SPA1. *Nature* **423**: 995-999.
- SEOL, J. H., R. M. FELDMAN, W. ZACHARIAE, A. SHEVCHENKO, C. C. CORRELL *et al.*, 1999 Cdc53/cullin and the essential Hrt1 RING-H2 subunit of SCF define a ubiquitin ligase module that activates the E2 enzyme Cdc34. *Genes Dev* **13**: 1614-1626.
- SHEARD, L. B., X. TAN, H. MAO, J. WITHERS, G. BEN-NISSAN *et al.*, 2010 Jasmonate perception by inositol-phosphate-potentiated COI1-JAZ co-receptor. *Nature* **468**: 400-405.

- SHEN, G., Z. ADAM and H. ZHANG, 2007a The E3 ligase AtCHIP ubiquitylates FtsH1, a component of the chloroplast FtsH protease, and affects protein degradation in chloroplasts. *Plant J* **52**: 309-321.
- SHEN, G., J. YAN, V. PASAPULA, J. LUO, C. HE *et al.*, 2007b The chloroplast protease subunit ClpP4 is a substrate of the E3 ligase AtCHIP and plays an important role in chloroplast function. *Plant J* **49**: 228-237.
- SHEN, W. H., Y. PARMENTIER, H. HELLMANN, E. LECHNER, A. DONG *et al.*, 2002 Null mutation of AtCUL1 causes arrest in early embryogenesis in Arabidopsis. *Mol Biol Cell* **13**: 1916-1928.
- SMALLE, J., J. KUREPA, P. YANG, E. BABYCHUK, S. KUSHNIR *et al.*, 2002 Cytokinin growth responses in Arabidopsis involve the 26S proteasome subunit RPN12. *Plant Cell* **14**: 17-32.
- SMALLE, J., J. KUREPA, P. YANG, T. J. EMBORG, E. BABYCHUK *et al.*, 2003 The pleiotropic role of the 26S proteasome subunit RPN10 in Arabidopsis growth and development supports a substrate-specific function in abscisic acid signaling. *Plant Cell* **15**: 965-980.
- SMALLE, J., and R. D. VIERSTRA, 2004 The ubiquitin 26S proteasome proteolytic pathway. *Annu Rev Plant Biol* **55**: 555-590.
- SPENCE, J., R. R. GALI, G. DITTMAR, F. SHERMAN, M. KARIN *et al.*, 2000 Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain. *Cell* **102**: 67-76.
- STARY, S., X. J. YIN, T. POTUSCHAK, P. SCHLOGELHOFER, V. NIZHYNKA *et al.*, 2003 PRT1 of Arabidopsis is a ubiquitin protein ligase of the plant N-end rule pathway with specificity for aromatic amino-terminal residues. *Plant Physiol* **133**: 1360-1366.
- STEINHAUSER, D., B. USADEL, A. LUEDEMANN, O. THIMM and J. KOPKA, 2004 CSB.DB: a comprehensive systems-biology database. *Bioinformatics* **20**: 3647-3651.
- STONE, S. L., E. M. ANDERSON, R. T. MULLEN and D. R. GORING, 2003 ARC1 is an E3 ubiquitin ligase and promotes the ubiquitination of proteins during the rejection of self-incompatible Brassica pollen. *Plant Cell* **15**: 885-898.
- TAN, X., L. I. CALDERON-VILLALOBOS, M. SHARON, C. ZHENG, C. V. ROBINSON *et al.*, 2007 Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* **446**: 640-645.
- TANG, Z., B. LI, R. BHARADWAJ, H. ZHU, E. OZKAN *et al.*, 2001 APC2 Cullin protein and APC11 RING protein comprise the minimal ubiquitin ligase module of the anaphase-promoting complex. *Mol Biol Cell* **12**: 3839-3851.
- TASAKI, T., L. C. MULDER, A. IWAMATSU, M. J. LEE, I. V. DAVYDOV *et al.*, 2005 A family of mammalian E3 ubiquitin ligases that contain the UBR box motif and recognize N-degrons. *Mol Cell Biol* **25**: 7120-7136.
- TEDESCO, D., J. LUKAS and S. I. REED, 2002 The pRb-related protein p130 is regulated by phosphorylation-dependent proteolysis via the protein-ubiquitin ligase SCF(Skp2). *Genes Dev* **16**: 2946-2957.
- THIMM, O., O. BLASING, Y. GIBON, A. NAGEL, S. MEYER *et al.*, 2004 MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J* **37**: 914-939.
- THROWER, J. S., L. HOFFMAN, M. RECHSTEINER and C. M. PICKART, 2000 Recognition of the polyubiquitin proteolytic signal. *EMBO J* **19**: 94-102.

- TRUJILLO, M., K. ICHIMURA, C. CASAIS and K. SHIRASU, 2008 Negative regulation of PAMP-triggered immunity by an E3 ubiquitin ligase triplet in Arabidopsis. *Curr Biol* **18**: 1396-1401.
- TURNER, G. C., F. DU and A. VARSHAVSKY, 2000 Peptides accelerate their uptake by activating a ubiquitin-dependent proteolytic pathway. *Nature* **405**: 579-583.
- UCHIDA, C., S. MIWA, K. KITAGAWA, T. HATTORI, T. ISOBE *et al.*, 2005 Enhanced Mdm2 activity inhibits pRB function via ubiquitin-dependent degradation. *EMBO J* **24**: 160-169.
- VANDEMARK, A. P., R. M. HOFMANN, C. TSUI, C. M. PICKART and C. WOLBERGER, 2001 Molecular insights into polyubiquitin chain assembly: crystal structure of the Mms2/Ubc13 heterodimer. *Cell* **105**: 711-720.
- VARSHAVSKY, A., G. TURNER, F. DU and Y. XIE, 2000 Felix Hoppe-Seyler Lecture 2000. The ubiquitin system and the N-end rule pathway. *Biol Chem* **381**: 779-789.
- VIERSTRA, R. D., 1996 Proteolysis in plants: mechanisms and functions. *Plant Mol Biol* **32**: 275-302.
- VIERSTRA, R. D., 2009 The ubiquitin-26S proteasome system at the nexus of plant biology. *Nat Rev Mol Cell Biol* **10**: 385-397.
- VIJAY-KUMAR, S., C. E. BUGG, K. D. WILKINSON, R. D. VIERSTRA, P. M. HATFIELD *et al.*, 1987 Comparison of the three-dimensional structures of human, yeast, and oat ubiquitin. *J Biol Chem* **262**: 6396-6399.
- WANG, H., K. I. PIATKOV, C. S. BROWER and A. VARSHAVSKY, 2009 Glutamine-specific N-terminal amidase, a component of the N-end rule pathway. *Mol Cell* **34**: 686-695.
- WANG, K. L., H. YOSHIDA, C. LURIN and J. R. ECKER, 2004 Regulation of ethylene gas biosynthesis by the Arabidopsis ETO1 protein. *Nature* **428**: 945-950.
- WARD, S. P., and O. LEYSER, 2004 Shoot branching. *Curr Opin Plant Biol* **7**: 73-78.
- WEI, W., J. JIN, S. SCHLISIO, J. W. HARPER and W. G. KAELIN, JR., 2005 The v-Jun point mutation allows c-Jun to escape GSK3-dependent recognition and destruction by the Fbw7 ubiquitin ligase. *Cancer Cell* **8**: 25-33.
- WEISSMAN, A. M., 2001 Themes and variations on ubiquitylation. *Nat Rev Mol Cell Biol* **2**: 169-178.
- WELCKER, M., A. ORIAN, J. JIN, J. E. GRIM, J. W. HARPER *et al.*, 2004 The Fbw7 tumor suppressor regulates glycogen synthase kinase 3 phosphorylation-dependent c-Myc protein degradation. *Proc Natl Acad Sci U S A* **101**: 9085-9090.
- WERTZ, I. E., K. M. O'ROURKE, Z. ZHANG, D. DORNAN, D. ARNOTT *et al.*, 2004 Human De-etiolated-1 regulates c-Jun by assembling a CUL4A ubiquitin ligase. *Science* **303**: 1371-1374.
- WIBORG, J., C. O'SHEA and K. SKRIVER, 2008 Biochemical function of typical and variant *Arabidopsis thaliana* U-box E3 ubiquitin-protein ligases. *Biochem J* **413**: 447-457.
- WILKINSON, K. D., 2000 Ubiquitination and deubiquitination: targeting of proteins for degradation by the proteasome. *Semin Cell Dev Biol* **11**: 141-148.
- WING, S. S., 2003 Deubiquitinating enzymes--the importance of driving in reverse along the ubiquitin-proteasome pathway. *Int J Biochem Cell Biol* **35**: 590-605.
- WOO, H. R., K. M. CHUNG, J. H. PARK, S. A. OH, T. AHN *et al.*, 2001 ORE9, an F-box protein that regulates leaf senescence in Arabidopsis. *Plant Cell* **13**: 1779-1790.

- WU, G., S. LYAPINA, I. DAS, J. LI, M. GURNEY *et al.*, 2001 SEL-10 is an inhibitor of notch signaling that targets notch for ubiquitin-mediated protein degradation. *Mol Cell Biol* **21**: 7403-7415.
- XIE, Q., H. S. GUO, G. DALLMAN, S. FANG, A. M. WEISSMAN *et al.*, 2002 SINAT5 promotes ubiquitin-related degradation of NAC1 to attenuate auxin signals. *Nature* **419**: 167-170.
- XU, L., Y. WEI, J. REBOUL, P. VAGLIO, T. H. SHIN *et al.*, 2003 BTB proteins are substrate-specific adaptors in an SCF-like modular ubiquitin ligase containing CUL-3. *Nature* **425**: 316-321.
- YADA, M., S. HATAKEYAMA, T. KAMURA, M. NISHIYAMA, R. TSUNEMATSU *et al.*, 2004 Phosphorylation-dependent degradation of c-Myc is mediated by the F-box protein Fbw7. *EMBO J* **23**: 2116-2125.
- YAN, J., J. WANG, Q. LI, J. R. HWANG, C. PATTERSON *et al.*, 2003 AtCHIP, a U-box-containing E3 ubiquitin ligase, plays a critical role in temperature stress tolerance in Arabidopsis. *Plant Physiol* **132**: 861-869.
- YAN, N., J. H. DOELLING, T. G. FALBEL, A. M. DURSKI and R. D. VIERSTRA, 2000 The ubiquitin-specific protease family from Arabidopsis. AtUBP1 and 2 are required for the resistance to the amino acid analog canavanine. *Plant Physiol* **124**: 1828-1843.
- YANAGAWA, Y., J. A. SULLIVAN, S. KOMATSU, G. GUSMAROLI, G. SUZUKI *et al.*, 2004 Arabidopsis COP10 forms a complex with DDB1 and DET1 in vivo and enhances the activity of ubiquitin conjugating enzymes. *Genes Dev* **18**: 2172-2181.
- YANG, P., H. FU, J. WALKER, C. M. PAPA, J. SMALLE *et al.*, 2004 Purification of the Arabidopsis 26 S proteasome: biochemical and molecular analyses revealed the presence of multiple isoforms. *J Biol Chem* **279**: 6401-6413.
- YI, C., and X. W. DENG, 2005 COP1 - from plant photomorphogenesis to mammalian tumorigenesis. *Trends Cell Biol* **15**: 618-625.
- YOO, S. D., Y. H. CHO, G. TENA, Y. XIONG and J. SHEEN, 2008 Dual control of nuclear EIN3 by bifurcate MAPK cascades in C2H4 signalling. *Nature* **451**: 789-795.
- YOSHIDA, S., M. ITO, J. CALLIS, I. NISHIDA and A. WATANABE, 2002 A delayed leaf senescence mutant is defective in arginyl-tRNA:protein arginyltransferase, a component of the N-end rule pathway in Arabidopsis. *Plant J* **32**: 129-137.
- ZENG, L. R., S. QU, A. BORDEOS, C. YANG, M. BARAOIDAN *et al.*, 2004 Spotted leaf11, a negative regulator of plant cell death and defense, encodes a U-box/armadillo repeat protein endowed with E3 ubiquitin ligase activity. *Plant Cell* **16**: 2795-2808.
- ZIMMERMANN, P., M. HIRSCH-HOFFMANN, L. HENNIG and W. GRUISSEM, 2004 GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiol* **136**: 2621-2632.

EIDESSTATTLICHE ERKLÄRUNG

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit, einschließlich Tabellen, Karten und Abbildungen, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie noch nicht veröffentlicht worden ist, sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Andreas Bachmair und Prof. Dr. George Coupland (Max-Planck-Institut für Pflanzenzüchtungsforschung) betreut worden.

Prabhavathi Talloji

Köln, d. 02.11.2011

CURRICULUM VITAE

Name: **Prabhavathi Talloji**
Geburtsdatum: 04 August 1975
Geburtsort: Metpally Indien
Nationalität: Indian
Email: talloji@mpiz-koeln.mpg.de
Prabhavathi.talloji@univie.ac.at

Dec 2006- Dec 2011 **PhD-Max-Planck-Institut für Züchtungsforschung, Köln, Deutschland / Max Perutz Laboratory, Wien**, in der Gruppe von Prof. Andreas Bachmair

Forschungsthema: Identification of novel components and links in ubiquitin dependent protein degradation pathways of *Arabidopsis thaliana*

2004 – 2006 Masters in molecular biology, Umea University, Umea, Sweden
2000 – 2004 Teaching, Intermediate college, India
1998 – 2000 M.Sc Botany, Osmania University, India
1997 – 1998 Bachelor of education, Osmania University, India
1996 – 1997 Teaching in highschool, India
1993 – 1996 Bachelor of science, University, India
1991 – 1993 Intermediate, India
1981 – 1991 School, India

Awards

Graduate Student Research Recognition Award, Selected for best graduate student abstract in the Metabolism/Biochemistry session.

ICAR- 2011 International conference on *Arabidopsis* Research, June 22nd-25th Madison, Wisconsin, USA.

Selected for a Talk at Society for Experimental Biology Annual Main Meeting (SEB).

SEB-2010, Prague, 30th June-3rd July, CZ.

Best Poster Award-3rd

1st Joint Retreat of the PhD Students in Experimental Plant Sciences, October 2nd -3rd Wageningen, The Netherlands.

Publikation, Konferenzen und Präsentationen

Holman TJ, Jones PD, Russell L, Medhurst A, Ubeda Tomás S, **Talloji P**, Marquez J, Schmuths H, Tung SA, Taylor I, Footitt S, Bachmair A, Theodoulou FL, Holdsworth MJ., 2009. The N-end rule pathway promotes seed germination and establishment through removal of ABA sensitivity in *Arabidopsis*. Proc Natl Acad Sci U S A. 106(11):4549-54.

Prabhavathi Talloji and Andreas Bachmair: Analysis of N-end rule pathway components in *Arabidopsis*: PRT8 is a novel E3 ubiquitin ligase that targets proteins with aliphatic hydrophobic amino terminal residue; 22nd ICAR meeting 2011, Madison, USA.

Prabhavathi Talloji and Andreas Bachmair: Molecular mapping of Ubiquitin Variant (ub K48R) suppressor mutant in *Arabidopsis thaliana*; 7th TNAM meeting 2010, Salzburg, Austria.

Prabhavathi Talloji and Andreas Bachmair: Analysis of *Arabidopsis* mutants in N-end rule pathway; SEB Annual meeting 2010, Prague, CZ.

Prabhavathi Talloji, Sedlazeck FJ and Andreas Bachmair: Study of Ubiquitin Proteasome Dependent Proteolysis in *Arabidopsis thaliana*; Molecular Aspects of Plant Development Meeting 2010, Vienna, Austria.

Prabhavathi Talloji and Andreas Bachmair: Analysis of UBR domain proteins and the N-end rule pathway; EMBO meeting 2009, Riva del Garda, Italy.

Prabhavathi Talloji and Andreas Bachmair: N-end rule pathway in *Arabidopsis thaliana*; 1st Joint Retreat of Phd students in Experimental Plant Sciences 2008, Wageningen, The Netherlands.