Pankaj Barah

Integrative systems approaches to study plant stress responses

Thesis for the degree of Philosophiae Doctor

Trondheim, April 2013

Norwegian University of Science and Technology
Faculty of Natural Sciences and Technology
Department of Biology

NTNU – Trondheim
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the degree of Philosophiae Doctor

Faculty of Natural Sciences and Technology
Department of Biology

ISBN 978-82-471-3673-7 (printed ver.)
ISBN 978-82-471-3674-4 (electronic ver.)
ISSN 1503-8181

Doctoral theses at NTNU, 2013:93

Printed by NTNU-trykk

Dedicated to my beloved parents…

Table of contents

- Gene Regulatory Networks
- Co-expression Networks
- Challenges and future prospects

CHAPTER 3: **Common and specific stress response to biotic stresses**

- Overall changes of *Arabidopsis* transcriptome in response to an insect and a bacterial attack
- Analysis of over represented gene-ontologies (GO) in *Arabidopsis* indicates rigorous reprogramming of several biological processes
- Mapping insect and bacterial specific responses on pathways and processes using MapMan
    - Metabolism overview map
    - Comparative overview of response to biotic stress during aphid and *P. syringae* treatment
    - Regulation overview map
    - Difference in the jasmonic acid (JA) biosynthesis pathway during aphid and *P. syringae* treatment
    - Salicylic acid (SA) regulates expression of aphid specific defense proteins and methyl salicylate activates *P. syringae* specific defense proteins
    - Overview of differences in secondary metabolism
    - A large number of transcription factors are differentially regulated and many of are unique for insect or bacterial stress
- microRNA mediated post-transcriptional regulations during insect and bacterial attack
    - MicroRNA connection to genes from the common category
    - MicroRNA connection to genes showing aphid specific responses
    - MicroRNA connection to genes showing *P. syringae* specific responses
- Cross-validation of differentially regulated aphid and Pseudomonas specific transcription factors by co-expression analysis of multiple biotic stress dataset

CHAPTER 4: **Single stress vs. combined stress response in *Arabidopsis*.**

- Responses to single stresses and intraspecific variation
- Responses to combined stresses
- Weighted gene co-expression network analysis (WGCNA)
- Intra-modular connectivity and identification of hub genes
- Eigengene networks and module membership
    - Biotic stress (FLG) response module
    - Cold stress response module
    - Cold and high light stress response module
    - Abiotic stress response module

CHAPTER 5: **Intraspecific variation in stress response and adaptive plasticity among *Arabidopsis* ecotypes**

- Different transcriptome signatures of 10 *Arabidopsis thaliana* ecotypes to cold stress.
- Ecotype specific cold responsive transcription factors (TFs) and transposable element (TEs)
- Gene Ontology analysis of the differentially regulated transcripts
- Probable co-evolution of stress signaling genes in response to cold and pathogen
- CBF regulon genes exhibits differential expression pattern in *Arabidopsis* ecotypes during cold treatment
- Sequence Polymorphisms in the CBF genes using data from *Arabidopsis thaliana* 1001 Genome project
- Reconstructing a transcriptional regulatory network during cold stress response in *Arabidopsis thaliana*
- Transcription factor activity dynamics under cold stress in ten *Arabidopsis* ecotypes.
- Correlation between activity profile and expression values of TFs

CHAPTER 6: **Transcriptional regulatory network in *Arabidopsis thaliana* during response to single and combined stresses**

- *Arabidopsis* ecotypes have different transcriptome signatures while responding to 11 stress conditions
- List of multiple stress regulated transcription factors (TFs)
- Very few TFs were regulated in more than one stress conditions
- Overlaps with stress reference gene (benchmarking the experimental data)
- Reconstruction of initial co-regulatory network
- Differential activities of the transcription factors under 11 stress conditions
- Stress wise regulatory sub-network (indicating predicted activation/repression)
- Stress specific TFs

CHAPTER 7: From laboratory to field (Model organism to Cultivated crops)

[a] Improving crop productivity and abiotic stress tolerance in cultivated *Fragaria* (Strawberry) using omics and systems biology approach

- Abiotic factors and agronomic aspects
- Botany and agricultural history of cultivated *Fragaria*
- Abiotic factors in strawberry production
- *Fragaria* breeding toward abiotic factors
- Genomic approaches toward *Fragaria*
- Proteomic approaches toward *Fragaria*
- Metabolomic approaches toward *Fragaria.*
- Systems biology as suitable tool for crop improvement
- Omics data integration for improving plant productivity / translational research
- Plant/Crops systems biology
- Pathway modelling and the concept of 'Virtual Plant'
- Network-based approaches

[b] Activation of plant defense in *Brassica napus L. cv. Westar* and transgenic *MINELESS* plants after attack by *Mamestra brassicae* (cabbage moth)

- *M. brassicae* larvae perform better on wild-type seedlings
- *M. brassicae* feeding reduces myrosinase activity
- Transcriptional responses of *B. napus* wild-type and MINELESS plants after *M. brassicae* feeding
- Mapman analysis of gene regulation in wild-type and MINELESS seedlings after *M. brassicae* feeding.
- Validation of microarray data.

Abstract

The world's population is growing at an alarming rate. As of 2010, out of 7 billion people in the world, 925 million are hungry. It represents 13.1 percent of the total world population, or almost 1 in 7 people are hungry (FAO). Climate change is increasingly viewed as a current and future cause of hunger and poverty. In the scenario of global climatic change, different biotic and abiotic stresses are severe threats to the agricultural production worldwide. In nature, plants are continuously stressed by exposure to multiple adverse conditions. The combined effect of multiple biotic and abiotic stresses is a major yield-limiting factor in agriculture. In such a situation, it is of utmost importance to take initiatives for genome scale molecular understanding of stress response mechanisms in plants, so that new stress resistant crop varieties can be developed. Recent developments in omics technologies (metabolomic, proteomic, transcriptomic, phenomics and more) have opened up a new dimension for conducting genome scale molecular studies to understand stress response mechanisms in plants. These studies have led to the revelation of extremely complex and interacting networks of various stress response processes. Statistical, mathematical and informatics driven analysis and integration of the enormous amount of data produced is a challenge. The combination of high throughput profiling techniques, bioinformatics tools and the knowledge of genetics will provide the ways by which to achieve a comprehensive understanding of biological processes related to stress responses in plants. Such knowledge can be translated further to develop better crop varieties.

This thesis presents a few such integrated studies, exploring different aspects of plant stress responses at the molecular and systems levels. I believe that the works presented in this thesis will significantly contribute towards a molecular understanding of plant stress response mechanisms at the systems level. The entire thesis has been divided into seven chapters.

Chapter 1 gives a brief introduction about the adverse effect of global climatic change on plant productivity due to intensified effects of various stress factors and its negative socio-economic impact on human society. This chapter also briefly summarises the background of seven research papers presented in this thesis along with a review of contemporary works.

Chapter 2 (Paper I) describes why systems biology is useful to study plant stress biology, reviewing various approaches and computational tools available to plant biologists till date.

Chapter 3 (Paper II) explores common and stress specific response signatures by the host plant to two different biotic stresses. It provides a comparative understandings of *Arabidopsis – Brevicoryne brassicae (aphid)* and *Arabidopsis –Pseudomonas syringae(bacteria)* interactions at the systems level.

Chapter 4 (Paper III) uncovers the molecular stress response patterns in plants during the co-occurrence of multiple abiotic and biotic stresses. The main outcome is that transcriptome changes in response to combined stresses could not be predicted from the responses to single stress treatments. This chapter also presents a modular network topology based approach to identify functionally related stress responsive gene modules.

Chapter 5 (Paper IV) presents the intraspecific variation in stress response patterns among 10 *Arabidopsis* ecotypes during cold stress exposure. Using an *in silico* transcriptional regulatory network model during cellular responses to cold stress in *Arabidopsis thaliana,* a hypothesis is presented that differentially evolving regulatory networks play a crucial role in climate adaptation of plants.

Chapter 6 (Paper V) presents an *in silico* transcriptional regulatory network model in responses to 11 stresses (5 single and 6 combined) conditions in *Arabidopsis thaliana* re-constructed from microarray data using a robust algorithm - Network Component Analysis (NCA).

Chapter 7 presents two application cases as examples of translational research, how knowledge developed in lab can be used in crop plants.

a. **(Paper VI)** demonstrates how the omics and systems biology approach is useful in improving crop productivity and abiotic stress tolerance in cultivated *Fragaria.*
b. **(Paper VII)** presents a case study on developing transgenic *Brassica napus MINELESS* as a new model system to study plant insect interactions. During this study, activation of plant defense in *Brassica napus L. cv. Westar* and transgenic MINELESS plants after attack by *Mamestra brassicae* (cabbage moth) were analysed.

Acknowledgements

The work presented in this thesis was carried out at Cell Molecular and Genomics Group, Department of Biology, Norwegian University of Science and Technology, during the period of August 2009 to November 2012. This thesis presents integrated systems biology approaches towards the global understanding of molecular level stress response mechanism in plants. The project was financially supported by the Biotechnology and Functional Genomics (FUGE) program of the Norwegian Research Council through grants NFR 184147/S10, 185173/V40 and 184146/S10.

I sincerely acknowledge my mentor Prof. Atle M. Bones, for his expert guidance, generous support, and for providing moral boost throughout my research period. He was a real motivational guru during my PhD. I am grateful to him for giving me the full freedom to develop, implement and present new scientific concepts during my PhD work.

I heartfully thank my co-supervisors Prof. Martin Kuiper and Dr. Per Winge for their supports and advices. I cannot forget the personal help provided by Per Winge during the initial period of my stay in Norway.

I thank all the lovely members of 'Boneslab' for the friendly working environment. Being a computational biologist, it wouldn't have been possible for me to learn so much detail in plant biology without the generous help from them.

My special thanks to all the administrative staffs at 'Institutt for biologi' and 'NT-faculty', NTNU for their support which made my life easy and comfortable in a new country.

Biweekly 'Systems Biology journal club' at NTNU was an excellent forum for learning new developments in the interdisciplinary field of systems biology and discussing our own research. My thanks to all the participating members of this journal club, especially professors Martin Kuiper and Finn Drabløs for coordinating such an awesome platform.

I thank Prof. John Mundy, University of Copenhagen, Drs. Henrik Bjørn Nielsen and Simon Rasmussen, Center for Biological Sequence Analysis, Technical University of Denmark, for their collaborating support during ERANET-PG, Multistress data analysis.

I owe special thanks to all my co-authors in the presented manuscripts for their great team spirit.

I thank all my friends especially to those with whom I had spent great time in Trondheim during the PhD period.

I take this opportunity to thank few of my previous mentors Prof. R Sowdhamini (NCBS, Bangalore, India), Prof. Indira Ghosh (JNU, New-Delhi, India), Prof. Somdatta Sinha

(CCMB, Hyderabad, India; now in IISER Mohali), Dr. Ram Rup Sarkar (NCL, Pune, India) for their learned guidance during my initial period of research career. I pay my sincere respect to all my teachers back in schools and colleges, who helped me to stand here today.

My special thanks to Drs. Jens Rohloff, Ralph Kissen, and Per Winge for reading my thesis and giving valuable suggestions for improvement.

I would like to acknowledge financial grants/fellowships from the following agencies:

- UNIFOR, Oslo for support during the 12th International Conference on Bioinformatics & Computational Biology held in Las Vegas, USA, July 2012.
- Fellowship from COST (European Cooperation in Science and Technology), to present my paper in the Fourth StatSeq workshop "Statistical Challenges of the 1000€ genome sequences in plants", held in Italy, April 2012.
- Fellowship from SPPS - Scandinavian Plant Physiology Society to present my research as an oral talk in XXIV SPPS Congress, held in Stavanger, August 2011.
- Fellowship from Faculty of Natural Science and Technology, Norwegian University of Science and Technology to present my paper in 10th International Conference in Systems Biology held in Heidelberg, Germany, October 2011.
- Welcome Trust (UK), for a full fee waiver to attend 10 days long Welcome Trust Advance Course in "Functional Genomics to Systems Biology", held at Sanger Institute, Cambridge, UK, June 2010.

Coming to my family members, I thank my sister Pallabi Bora and brother-in-laws Dr. Prasanta Bora and Sumit Bhuyan for their continuous encouragements and love. Sincere thanks to my fiancé Kaveri Bhuyan, for her highest-level of tolerance, love and support towards me during the entire PhD period. You have been with me all the time to share my success, failure, frustrations and everything. You make me feel that there is life beyond PhD too. We have stayed apart from each other for long. Now it is time to get united.

My sincere thanks to parents-in-law Mr. Lohit Ch. Bhuyan and Mrs. Bima Bhuyan for their trust and confidence on me.

Finally, I would like to pay gratitude to my wonderful parents Mr. Siddheswar Bora and Mrs. Renu Phukan Bora, who initialized my way of logical thinking, scientific spirit and pursuit of knowledge since my childhood. I am blessed to have mom and dad like you. I am dedicating this thesis to both of you.

- Pankaj Barah

পংকজ বৰা

Manuscripts included in this thesis

Paper I. Konika Chawla*, **Pankaj Barah*,** Martin Kuiper, and Atle M. Bones**,** "Systems Biology: a promising tool to study abiotic stress responses" in "Omics and Plant Abiotic Stress Tolerance", Bentham Publishers, USA (ISBN: No.: 978-1-60805-058-1).

Paper II. Pankaj Barah, Per Winge, Anna Kusnierczyk, Diem Hong Tran and Atle M Bones, Molecular signatures of specific and common defense responses in *Arabidopsis thaliana* during an insect feeding and a bacterial infection **(**accepted in ***Plos One*****).**

Paper III. Simon Rasmussen, **Pankaj Barah,** Maria Cristina Suarez-Rodriguez, Simon Bressendorff, Pia Friis, Paolo Costantino, Atle M Bones, Henrik Bjorn Nielsen and John Mundy**,** Transcriptome response to combinations of stress in *Arabidopsis thaliana* (accepted in ***Plant Physiology***).

Paper IV. Pankaj Barah, Naresh Doni Jayavelu**,** Simon Rasmussen, Henrik Bjørn Nielsen, John Mundy, Atle M Bones**,** Natural variation of genome scale cold response regulatory networks in ten *Arabidopsis thaliana* ecotypes **(**manuscript submitted to ***New Phytologist*** **).**

Paper V. Pankaj Barah, Naresh Doni Jayavelu, Simon Rasmussen, Henrik Bjørn Nielsen, John Mundy and Atle M Bones, Transcriptional regulatory network in *Arabidopsis thaliana* during response to single and combined stresses (manuscript).

Paper VI. Jens Rohloff, **Pankaj Barah** and Atle M. Bones, Improving crop productivity and abiotic stress tolerance in cultivated Fragaria using omics and systems biology approach., in "Improving Crop Productivity in Sustainable Agriculture", Wiley-VCH Germany, ISBN: 978-3-527-33242-7, 2012

Paper VII. Ishita Ahuja, Per Winge, Marianne Trælnes, **Pankaj Barah**, Nicole M. van Dam and Atle Magnar Bones, Differential activation of plant defense in *Brassica napus L.* cv. Westar and transgenic MINELESS plants after attack by *Mamestra brassicae* (cabbage moth) (submitted).

My Contribution to the papers

Paper I:

-Equal contribution as joint first author.

-Contributed equally towards the planning of the manuscript.

-Analyzed the data together with other co-authors and the corresponding author.

-Wrote the paper together with the corresponding author.

Paper II:

-Contributed significantly towards the planning of the manuscript.

-Analyzed the data together with other co-authors and the corresponding author.

-Wrote the paper together with the corresponding author.

Paper III:

- Contributed to the data analysis.

- Contributed to the writing of the manuscript.

Paper IV:

- Developed the concept of the paper.

- Analyzed the data together with other co-authors and the corresponding author.

- Wrote the paper together with the co-authors.

Paper V:

- Developed the concept of the paper.

- Analyzed the data together with other co-authors and the corresponding author.

- Wrote the paper together with the co-authors.

Paper VI:

- First two authors contributed equally towards the planning of the manuscript.

- Analyzed the data together with other co-authors and the corresponding author.

- Wrote the paper together with the corresponding author.

Paper VII:

-Statistical and bioinformatics analysis of the data.

CHAPTER 1

(General introduction)

1.1 Plants

Plants (*Viridiplantae* in Latin) are living organisms of the kingdom *Plantae,* which includes multicellular groups like flowering plants, conifers, ferns and mosses, as well as the green algae. There are about 300,000 plant species on Earth (Mendelsohn, 1963). Among the life forms, plants, algae and cyanobacteria are the major groups that can produce their own food using energy from sunlight. An environment without plants is impossible to imagine. Even extreme environments like the hot and dry deserts or freezing polar regions have plants. These plants have adaptations that help them to survive the harsh conditions. Photosynthesis produce almost all of the oxygen in the air that human and other animals breathe. Plants are also an important source of food, building materials, and other resources that make life possible for the Earth's animals and humans.

1.2 Global climate change may have significant impacts on crop yields

World population is increasing exponentially and is expected to reach more than nine billion by the end of 2050. But, agriculture productivity is being seriously limited by adverse environmental factors and various biotic invasions. Most plants grow in suboptimal environments, which prevent them from attaining their full potential for growth and reproduction. This is reflected clearly in the difference between maximum crop yields statistics and the statistics of average yield for that crop (Boyer, 1982). Such difference in yields can mainly be explained by adverse environmental conditions, that potentially affect physiological processes within plants (Ahuja *et al.*, 2010). In a simple way, these adverse conditions are known as stresses on plants (Hirt, 2009). Environmental stress is a major cause of crop loss worldwide, resulting in average yield losses of more than 70% for major crops every year (Boyer, 1982), and plays a major role in determining the geographic distribution of crops (Trontin *et al.*, 2011).

Climate change may have significant impacts on society and ecosystems over the next decades (Brown & Funk, 2008). The global climatic pattern is becoming more unpredictable with increased occurrence of global warming, drought, cold, flood, chemical pollutions, high salinity, elevated CO_2. Lobell, Bruke et al. have used a mathematical model based on available agronomic and climatic data to calculate the trend in agricultural production up to the year 2030. The results show that climate change is likely to reduce agricultural production

yields of the major staple crops like corn, wheat, rice, maize in the near future, thus reducing food availability to an increasing world population (Lobell *et al.*, 2008; Lobell *et al.*, 2011). It has been estimated that abiotic stresses were the principal cause of decreasing the average yield of major crops by more than 50%, which caused losses worth hundreds of billions of US dollars each year (Ney *et al.*, 2000). Invasions by organisms such as bacteria, viruses, fungi, parasites, weeds on cultivated plants also cause vast economic losses (Peterson & Higley, 2000). The damage becomes more severe due to the co-occurrence of multiple abiotic stresses or the interaction of multiple abiotic and biotic stresses (Mittler, 2006; Atkinson & Urwin, 2012). Molecular effects resulting from combinations of stresses have not received much attention from plant molecular biologists.

1.3 What is stress in plants?

Stress is a frequently used term today, but has become increasingly difficult to define. In general, the term stress was associated with the mechanical concept of a force being applied to a body. In this context, stress is a measure of the internal forces acting within a deformable body. However, stress in biological systems is typically described as a negative event that can have an impact on normal physical stability of a living system. Robustness is a key property of any healthy living system. Most biochemical processes inside an organism try to maintain equilibrium, which is a steady state that exists more as an ideal and less as an achievable condition (Kitano, 2007). Such optimal condition in constant flux wavering point of physiological and biochemical processes in an organism is known as homeostasis (Cannon, 1929). Environmental factors, internal or external stimuli, continuously disrupt such homeostasis. Any such factors causing an organism's condition to deviate from homeostasis can simply be defined as stress (Nilsen & Orcutt, 1996).

The stress concept in plants is described according to physiological and ecological requirements of an organism throughout its life-cycle (Godbold, 1998). Grime et al. has defined stress as "Constraints which limit the utilization of resources, growth and reproduction" (J.P. Grime, 1991). The mentioned required resources in the above definition can be any environmental factor, and hence include chemical, physical and also biotic factors (**Figure1**). Such stress factors are defined as extreme environmental conditions that induce functional changes in plants to such an extent that stress on the organism develops, resulting in inhibited growth, reduced bioproduction, physiological acclimatization,

adaptation of species or some combination of these changes. Again, these factors can be divided into essential and non-essential factors. Essential factors can be nutrients, water, temperature and even other interacting organisms, for example mycorrhizae. Even for the essential factors most plants have a defined optimum above and below which growth is inhibited. Higley et al. (Higley *et al.*, 1993) proposed that plant stress be defined as "a departure from optimal physiological conditions" due to the adverse reaction generated from involvement of two types of environmental factors - mainly biotic (living organisms) and abiotic (climatic condition).

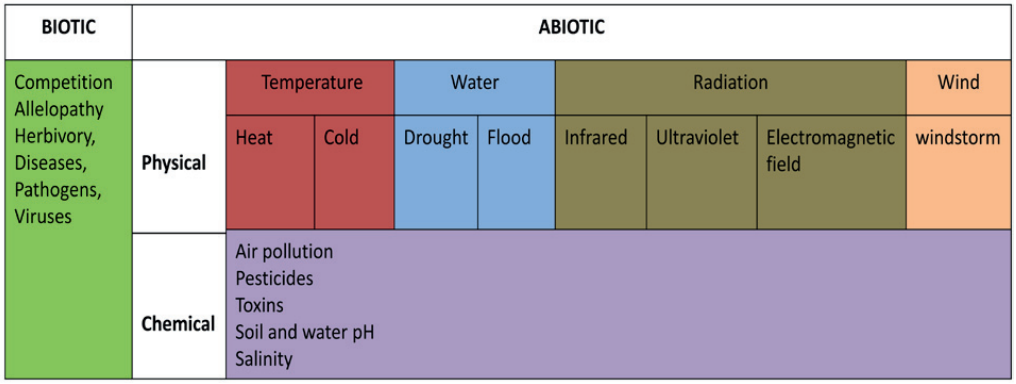


Figure1: Examples of different types of stress factors that can affect the plants homeostasis.

*Figure redrawn from (Nilsen & Orcutt, 1996).

1.4 Stress interaction

In a natural environment, plants are exposed to multiple stresses simultaneously, rather than a particular stress at a time. Interaction effects of multiple stresses are more severe to plants. Surprisingly, very few studies have been conducted till date to study the molecular responses of plants to a combination of different stresses, and these studies reported that the responses of plants to a combination of stresses were unique and could not be directly inferred from the response of plants to each individual stress (Rizhsky *et al.*, 2002). Exposure of plants to a combination of stress factors may trigger agonistic, antagonistic, or potentially unrelated responses. Such interaction between multiple biotic and/or abiotic stresses is coordinated by a complex signalling crosstalk of phyto-hormones (Mundy *et al.*, 2006). Phytohormones such

as salicylic acid (SA), jasmonic acid (JA), ethylene (ET), and abscisic acid (ABA) are major players that regulate the defense responses of plants against both biotic and abiotic stresses via synergistic and antagonistic actions, which are referred to as signalling crosstalk (Fujita *et al.*, 2006). **Figure 2** schematically explains the role of plant hormones in regulating the interaction between biotic and abiotic stress. Even though the complete molecular mechanism of stress signalling cross-talk is not fully understood yet, still it partially represents the summary of available knowledge about interactions taking place among hormones, transcription factors, and other regulatory components when biotic and abiotic stresses occur concurrently. More on phytohormone mediated stress signalling is explained under the section 1.8.

Some of the stress interactions are mutually positive to each other, but beyond a threshold they are antagonistic to one another. For example, temperature is known to influence disease resistance in plants against bacteria, fungi, virus, and insects (Zhu *et al.*, 2010). Different host-pathogen interactions respond differently to different temperature ranges. A high temperature very often inhibits disease resistance or plant immunity, although low temperature also leads to reduce plant defense in some cases. In most of the cases, long-term exposure of plants to abiotic stress conditions results in the weakening of plant's defense machinery and thus makes them more susceptible to pests or pathogen attacks (Wang *et al.*, 2009). Again, the molecular and metabolic response of plants to a combination of drought and heat cannot be directly predicted by combining the response of plants to each of these different stresses when applied individually (Rizhsky *et al.*, 2002).

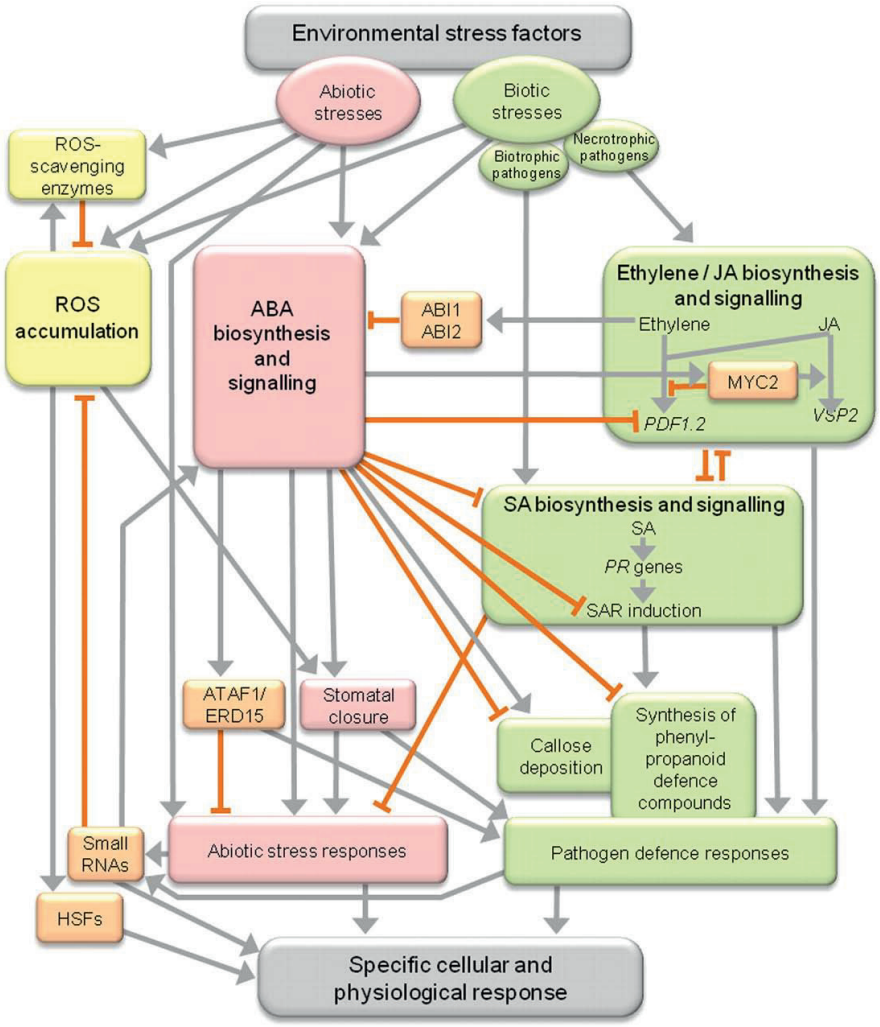


Figure 2: Phytohormones are major players that regulate defense responses of plants against both biotic and abiotic stresses via signalling crosstalk.

*Figure reproduced from (Atkinson & Urwin, 2012) , The interaction of plant biotic and abiotic stresses: from genes to the field, Journal of Experimental Botany, 2012, 63 (10): 3523-3543, by permission of Oxford University Press.

Mittler et al. reported that in the US, total agricultural losses during 1980 and 2004 due to drought stress were worth $20 billion, but loss due to combination of drought with a heat wave raised the figure up to $120 billion (Mittler, 2006). These figures suggest that the occurrence of a second stress factor can worsen the damage several folds. With the limited data available on combined stress treatment in plants, Mittler et al. (Mittler, 2006) have also presented a statistical model that they named as 'stress matrix', which summarized some of the stress combinations that could have a significant impact on agricultural production **(Figure 3).** This model reflects pair wise interactions between nine different combinations of biotic and abiotic stresses in the form of a matrix to show potential interactions that can have significant importance for agriculture. In some previous studies it has been reported that pre-exposure to a particular abiotic stress condition enhances the tolerance of plants to consecutive pathogen attacks (Bowler & Fluhr, 2000; Park *et al.*, 2001).

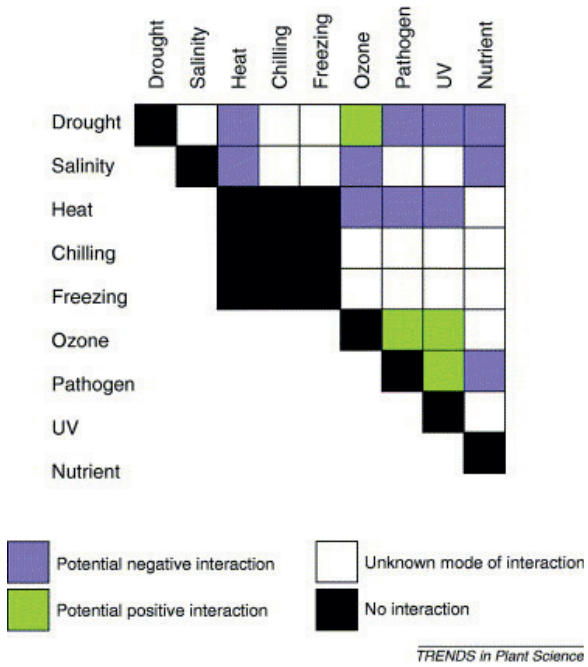


Figure 3: Stress interaction matrix summarizes potential positive, negative, unknown or neutral impacts of pair wise interactions among nine single stresses on plants. Different interactions are color coded to indicate potential negative (purple) or potential positive (green), effects of the stress combination on plant health.

*Figure reprinted from 'Abiotic stress, the field environment and stress combination'(Mittler, 2006), Trends in Plant Science, Volume 11, Issue 1, January 2006, Pages 15–19, with permission from Elsevier.

1.5 The complexity of climatic factors and plant environment interaction

Climate is a complex system constituted by many inter-related variables. A small perturbation in one variable may produce amplified changes in other components and hence in the whole system (Rind, 1999). With such tight inter-connections, the risks associated with climate change lie in the interaction of several systems with many variables that must be collectively considered (Ferro *et al.*, 2012). Agriculture (including crop cultivation, animal husbandry, forestry and fisheries) can be defined as one of the systems, and climate the other (Betts, 2005). The situation is more global. If the components of the system were treated independently, this would lead to an approach that is incomplete. It is now an established consensus that human activities affect climate (Vitousek *et al.*, 1997). Climate in turn affects agriculture production, the source of all food consumed by human beings and domestic animals (Falloon & Betts, 2010). It is now evident not only that climate changes but also that evolving patterns of human societies and agriculture practices develop trends and constraints of their own that might magnify the impacts of climate change (Gornall *et al.*, 2010). Hence, new scientific studies must be designed to take such issues into consideration.

As systems theory says, high levels of organization exhibit emergent properties (Bertalanffy, 1968). The relationship between complexity and physiological stability has been observed among different kinds of biological systems. Interactions among modular components of a complex biological network can facilitate predictions of behavior under environmental perturbations. Like other biological systems, plants are highly complex systems, composed of highly interconnected elements, arranged in a hierarchical manner from molecular to the whole plant and ecosystem level. As expected, they show some properties that may not be understood by studying the isolated elements (Spiertz, 2007). Considering plant as a complex system, temporal dynamics of parameters related to processes such as photosynthesis, enzymatic reactions and a broad class of fluxes could be associated with a greater capacity of system homeostasis and successive adaptation (Hesse & Hoefgen, 2006). Utilizing enormous amounts of high throughput omics data (for example genomic, proteomic, transcriptomic, metabolomic, phenomic, interactomic, ionomic) along with robust bioinformatics and data mining tools, scientists can now explore relevant correlations and construct mathematical or statistical models describing different physiological states. Models of the various cellular processes such as enzyme activities signaling cascades, gene expression, metabolite pools or pathway flux modes can help us dealing with the complexity of the plant system. Particularly integration of multidimensional heterogeneous data from

transcriptomics, metabolomics and proteomics experiments into consistent models will be tremendously helpful in describing and predicting the behavior of biological systems.

1.6 Plant Systems Biology

Plants are sessile systems unable to escape biotic and abiotic stresses. As a result, they have developed intricate mechanisms to perceive external signals, allowing optimal response to stress conditions. Understanding the systems level responses of whole plants to environmental conditions is essential if we are to use genetic and molecular approaches to develop crops that grow well in harsh environments. Some responses of the plants to different stress conditions are very general and provide protection from a variety of stress conditions, whereas others are more specific against a particular stress type. The multidimensional level of a network's crosstalk makes it challenging to recognize which of the observed responses are general and which are more stress-specific. Understanding the mechanisms of how plants respond to various single and combinations of stresses is therefore essential and needed to develop broad-spectrum stress-tolerant crops (Mittler & Blumwald, 2010). To meet the food requirement for the human population, major crops must be improved through selective breeding or genetic modifications to ensure productivity in rapidly changing field environments (Zhang *et al.*, 2000; Takeda & Matsuoka, 2008; Iglesias *et al.*, 2011).The molecular components of cellular life forms such as proteins, nucleic acids and metabolites, have largely been studied in isolation or as parts of individual pathways. In reality, they are tied together to form a large, interlinked, complex system in the cell, very much like a densely connected network (Yuan *et al.*, 2008). Systems biology is based on the idea that properties of a complex biological system cannot be understood by focusing on any one aspect of their highly interacting components (Kitano, 2002). Being a biology-based inter-disciplinary field, systems biology focuses on complex interactions among different components in the biological systems **(Figure 4).** It uses a new perspective 'holism instead of reductionism' (Raikhel & Coruzzi, 2003; Hammer *et al.*, 2004).

The integrative systems approach has been getting large attention of plant biologists in the last few years, concomitant with the increase in large amount of molecular data. But integration and interpretation of these huge amounts of omics data to create a holistic view of a biological process has been limited to date. To meet the challenges involved in integrating the omics information, communication between plant biologists and computational scientists

is necessary. In recent years, integrated approaches like systems biology have been evolving as promising tools to study plant stress responses and adaptation (Fukushima *et al.*, 2009; Kliebenstein, 2010; Mochida & Shinozaki, 2010; Cramer *et al.*, 2011; Mochida & Shinozaki, 2011). Crop scientists have been using systems approaches to investigate whole-crop physiology, crop ecology and morphology (Trewavas, 2006). The term *Plant Systems Biology* was first defined at the 22nd Symposium on Plant Biology, using computational modelling approaches to predict a plant cell (ome) from underlying genomic understanding (Minorsky, 2003; Raikhel & Coruzzi, 2003). So, *plant systems biology* is not to be considered an entirely new field (Gutierrez *et al.*, 2005). A new term *crop systems biology* has been proposed recently, which aims at modelling complex crop-level traits relevant to global food production and energy supply, by integrating omics-level information, underlying biochemical understanding, and physiological component processes (Yin & Struik, 2007).

Chapter 2 in this thesis elaborates in more detail, why and how systems biology is useful to study plant stress biology reviewing various approaches and tools available to plant biologists (Chawla *et al.*, 2011). In consecutive chapters of this thesis we illustrate case-based-examples of various such integrated approaches to understand the diverse range of plant stress response mechanisms.

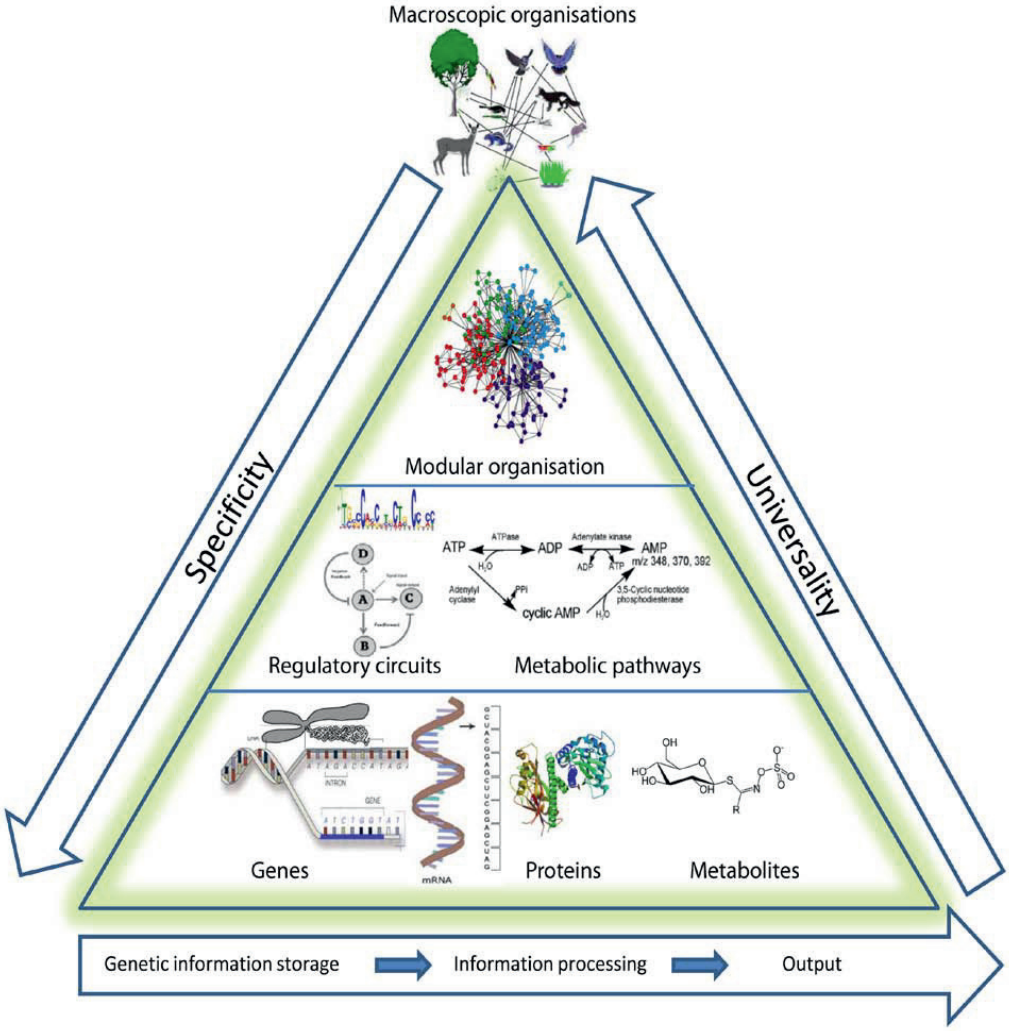


Figure 4: Life's complexity pyramid: It shows the complexity of hierarchical modular and interconnected organisation of living systems from macroscopic level to molecular level.

* Figure modified from (Oltvai & Barabasi, 2002).

1.7 Molecular picture behind plant stress response

Different plant species may react differently to the same extreme condition. Moreover, the response and sensitivity towards stressors depend on the age and development stage of the plant in question. For example, young tree seedlings are sensitive towards the water content in the upper layers of the soil, and they may die as a result of flood or drought, while adults of the same species need not have any sensitivity towards the water content of the upper soil layers (Kozlowski, 1991; Nilsen & Orcutt, 1996). In addition, the intensity of responses may vary in time and space, being different in the organs of the same plant.

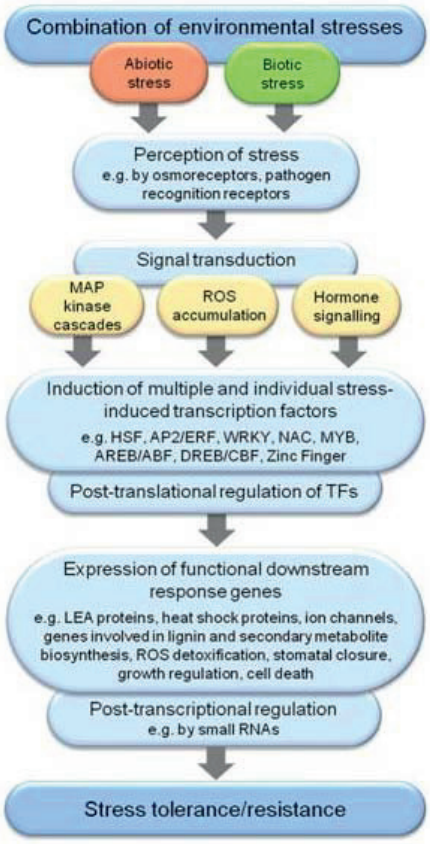


Figure 5: Key events in the signal transduction pathway activated in response to combined biotic and abiotic stresses.

* Figure reproduced from (Atkinson & Urwin, 2012) , The interaction of plant biotic and abiotic stresses: from genes to the field, Journal of Experimental Botany, 2012, 63 (10): 3523-3543, by permission of Oxford University Press.

During their long course of evolution and artificial domestication, plants have evolved a series of fine mechanisms for responding to different types of stresses. Such mechanisms include many aspects of anatomy, physiology, biochemistry, genetics, development, evolution and molecular biology. The molecular mechanisms developed in higher plants as response to stress conditions starts with environmental signal recognition (input), signal transduction (several biochemical cascades and crosstalks are involved in this process), signal output, signal responses and phenotype realization, which is a multi-dimensional network system and contains many levels of gene expression and regulation **(Figure 5)**. From several microarray experiments conducted to uncover transcriptional response pattern(s) to different abiotic and biotic stresses, it is now understood that the transcriptional response initially is composed of a core set of genes responsive to multiple stresses, but becomes gradually more stress specific as time progresses (De Vos *et al.*, 2005; Eulgem, 2005; Bohnert *et al.*, 2006; Kilian *et al.*, 2007). Such general stress response has also been referred to as the cellular stress response by Kultz et al. (Kultz, 2005) or core stress response by Lopez-Maury et al. (Lopez-Maury *et al.*, 2008). Interestingly, key molecular components of this general stress response have shown to be evolutionarily conserved in all organisms (Singh *et al.*, 2008).

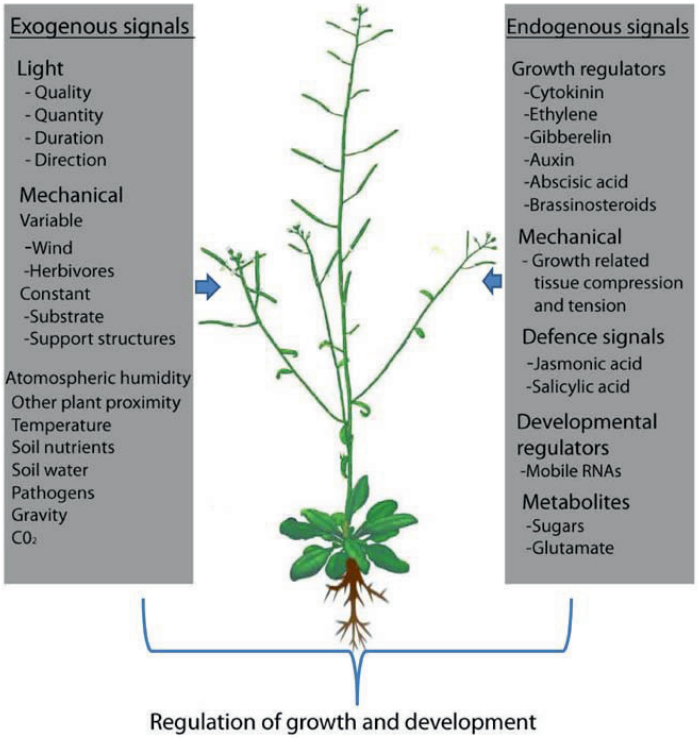
In **Chapter 3,** we have explored the common and attacker-specific defense responses in *Arabidopsis thaliana* while they are attacked by an insect or by a pathogenic bacterium.

1.8 Stress signal perception and hormone mediated signalling in plants

Plants might perceive the stresses in different ways, such as by plasma membrane located receptors, intracellular or cytoskeleton-associated proteins. Perceived stress signals are transmitted by signalling cascades, which lead to changes in gene expression patterns, ultimately resulting in metabolic re-programming and altered physiological responses. There are multiple stress perception and signalling pathways, some of which are specific, but others may cross-talk at various steps (Tuteja & Sopory, 2008a). The general stress response works in a rapid and transient manner in response to a range of stresses and responds to strains imposed by environmental forces on macromolecules such as membrane lipids, proteins and DNA.

All plants are able to detect the intensity in signals (such as light) and nutrient resources (such as nitrate and water) (Gilroy & Trewavas, 2001). Compared to animals, tissue and cell level functional specialization is minimized in plants. Most plant cells can sense

nearly all the signals to which the individual plant responds **(Figure 6).** Specific membranes of higher plants are equipped with receptors, channels, G-protein coupled receptors (GPCRs), and receptor-like kinases (RLKs). Some signalling protein complexes are permanent, such as the comparatively stable COP9 signalosome (Wei & Deng, 2003; Tuteja & Sopory, 2008b). Other signalling protein complexes are likely to be transient and formed immediately as a result of signaling. There are more than 600 receptor-like kinases in the most popular model plant *Arabidopsis thaliana*, and most of them are membrane bound (Shiu & Bleecker, 2001). After ligand binding and autophosphorylation, such kinases may act as nucleation sites for the construction of temporary signalling complexes that contain many proteins.



F**igure 6:** Plants can sense a wide range of different external and internal signals that are used to control appropriate growth and developmental responses. The molecular components of the plant sensory machinery and signal-transduction systems can incorporate these signals and make a stable decision as to how to grow and develop by successful utilization of available resources and constrains.

*Figure modified from(Gilroy & Trewavas, 2001) by permission from Macmillan Publishers Ltd: [Nature Reviews Molecular Cell Biology].

Signal transduction during multiple biotic and abiotic stress results from a complex array of interacting components (Fujita *et al.*, 2006). A common signal perception and signal transduction model for stress transduction pathways always exists in higher plant, with few exceptions. A simplified view of this model is shown in **Figure 5**. It begins with the perception of stress signals from environments, followed by the generation of secondary messengers (such as inositol phosphates and reactive oxygen species). Secondary messengers can modulate intracellular Ca^{2+} levels, often initiating a protein phosphorylation cascade that finally targets proteins directly involved in cellular protection or transcription factors controlling specific sets of stress-regulated genes. The products of these genes may participate in the production of regulatory molecules like the plant hormones abscisic acid (ABA), ethylene, and salicylic acid (SA). Previously, responses to abiotic stress were known to be mainly controlled by the hormone ABA, while stress responses against different biotic attackers was known to be regulated by an antagonistic interaction between the salicylic acid (SA) and jasmonic acid (JA)/ethylene signalling pathways (Fujita *et al.*, 2006). But recent findings suggest that ABA acts both synergistically and antagonistically with biotic stress signalling, through a complex network of interacting pathways which cross-talk at different levels of signal transduction (Yasuda *et al.*, 2008). Now ABA is regarded as a global stress regulator, that can dominantly switch the priority between the response to biotic or abiotic stress and allowing plants to respond to the most severe threat (Asselbergh *et al.*, 2008).

In **Chapter 4,** we have explored the transcriptome level difference in stress response pattern when plants were exposed to 5 single stresses and 6 combinations of stresses.

1.9 *Arabidopsis* as a model plant

In biological science, models are those organisms with a huge amount of biological existing information that make them attractive to study as examples for other species and/or natural phenomena that are more difficult to study directly. Such models are widely used in genetic studies because they possess characteristics, such as short generation time and large numbers of progeny that make it well suited to genetic analysis. Such models can be used to study different levels of biological systems; from ecology, behavior, and biomechanics, down to the tiny functional scale of individual tissues, organelles, and proteins (Fields & Johnston, 2005). Advancement of modern omics techniques has enhanced the generation of huge amounts of heterogeneous data, and consequently the possibility of exploring any biological systems in a

more holistic way (Joyce & Palsson, 2006). There are several model systems available in plants. Among all of them, *Arabidopsis thaliana* has been the most widely studied 'reference system', for nearly all biological processes by the plant science community (Van Norman & Benfey, 2009). *Arabidopsis* is now a well-established model system in plant biology or 'a fortunate choice' to study fundamental mechanism of stress responses and to translate such knowledge to other cultivating crops (Somerville & Koornneef, 2002). As of September 2012, 23913 people and 9968 unique institutes/groups are registered as *Arabidopsis* researchers in The Arabidopsis Information Resource (TAIR) (Lamesch *et al.*, 2012). Apart from that, *Arabidopsis* has crossed the border of plant sciences and evolved into a beneficial model system even for molecular mechanism related to human health and diseases (Martin *et al.*, 2011).

Arabidopsis is a member of the mustard (*Brassicaceae*) family, which includes cultivated species such as oilseed rape, cabbage and radish. The first genome sequence of the model plant *Arabidopsis thaliana* was published in 2000 by the *Arabidopsis* Genome Initiatives (AGI), and it has approximately 115 Mb of the 125 Mb genome (Alonso-Blanco & Koornneef, 2000). Comprehensive genetic and physical maps of all 5 chromosomes in *Arabidopsis* are now available. If we compare the situation a decade later, there are 503 instances of genome projects (as of October 2012) in species of 'Viridiplantae' that includes both green algae and land plants, including many agronomically important crops (http://www.ncbi.nlm.nih.gov/genome?term=Viridiplantae).

A physician Johannes Thal first discovered the *Arabidopsis* species in 1577 in the Harz Mountains. He called it *Pilosella siliquosa*. In 1841, the plant was renamed *Arabidopsis thaliana* by German botanist Gustav Heynhold in honor of Thal. The genus name, *Arabidopsis* comes from Greek, meaning "resembling Arabis". As reported in TAIR (http://www.*Arabidopsis*.org), systematic collection of the plant began in 1900 and genetic experiments were performed in 1907 by Friedrich Laibach in the University of Bonn, Germany, for his Ph.D. thesis (Meyerowitz, 2001). In 1943, Laibach as a Professor in Botany at the University of Frankfurt proposed the suitability of *Arabidopsis* as a model for genetic and developmental biology research. Since then, hundreds of research groups across the globe have adapted this weed as a model system for studying different mechanisms in plant sciences. During the XIth Genetics Congress held in The Hague in 1963, many scientists working with *Arabidopsis* agreed upon establishing an *Arabidopsis* Information Service (AIS) to exchange information on *Arabidopsis*. The first international *Arabidopsis* Symposium was

organized in Göttingen, Germany, in 1965. In the year 1986, the first *Arabidopsis* gene transformation was performed by Lloyd et al. (Lloyd *et al.*, 1986). The first ever quantitative monitoring of gene expression patterns with a complementary DNA microarray was published on 45 *Arabidopsis* genes in 1995 (Schena *et al.*, 1995). The near completion 1001 Genomes project was launched at the beginning of 2008, with a goal to discover the whole-genome sequence variation in 1001 strains (accessions) of the reference plant *Arabidopsis thaliana* (Cao *et al.*, 2011). Recently, a large scale experiemnts was conducted taking *Arabidopsis thaliana* as a model system to understand multiple stress responses and adaptation mechanisms in plants. The general objective of this project was to analyze responses and adaptations of plants to multiple stresses and to identify the level and functions of stress regulatory networks and crosstalk (Rasmussen *et al.*, in press). Both biotic and abiotic (unfavourable environmental conditions) stresses were taken into account. The generated unique dataset has been extensively used in 3 manuscripts **(Paper III, IV and V)** presented in this thesis.

(http://www.eracaps.org/joint-calls/era-pg-funded-projects/2006-sub-call/multiple-stress-responses-and-adaptations).

1.10 Natural variation as a key tool in plant stress biology

Environmental stress is a key factor to determine the genome regulation, evolutionary history and geographical distribution of any living organisms including plants (Alonso-Blanco *et al.*, 2009; Becker & Weigel, 2012). Intraspecific natural variation or within-species phenotypic variation caused by spontaneously arising favorable mutations that have been maintained in nature to facilitate evolutionary process, contribute towards the local adaptation of the plant for survival. Such natural variation contributes to plant development and physiology, germination and flowering, plant growth and morphology, defense responses to stress, primary metabolism and mineral accumulation (Mitchell-Olds & Schmitt, 2006). The vast diversity within wild plant species as well as most of the genetic found in domesticated cultivated plants are mainly due the combined effect of natural variation and evolutionary processes. Such natural variation present in crop plants has been utilized by human society for the last thousands of years for genetic selection of developmental traits and physiological features beneficial for agriculture (Doebley *et al.*, 2006). Additionally, studying natural variation in wild species can tell us about the molecular basis of phenotypic differences

related to plant's adaptation to diverse natural environments (Borevitz & Nordborg, 2003; Weigel & Nordborg, 2005). It can explain the underlying molecular mechanism that determines the ecological and evolutionary plasticity of a species through such variations. The hidden potential of studying genetic variation to different areas of plant biology or crop sciences was strongly highlighted by Maarten Koornneef and his co-workers (Alonso-Blanco & Koornneef, 2000). Natural systems are highly optimized and robustly engineered to face any adverse situation through the strict selection processes of evolution. The existing natural diversities among plants in nature can easily be harnessed for developing better traits in cultivated crops.

The most straightforward approach to identify the causal genes underneath such natural variation is genetic mapping. It uses statistical methods to find regions of the genome associated with the trait of interest, an approach that is known as quantitative trait locus (QTL) mapping (Borevitz & Chory, 2004). New high throughput technologies such as genome sequencing, microarray, RNA sequencing, SNP arrays and metabolomics have been proven as a great aid towards studying natural variation at a systems level (Atwell *et al.*, 2010; Chan *et al.*, 2011; Filiault & Maloof, 2012; Horton *et al.*, 2012; Weigel, 2012). Genome-wide association (GWA) studies have evidently become a powerful approach for studying the genetics of natural variation and traits of agricultural importance (Atwell *et al.*, 2010). The main advantage of studying natural variations at the systems level is, apart from identifying a single gene or protein, that we now can globally look at the variation in pathways or processes (Chawla *et al.*, 2011).

In **Chapters 5** and **6**, natural variation in stress response patterns among *Arabidopsis* ecotypes will be presented.

1.11 Transcription factors (TFs) and regulation of stress gene expression

As the central dogma of molecular biology says, transcription of mRNA from DNA and subsequent translation of mRNA into protein transform genetic blueprints into cellular functions (Crick, 1970). Regulation of gene expression is a key component in development and evolution of living beings along with genome composition and structure (Chen *et al.*, 2005; Salse, 2012). Being highly dynamic in nature, any biological system continuously changes responding to environmental and genetic perturbations. A single transcription factor (TF) can control the expression of many target genes through specific binding of the TF to

cis-acting elements in the promoter of respective target genes. Often, genes that respond to specific stresses can be activated or repressed by several closely related transcription factors. Transcriptional re-programming is a key step of plant response to various stresses (Singh *et al.*, 2002). The *Arabidopsis* genome encodes more than 1900 transcription factors and they generally belong to large gene families, which in some cases are unique to plants (Guo *et al.*, 2005). The Database of *Arabidopsis* Transcription Factors (DATF) collects all *Arabidopsis* transcription factors (in total 1922 loci; 2290 gene models) and classifies them into 64 families (Guo *et al.*, 2005). Significant progress has been achieved in the past few years in characterizing stress inducible transcription factors in plants (Shameer *et al.*, 2009). But experimental validation and evidence (ChIP-chip, flow cytometry) about how many of putative TF binding sites actually bind a TF to result in regulation of their downstream gene in vivo is still lacking (Yilmaz *et al.*, 2011).

TFs are of key importance in generating specificity in plant stress responses (Chen *et al.*, 2002). Information regarding activity dynamics of TFs and their dynamic regulatory relationships with target genes are presently not yet available for *Arabidopsis* at the genome scale. The AGRIS database has collected regulatory relations for near about 100 TFs only (~5%) (Davuluri *et al.*, 2003). Shameer et al. integrated 2,269 genes upregulated in different stress related microarray experiments and surveyed their 1,000 bp and 100 bp upstream regions and 5′UTR regions using the 'STIF' algorithm and identified putative abiotic stress responsive transcription factor binding sites, which are now compiled in the STIFDB database (Arabidopsis Stress Responsive Transcription Factor DataBase) (Shameer *et al.*, 2009). Ahuja et al. (2010) have listed a compendium of plant TFs responsive during single stresses and multiple stress treatments **(Figure 7)**. TFs responsive during multiple stress treatments will be of tremendous importance in engineering multiple stress resistant crop varieties to face the abruptly changing global climate (Ahuja *et al.*, 2010). Microarray analyses combined with genetic and biochemical approaches are now enabling us to study basic principles and details of regulatory mechanisms controlling the defense transcriptome in *Arabidopsis* (Eulgem, 2005).

Different types of physical and genetic interaction networks generated from 'omic' data provide key insights into complex biological systems, from how different processes interact to the function of individual residues on a single protein (Barabasi & Oltvai, 2004). Differential dynamic network mapping of such processes facilitates the exploration of previously unknown interactions (Ideker & Krogan, 2012). To compensate the lack of

experimental data for transcription factor activity on genome-scale, several computational algorithms have been developed to identify regulatory modules and their condition-specific regulators from gene expression data (Segal *et al.*, 2003; Herrgard *et al.*, 2004; Kao *et al.*, 2004; Tirosh & Barkai, 2011).

Chapters 5 and **6** will describe use of a computational algorithm to re-construct transcriptional regulatory network model in responses to single and combined stress conditions in *Arabidopsis thaliana*.

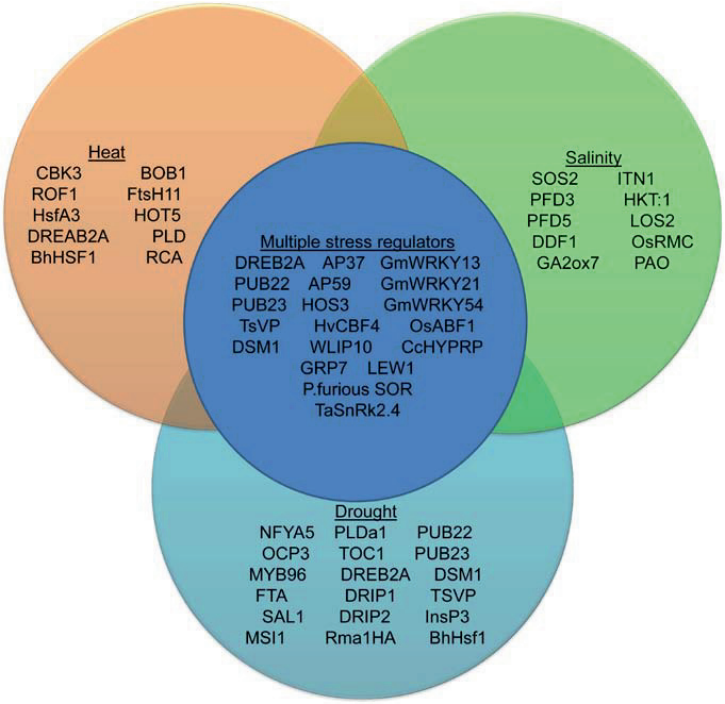


Figure 7: Transcriptions factors **(TFs)** play a crucial role during stress responses in plants. Some TFs are stress-specific and some are differentially regulated during multiple stresses. In this Venn-diagram, some of the stress specific TFs in plants during heat, salt and drought stresses are listed inside the brown, green and cyan circle respectively. The TFs in the blue circle are regulated in multiple stress conditions.

*Redrawn from (Ahuja *et al.*, 2010), Plant molecular stress responses face climate change, Trends in Plant Science, Volume 15, Issue 12, 664-674, with permission from Elsevier.

1.12 From laboratory to agriculture field and medicine: Translational perspectives of plant research

The term 'translational research', was not so popular a decade ago, but is now becoming very popular because it is seen as the solution to different problems faced by human society. It is highly expected that the findings of modern cutting edge science to be 'translated' into benefits of the everyday world (Anonymous, 2008). The utmost necessity of the present time is to convert knowledge gathered from basic sciences to practical applications that can enhance human society. In plant sciences, the main challenge is how to exploit enormous information gained from model systems like *Arabidopsis* to produce new crop plant varieties (Zhang *et al.*, 2004; Chen *et al.*, 2012). The number of genes in diploid flowering plants is surprisingly similar (Hall *et al.*, 2002). This commonality gives the opportunity of employing the strength of comparative genomics to transfer knowledge from model plants like *Arabidopsis* to crops (Caicedo & Purugganan, 2005). Using powerful techniques available in systems biology, knowledge and hypothesis can also be generated for plant models by integrating information and resources available from other species or among different plant species (Lee, I *et al.*, 2010; Ficklin & Feltus, 2011). *Arabidopsis* research has greatly contributed in unraveling abiotic stress responsive processes (Bartel & Last, 2004). Zhang et al. listed a couple of such success stories that had been implemented on other plant species based on findings in *Arabidopsis* research (**Table1**) (Zhang *et al.*, 2004).

The medical science communities now acknowledge that research on a plant model like *Arabidopsis* can significantly contribute to human health and medicine (Jones *et al.*, 2008; Eckardt, 2011; Martin *et al.*, 2011). For example, researchers from the Flanders Interuniversity Institute for Biotechnology (VIB) connected with Ghent University, studying cell division in plants have revealed the importance of the DEL1 protein, which controls cell division in *Arabidopsis*. They have shown that the human variant of this protein, E2F7, performs the same essential function in human cells (Vlieghe *et al.*, 2005). Through BLAST search, it has been found that, among cancer genes, 70% (E-value cutoffs of less than E^10) of genes implicated in cancer have *Arabidopsis* orthologs (Jones *et al.*, 2008). Such finding will help cancer research in order to better understand the factors that control the cell division during cancer. Ting et al. have reported the presence of equivalent human orthologous and paralogous genes of the *Arabidopsis* NB-LRR genes (~150 genes) in the animal innate immune system called NOD/CARD/CATERPILLER (Ting *et al.*, 2006). Research on plant stem cells significantly aids medical science, as there are intriguing similarities in the way

stem cells function in both plants and animals to sustain growth and replace tissues (Sablowski, 2004; Lee, EK *et al.*, 2010). Studying light signaling complexes like COP9 signalosome and COP1 in plants, contributed towards understanding mammalian tumorigenesis, DNA damage, and lipid metabolism (Dornan *et al.*, 2004; Qi *et al.*, 2006). It was shown in *Arabidopsis* that auxin regulates gene expression by promoting the ubiquitin-dependent degradation of transcriptional repressors called Aux/IAA proteins which is similar to the regulation of animals cells (e.g. NFkB)(Parry & Estelle, 2006).

In **Chapter 7** of this thesis, two application case studies will be presented that demonstrate the power of integrated systems approaches towards translating knowledge from basic plant research to application oriented translational research.

Gene Name	Gene Source	Transgenic Species	Intervention Method	Traits
		Ion Transporters		
AtNHX1 (vacuolar Na^+/H^+ antiporter)	Arabidopsis	Arabidopsis	Overexpression	Salt tolerance
AtNHX1 (vacuolar Na^+/H^+ antiporter)	Arabidopsis	*B. napus*	Overexpression	Salt tolerance
AtNHX1 (vacuolar Na^+/H^+ antiporter)	Arabidopsis	Tomato	Overexpression	Salt tolerance
AtNHX1 (vacuolar Na^+/H^+ antiporter)	*A. gmelini*	Rice	Overexpression	Salt tolerance
SOS1 (plasma membrane Na^+/H^+ antiporter)	Arabidopsis	Arabidopsis	Overexpression	Salt tolerance
AVP1 (vacuolar H^+-ATPase)	Arabidopsis	Arabidopsis	Overexpression	Salt tolerance
HKT1 (high affinity K^+ transporter)	Wheat	Wheat	Antisense	Salt tolerance
		Transcription Factors		
CBF1, DREB1a (CBF3)	Arabidopsis	Arabidopsis	Overexpression	Freezing, salt, and drought tolerance
CBF1, CBF2, CBF3	Arabidopsis	*B. napus*	Overexpression	Freezing and drought tolerance
CBF1	Arabidopsis	Tomato	Overexpression	Drought, chilling, and oxidative stress tolerance
CBF1	Arabidopsis	Strawberry	Overexpression	Freezing tolerance
ZmCBF	Maize	Maize	Overexpression	Cold tolerance
DREB1a (CBF3)	Arabidopsis	Wheat	Overexpression	Drought tolerance
OsDREB1A	Rice	Arabidopsis	Overexpression	Drought, salt, and freezing tolerance
CBF3	Arabidopsis	Rice	Overexpression	Stress tolerance
ICE1	Arabidopsis	Arabidopsis	Overexpression	Freezing tolerance
ABF3/4	Arabidopsis	Arabidopsis	Overexpression	Drought tolerance
Tsi1	Tobacco	Tobacco	Overexpression	Salt tolerance
SCOF-1	Soybean	Arabidopsis, Tobacco	Overexpression	Low temperature stress tolerance
AtMYC2/AtMYB2	Arabidopsis	Arabidopsis	Overexpression	Drought tolerance

Table1: Few examples of translational research success stories in *Arabidopsis*.

*Table reprinted from (Zhang *et al.*, 2004), Plant Physiology 135(2): 615-621, with Copyright permission from American Society of Plant Biologists.

Aims of the study

Paper I: The main aim of this review paper (book chapter) is to conduct an extensive review of existing computational tools (software and databases) and methods suitable for studying abiotic stress responses in plants.

Paper II: The main aim of this paper is to explore common and stress specific response signatures in the host plant *Arabidopsis thaliana* during attack by insect *Brevicoryne brassicae* and infection by bacteria *Pseudomonas syringae pv. tomato strain DC3000* .

Paper III: The aim of the study is to analyze transcriptome level changes in ten *Arabidopsis thaliana* ecotypes during 5 single and 6 combinations of stresses. This paper also aims at using network based approaches to identify hub genes that might be crucial during responses to combined stresses in plant.

Paper IV: The main aim of this study is to analyze intraspecific variation in transcriptomic response signatures during cold stress treatment among 10 *Arabidopsis thaliana* ecotypes originated from different geographic locations. Another aim is to re-construct an in silico transcriptional regulatory network model during cellular responses to cold stress in *Arabidopsis thaliana* and to explore the probable effect of sequence polymorphism on gene-expression pattern in the core cold stress regulon genes.

Paper V: The aim is to re-construct an *in silico* transcriptional regulatory network model in responses to 11 stress (5 single and 6 combined) conditions in *Arabidopsis thaliana* from microarray data. Another aim is to identify common and stress specific transcription factors and their differential regulatory activities while responding to 11 stress conditions.

Paper VI and VII: The common aim for both of these studies is to show that different systems biology approaches primarily used and developed for model plant *Arabidopsis thaliana* can also be successfully utilised for studying stress response mechanism in cultivating crops like *Fragaria* and *Brassica napus.*

REFERENCES

Ahuja I, de Vos RC, Bones AM, Hall RD. 2010. Plant molecular stress responses face climate change. *Trends Plant Sci* **15**(12): 664-674.

Alonso-Blanco C, Aarts MGM, Bentsink L, Keurentjes JJB, Reymond M, Vreugdenhil D, Koornneef M. 2009. What Has Natural Variation Taught Us about Plant Development, Physiology, and Adaptation? *Plant Cell* **21**(7): 1877-1896.

Alonso-Blanco C, Koornneef M. 2000. Naturally occurring variation in Arabidopsis: an underexploited resource for plant genetics. *Trends Plant Sci* **5**(1): 22-29.

Anonymous. 2008. Translational research: Getting the message across. *Nature* **453**(7197): 839-839.

Asselbergh B, De Vleesschauwer D, Hofte M. 2008. Global switches and fine-tuning-ABA modulates plant pathogen defense. *Mol Plant Microbe Interact* **21**(6): 709-719.

Atkinson NJ, Urwin PE. 2012. The interaction of plant biotic and abiotic stresses: from genes to the field. *J Exp Bot*.

Atwell S, Huang YS, Vilhjalmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT, Jiang R, Muliyati NW, Zhang X, Amer MA, Baxter I, Brachi B, Chory J, Dean C, Debieu M, de Meaux J, Ecker JR, Faure N, Kniskern JM, Jones JD, Michael T, Nemri A, Roux F, Salt DE, Tang C, Todesco M, Traw MB, Weigel D, Marjoram P, Borevitz JO, Bergelson J, Nordborg M. 2010. Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. *Nature* **465**(7298): 627-631.

Barabasi AL, Oltvai ZN. 2004. Network biology: understanding the cell's functional organization. *Nat Rev Genet* **5**(2): 101-113.

Bartel B, Last RL. 2004. Weed power, translating Arabidopsis. *Plant Physiology* **135**(2): 601-601.

Becker C, Weigel D. 2012. Epigenetic variation: origin and transgenerational inheritance. *Current Opinion in Plant Biology*.

Bertalanffy L. 1968. *General System Theory, Foundations, Development, Applications* New York: George Braziller, Inc.

Betts RA. 2005. Integrated approaches to climate-crop modelling: needs and challenges. *Philos Trans R Soc Lond B Biol Sci* **360**(1463): 2049-2065.

Bohnert HJ, Gong Q, Li P, Ma S. 2006. Unraveling abiotic stress tolerance mechanisms--getting genomics going. *Current Opinion in Plant Biology* **9**(2): 180-188.

Borevitz JO, Chory J. 2004. Genomics tools for QTL analysis and gene discovery. *Current Opinion in Plant Biology* **7**(2): 132-136.

Borevitz JO, Nordborg M. 2003. The impact of genomics on the study of natural variation in Arabidopsis. *Plant Physiology* **132**(2): 718-725.

Bowler C, Fluhr R. 2000. The role of calcium and activated oxygens as signals for controlling cross-tolerance. *Trends Plant Sci* **5**(6): 241-246.

Boyer JS. 1982. Plant Productivity and Environment. *Science* **218**(4571): 443-448.

Brown ME, Funk CC. 2008. Climate. Food security under climate change. *Science* **319**(5863): 580-581.

Caicedo AL, Purugganan MD. 2005. Comparative plant genomics. Frontiers and prospects. *Plant Physiology* **138**(2): 545-547.

Cannon WB. 1929. Organization for physiological homeostasis *Physiol Rev* **9:(3)**: 399-431.

Cao J, Schneeberger K, Ossowski S, Gunther T, Bender S, Fitz J, Koenig D, Lanz C, Stegle O, Lippert C, Wang X, Ott F, Muller J, Alonso-Blanco C, Borgwardt K, Schmid KJ, Weigel D. 2011. Whole-genome sequencing of multiple Arabidopsis thaliana populations. *Nat Genet* **43**(10): 956-963.

Chan EK, Rowe HC, Corwin JA, Joseph B, Kliebenstein DJ. 2011. Combining genome-wide association mapping and transcriptional networks to identify novel genes controlling glucosinolates in Arabidopsis thaliana. *PLoS Biol* **9**(8): e1001125.

Chawla K, Barah P, Kuiper M, Bones AM 2011. Systems Biology: a promising tool to study abiotic stress responses.In Tuteja N GSTR. *Omics and Plant Abiotic Stress Tolerance*. USA Bentham Publishers. 163-172.

Chen WJ, Chang SH, Hudson ME, Kwan WK, Li J, Estes B, Knoll D, Shi L, Zhu T. 2005. Contribution of transcriptional regulation to natural variations in Arabidopsis. *Genome Biol* **6**(4): R32.

Chen WQ, Provart NJ, Glazebrook J, Katagiri F, Chang HS, Eulgem T, Mauch F, Luan S, Zou GZ, Whitham SA, Budworth PR, Tao Y, Xie ZY, Chen X, Lam S, Kreps JA, Harper JF, Si-Ammour A, Mauch-Mani B, Heinlein M, Kobayashi K, Hohn T, Dangl JL, Wang X, Zhu T. 2002. Expression profile matrix of Arabidopsis transcription factor genes suggests their putative functions in response to environmental stresses. *Plant Cell* **14**(3): 559-574.

Chen XY, Li LG, Huang JR, Ruan YL, Wang XM, Li L. 2012. Translate Plant Metabolism into Modern Agriculture: A Starting Point. *Molecular Plant* **5**(2): 291-293.

Cramer GR, Urano K, Delrot S, Pezzotti M, Shinozaki K. 2011. Effects of abiotic stress on plants: a systems biology perspective. *BMC Plant Biol* **11**: 163.

Crick F. 1970. Central dogma of molecular biology. *Nature* **227**(5258): 561-563.

Davuluri RV, Sun H, Palaniswamy SK, Matthews N, Molina C, Kurtz M, Grotewold E. 2003. AGRIS: Arabidopsis gene regulatory information server, an information resource of Arabidopsis cis-regulatory elements and transcription factors. *BMC Bioinformatics* **4**: 25.

De Vos M, Van Oosten VR, Van Poecke RM, Van Pelt JA, Pozo MJ, Mueller MJ, Buchala AJ, Metraux JP, Van Loon LC, Dicke M, Pieterse CM. 2005. Signal signature and transcriptome changes of Arabidopsis during pathogen and insect attack. *Mol Plant Microbe Interact* **18**(9): 923-937.

Doebley JF, Gaut BS, Smith BD. 2006. The molecular genetics of crop domestication. *Cell* **127**(7): 1309-1321.

Dornan D, Bheddah S, Newton K, Ince W, Frantz GD, Dowd P, Koeppen H, Dixit VM, French DM. 2004. COP1, the negative regulator of p53, is overexpressed in breast and ovarian adenocarcinomas. *Cancer Research* **64**(20): 7226-7230.

Eckardt NA. 2011. Plant Science in the Service of Human Health and Nutrition. *Plant Cell* **23**(7): 2476-2476.

Eulgem T. 2005. Regulation of the Arabidopsis defense transcriptome. *Trends Plant Sci* **10**(2): 71-78.

Falloon P, Betts R. 2010. Climate impacts on European agriculture and water management in the context of adaptation and mitigation-The importance of an integrated approach. *Science of the Total Environment* **408**(23): 5667-5687.

Ferro CA, Jupp TE, Lambert FH, Huntingford C, Cox PM. 2012. Model complexity versus ensemble size: allocating resources for climate prediction. *Philos Transact A Math Phys Eng Sci* **370**(1962): 1087-1099.

Ficklin SP, Feltus FA. 2011. Gene Coexpression Network Alignment and Conservation of Gene Modules between Two Grass Species: Maize and Rice. *Plant Physiology* **156**(3): 1244-1256.

Fields S, Johnston M. 2005. Cell biology. Whither model organism research? *Science* **307**(5717): 1885-1886.

Filiault DL, Maloof JN. 2012. A genome-wide association study identifies variants underlying the Arabidopsis thaliana shade avoidance response. *PLoS Genet* **8**(3): e1002589.

Fujita M, Fujita Y, Noutoshi Y, Takahashi F, Narusaka Y, Yamaguchi-Shinozaki K, Shinozaki K. 2006. Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. *Current Opinion in Plant Biology* **9**(4): 436-442.

Fukushima A, Kusano M, Redestig H, Arita M, Saito K. 2009. Integrated omics approaches in plant systems biology. *Current Opinion in Chemical Biology* **13**(5-6): 532-538.

Gilroy S, Trewavas A. 2001. Signal processing and transduction in plant cells: the end of the beginning? *Nat Rev Mol Cell Biol* **2**(4): 307-314.

Godbold DL. 1998. Stress concepts and forest trees. *Chemosphere* **36**(4-5): 859-864.

Gornall J, Betts R, Burke E, Clark R, Camp J, Willett K, Wiltshire A. 2010. Implications of climate change for agricultural productivity in the early twenty-first century. *Philosophical Transactions of the Royal Society B-Biological Sciences* **365**(1554): 2973-2989.

Guo A, He K, Liu D, Bai S, Gu X, Wei L, Luo J. 2005. DATF: a database of Arabidopsis transcription factors. *Bioinformatics* **21**(10): 2568-2569.

Gutierrez RA, Shasha DE, Coruzzi GM. 2005. Systems biology for the virtual plant. *Plant Physiology* **138**(2): 550-554.

Hall AE, Fiebig A, Preuss D. 2002. Beyond the Arabidopsis genome: Opportunities for comparative genomics. *Plant Physiology* **129**(4): 1439-1447.

Hammer GL, Sinclair TR, Chapman SC, van Oosterom E. 2004. On systems thinking, systems biology, and the in silico plant. *Plant Physiology* **134**(3): 909-911.

Herrgard MJ, Covert MW, Palsson BO. 2004. Reconstruction of microbial transcriptional regulatory networks. *Current Opinion in Biotechnology* **15**(1): 70-77.

Hesse H, Hoefgen R. 2006. On the way to understand biological complexity in plants: S-nutrition as a case study for systems biology. *Cellular & Molecular Biology Letters* **11**(1): 37-55.

Higley LG, Browde JA, Higley PM 1993. Moving towards new understandings of biotic stress and stress interactions,. In: Buxton DR, Shibles R, Forsberg RA, Blad BL, Asay KH, Paulson GM, Wilson RF eds. *International Crop Science*. Madison, WI: Crop Science Society of America, 749.

Hirt H, ed. 2009. *Plant Stress Biology:From Genomics to Systems Biology* . Willey-VCH.

Horton MW, Hancock AM, Huang YS, Toomajian C, Atwell S, Auton A, Muliyati NW, Platt A, Sperone FG, Vilhjalmsson BJ, Nordborg M, Borevitz JO, Bergelson J. 2012. Genome-wide patterns of genetic variation in worldwide Arabidopsis thaliana accessions from the RegMap panel. *Nat Genet* **44**(2): 212-216.

Ideker T, Krogan NJ. 2012. Differential network biology. *Mol Syst Biol* **8**: 565.

Iglesias A, Quiroga S, Diz A. 2011. Looking into the future of agriculture in a changing climate. *European Review of Agricultural Economics* **38**(3): 427-447.

J.P. Grime, ed. 1991. *Nutrition, environment and plant ecology: an overview*. Plant growth: interactions with nutrition and environment. Cambridge: Cambridge University Press.

Jones AM, Chory J, Dangl JL, Estelle M, Jacobsen SE, Meyerowitz EM, Nordborg M, Weigel D. 2008. The impact of Arabidopsis on human health: Diversifying our portfolio. *Cell* **133**(6): 939-943.

Joyce AR, Palsson BO. 2006. The model organism as a system: integrating 'omics' data sets. *Nature Reviews Molecular Cell Biology* **7**(3): 198-210.

Kao KC, Yang YL, Boscolo R, Sabatti C, Roychowdhury V, Liao JC. 2004. Transcriptome-based determination of multiple transcription regulator activities in Escherichia coli by using network component analysis. *Proc Natl Acad Sci U S A* **101**(2): 641-646.

Kilian J, Whitehead D, Horak J, Wanke D, Weinl S, Batistic O, D'Angelo C, Bornberg-Bauer E, Kudla J, Harter K. 2007. The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *Plant J* **50**(2): 347-363.

Kitano H. 2002. Systems biology: a brief overview. *Science* **295**(5560): 1662-1664.

Kitano H. 2007. Towards a theory of biological robustness. *Mol Syst Biol* **3**: 137.

Kliebenstein DJ. 2010. Systems biology uncovers the foundation of natural genetic diversity. *Plant Physiology* **152**(2): 480-486.

Kozlowski TT, Kramer, P.J., Pallardy, S.G. 1991. *The Physiological Ecology of Woody Plants*. USA: San Diego, Academic Press.

Kultz D. 2005. Molecular and evolutionary basis of the cellular stress response. *Annu Rev Physiol* **67**: 225-257.

Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander DL, Garcia-Hernandez M, Karthikeyan AS, Lee CH, Nelson WD, Ploetz L, Singh S, Wensel A, Huala E. 2012. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Research* **40**(Database issue): D1202-1210.

Lee EK, Jin YW, Park JH, Yoo YM, Hong SM, Amir R, Yan Z, Kwon E, Elfick A, Tomlinson S, Halbritter F, Waibel T, Yun BW, Loake GJ. 2010. Cultured cambial meristematic cells as a source of plant natural products. *Nature Biotechnology* **28**(11): 1213-1217.

Lee I, Ambaru B, Thakkar P, Marcotte EM, Rhee SY. 2010. Rational association of genes with traits using a genome-scale gene network for Arabidopsis thaliana. *Nature Biotechnology* **28**(2): 149-U114.

Lloyd AM, Barnason AR, Rogers SG, Byrne MC, Fraley RT, Horsch RB. 1986. Transformation of Arabidopsis thaliana with Agrobacterium tumefaciens. *Science* **234**(4775): 464-466.

Lobell DB, Burke MB, Tebaldi C, Mastrandrea MD, Falcon WP, Naylor RL. 2008. Prioritizing climate change adaptation needs for food security in 2030. *Science* **319**(5863): 607-610.

Lobell DB, Schlenker W, Costa-Roberts J. 2011. Climate Trends and Global Crop Production Since 1980. *Science* **333**(6042): 616-620.

Lopez-Maury L, Marguerat S, Bahler J. 2008. Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nat Rev Genet* **9**(8): 583-593.

Martin C, Butelli E, Petroni K, Tonelli C. 2011. How Can Research on Plants Contribute to Promoting Human Health? *Plant Cell* **23**(5): 1685-1699.

Mendelsohn E. 1963. A History of Biology to About the Year 1900 - a General Introduction to the Study of Living Things - Singer,C. *Isis* **54**(177): 415-417.

Meyerowitz EM. 2001. Prehistory and history of Arabidopsis research. *Plant Physiology* **125**(1): 15-19.

Minorsky PV. 2003. Frontiers of plant cell biology: signals and pathways, system-based approaches 22nd Symposium in Plant Biology (University of California--Riverside). *Plant Physiology* **132**(2): 428-435.

Mitchell-Olds T, Schmitt J. 2006. Genetic mechanisms and evolutionary significance of natural variation in Arabidopsis. *Nature* **441**(7096): 947-952.

Mittler R. 2006. Abiotic stress, the field environment and stress combination. *Trends Plant Sci* **11**(1): 15-19.

Mittler R, Blumwald E. 2010. Genetic engineering for modern agriculture: challenges and perspectives. *Annu Rev Plant Biol* **61**: 443-462.

Mochida K, Shinozaki K. 2010. Genomics and bioinformatics resources for crop improvement. *Plant Cell Physiol* **51**(4): 497-523.

Mochida K, Shinozaki K. 2011. Advances in omics and bioinformatics tools for systems analyses of plant functions. *Plant Cell Physiol* **52**(12): 2017-2038.

Mundy J, Nielsen HB, Brodersen P. 2006. Crosstalk. *Trends Plant Sci* **11**(2): 63-64.

Ney B, Tivoli B, Bancal MO. 2000. Effect of diseases on plant growth and crop development - Modeling of losses. *Behavior of Plants Subject to Environmental Stress*(93): 481-496.

Nilsen ET, Orcutt DM. 1996. *The physiology of plants under stress*. New York: Wiley.

Oltvai ZN, Barabasi AL. 2002. Systems biology. Life's complexity pyramid. *Science* **298**(5594): 763-764.

Park JM, Park CJ, Lee SB, Ham BK, Shin R, Paek KH. 2001. Overexpression of the tobacco Tsi1 gene encoding an EREBP/AP2-type transcription factor enhances resistance against pathogen attack and osmotic stress in tobacco. *Plant Cell* **13**(5): 1035-1046.

Parry G, Estelle M. 2006. Auxin receptors: a new role for F-box proteins. *Current Opinion in Cell Biology* **18**(2): 152-156.

Peterson RKD, Higley LG. 2000. *Biotic Stress and Yield Loss*. USA: CRC Press.

Qi L, Heredia JE, Altarejos JY, Screaton R, Goebel N, Niessen S, MacLeod IX, Liew CW, Kulkarni RN, Bain J, Newgard C, Nelson M, Evans RM, Yates J, Montminy M. 2006. TRB3 links the E3 ubiquitin ligase COP1 to lipid metabolism. *Science* **312**(5781): 1763-1766.

Raikhel NV, Coruzzi GM. 2003. Achieving the in silico plant. Systems biology and the future of plant biological research. *Plant Physiology* **132**(2): 404-409.

Rasmussen S, Barah P, Suarez-Rodriguez MC, Friis P, Costantino P, Bones AM, Nielsen HB, Mundy J. in press. Transcriptome response to combinations of stress in Arabidopsis thaliana. *Plant Physiology*: (In press).

Rind D. 1999. Complexity and climate. *Science* **284**(5411): 105-107.

Rizhsky L, Liang H, Mittler R. 2002. The combined effect of drought stress and heat shock on gene expression in tobacco. *Plant Physiology* **130**(3): 1143-1151.

Sablowski R. 2004. Plant and animal stem cells: conceptually similar, molecularly distinct? *Trends Cell Biol* **14**(11): 605-611.

Salse J. 2012. In silico archeogenomics unveils modern plant genome organisation, regulation and evolution. *Current Opinion in Plant Biology* **15**(2): 122-130.

Schena M, Shalon D, Davis RW, Brown PO. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**(5235): 467-470.

Segal E, Shapira M, Regev A, Pe'er D, Botstein D, Koller D, Friedman N. 2003. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nature Genetics* **34**(2): 166-176.

Shameer K, Ambika S, Varghese SM, Karaba N, Udayakumar M, Sowdhamini R. 2009. STIFDB-Arabidopsis Stress Responsive Transcription Factor DataBase. *Int J Plant Genomics* **2009**: 583429.

Shiu SH, Bleecker AB. 2001. Receptor-like kinases from Arabidopsis form a monophyletic gene family related to animal receptor kinases. *Proc Natl Acad Sci U S A* **98**(19): 10763-10768.

Singh AH, Wolf DM, Wang P, Arkin AP. 2008. Modularity of stress response evolution. *Proc Natl Acad Sci U S A* **105**(21): 7500-7505.

Singh KB, Foley RC, Onate-Sanchez L. 2002. Transcription factors in plant defense and stress responses. *Current Opinion in Plant Biology* **5**(5): 430-436.

Somerville C, Koornneef M. 2002. Timeline - A fortunate choice: the history of Arabidopsis as a model plant. *Nature Reviews Genetics* **3**(11): 883-889.

Spiertz JHJS, P.C.; Laar, H.H. van (Eds.). 2007. *Scale and Complexity in Plant Systems Research*: Springer

Takeda S, Matsuoka M. 2008. Genetic approaches to crop improvement: responding to environmental and population changes. *Nat Rev Genet* **9**(6): 444-457.

Ting JPY, Kastner DL, Hoffman HM. 2006. CATERPILLERs, pyrin and hereditary immunological disorders. *Nature Reviews Immunology* **6**(3): 183-195.

Tirosh I, Barkai N. 2011. Inferring regulatory mechanisms from patterns of evolutionary divergence. *Mol Syst Biol* **7**: 530.

Trewavas A. 2006. A brief history of systems biology. "Every object that biology studies is a system of systems." Francois Jacob (1974). *Plant Cell* **18**(10): 2420-2430.

Trontin C, Tisne S, Bach L, Loudet O. 2011. What does Arabidopsis natural variation teach us (and does not teach us) about adaptation in plants? *Current Opinion in Plant Biology* **14**(3): 225-231.

Tuteja N, Sopory SK. 2008a. Chemical signaling under abiotic stress environment in plants. *Plant Signal Behav* **3**(8): 525-536.

Tuteja N, Sopory SK. 2008b. Plant signaling in stress: G-protein coupled receptors, heterotrimeric G-proteins and signal coupling via phospholipases. *Plant Signal Behav* **3**(2): 79-86.

Van Norman JM, Benfey PN. 2009. Arabidopsis thaliana as a Model Organism in Systems Biology. *Wiley Interdiscip Rev Syst Biol Med* **1**(3): 372-379.

Vitousek PM, Mooney HA, Lubchenco J, Melillo JM. 1997. Human domination of Earth's ecosystems. *Science* **277**(5325): 494-499.

Vlieghe K, Boudolf V, Beemster GTS, Maes S, Magyar Z, Atanassova A, Engler JD, De Groodt R, Inze D, De Veylder L. 2005. The DP-E2F-like gene DEL1 controls the endocycle in Arabidopsis thaliana. *Current Biology* **15**(1): 59-63.

Wang Y, Bao Z, Zhu Y, Hua J. 2009. Analysis of temperature modulation of plant defense against biotrophic microbes. *Mol Plant Microbe Interact* **22**(5): 498-506.

Wei N, Deng XW. 2003. The COP9 signalosome. *Annual Review of Cell and Developmental Biology* **19**: 261-286.

Weigel D. 2012. Natural variation in Arabidopsis: from molecular genetics to ecological genomics. *Plant Physiology* **158**(1): 2-22.

Weigel D, Nordborg M. 2005. Natural variation in arabidopsis. How do we find the causal genes? *Plant Physiology* **138**(2): 567-568.

Yasuda M, Ishikawa A, Jikumaru Y, Seki M, Umezawa T, Asami T, Maruyama-Nakashita A, Kudo T, Shinozaki K, Yoshida S, Nakashita H. 2008. Antagonistic interaction between systemic acquired resistance and the abscisic acid-mediated abiotic stress response in Arabidopsis. *Plant Cell* **20**(6): 1678-1692.

Yilmaz A, Mejia-Guerra MK, Kurz K, Liang XY, Welch L, Grotewold E. 2011. AGRIS: the Arabidopsis Gene Regulatory Information Server, an update. *Nucleic Acids Res* **39**: D1118-D1122.

Yin X, Struik P 2007. Crop systems biology: an approach to connect functional genomics with crop modelling. In: J H J Spiertz PCS, H H van Laar ed. *Scale and complexity in plant systems research: gene plant-crop relations*. Wageningen, The Netherlands Springer, 63-73.

Yuan JS, Galbraith DW, Dai SY, Griffin P, Stewart CN, Jr. 2008. Plant systems biology comes of age. *Trends Plant Sci* **13**(4): 165-171.

Zhang JX, Klueva NY, Wang Z, Wu R, Ho THD, Nguyen HT. 2000. Genetic engineering for abiotic stress resistance in crop plants. *In Vitro Cellular & Developmental Biology-Plant* **36**(2): 108-114.

Zhang JZ, Creelman RA, Zhu JK. 2004. From laboratory to field. Using information from Arabidopsis to engineer salt, cold, and drought tolerance in crops. *Plant Physiology* **135**(2): 615-621.

Zhu Y, Qian W, Hua J. 2010. Temperature modulates plant defense responses through NB-LRR proteins. *PLoS Pathog* **6**(4): e1000844.

CHAPTER 2

(Paper I)

Is not included due to copyright

CHAPTER 3

(Paper II)

OPEN ACCESS Freely available online

PLOS ONE

Molecular Signatures in *Arabidopsis thaliana* in Response to Insect Attack and Bacterial Infection

Pankaj Barah, Per Winge, Anna Kusnierczyk, Diem Hong Tran, Atle M. Bones*

Department of Biology, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

Abstract

Background: Under the threat of global climatic change and food shortages, it is essential to take the initiative to obtain a comprehensive understanding of common and specific defence mechanisms existing in plant systems for protection against different types of biotic invaders. We have implemented an integrated approach to analyse the overall transcriptomic reprogramming and systems-level defence responses in the model plant species *Arabidopsis thaliana* (*A. thaliana* henceforth) during insect *Brevicoryne brassicae* (*B. brassicae* henceforth) and bacterial *Pseudomonas syringae pv. tomato strain DC3000* (*P. syringae* henceforth) attacks. The main aim of this study was to identify the attacker-specific and general defence response signatures in *A. thaliana* when attacked by phloem-feeding aphids or pathogenic bacteria.

Results: The obtained annotated networks of differentially expressed transcripts indicated that members of transcription factor families, such as *WRKY, MYB, ERF, BHLH and bZIP,* could be crucial for stress-specific defence regulation in *Arabidopsis* during aphid and *P. syringae* attack. The defence response pathways, signalling pathways and metabolic processes associated with aphid attack and *P. syringae* infection partially overlapped. Components of several important biosynthesis and signalling pathways, such as salicylic acid (SA), jasmonic acid (JA), ethylene (ET) and glucosinolates, were differentially affected during the two the treatments. Several stress-regulated transcription factors were known to be associated with stress-inducible microRNAs. The differentially regulated gene sets included many signature transcription factors, and our co-expression analysis showed that they were also strongly co-expressed during 69 other biotic stress experiments.

Conclusions: Defence responses and functional networks that were unique and specific to aphid or *P. syringae* stresses were identified. Furthermore, our analysis revealed a probable link between biotic stress and microRNAs in *Arabidopsis* and, thus gives indicates a new direction for conducting large-scale targeted experiments to explore the detailed regulatory links between them. The presented results provide a comparative understanding of *Arabidopsis – B. brassicae* and *Arabidopsis – P. syringae* interactions at the transcriptomic level.

Citation: Barah P, Winge P, Kusnierczyk A, Tran DH, Bones AM (2013) Molecular Signatures in *Arabidopsis thaliana* in Response to Insect Attack and Bacterial Infection. PLoS ONE 8(3): e58987. doi:10.1371/journal.pone.0058987

Editor: Boris Alexander Vinatzer, Virginia Tech, United States of America

Received September 23, 2012; **Accepted** February 9, 2013; **Published** March 25, 2013

Funding: This work was supported by the Biotechnology and Functional Genomics (FUGE) program of the Norwegian Research Council through grants NFR 184147/S10, 185173/V40 and 184146/S10. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: atle.bones@bio.ntnu.no

Introduction

Plants are sessile organisms that are unable to escape biotic and abiotic stresses. As a result, they have evolved flexibility in their responses to changing environmental conditions, such as light, drought, temperature, the available nutritional supply and biotic invasion. Different types of biotic invasions, such as insect, bacterial, fungal and viral invasions, represent a severe threat to agricultural production worldwide [1]. Some responses of host plants to different stress conditions are very general and provide protection from a variety of invading organisms, whereas others are more specific and target particular types of attackers. Highly complex and often connected signalling pathways, regulating numerous metabolic networks, coordinate plant responses to different stress conditions. Over the last decade or so, clear advances have been made in understanding how defence responses are orchestrated in higher plants. The development of microarray technology has allowed monitoring of expressional changes in thousands of genes simultaneously, and this technology has now become a major tool for examining plant stress biology. Most of these studies have adopted *A. thaliana* as a model plant organism because of the vast amount of genomic information made available for this species with the completion of the *A. thaliana* genome sequence and advanced annotation of *A. thaliana* genes [2]. Analysing the regulation of gene expression under various stress conditions has revealed that the early defence responses of a plant to different stress factors often overlap and engage the same sets of genes [3]. It has also become evident that different types of plant invaders may induce substantially different changes in the host plant transcriptome. Furthermore, studies on plants subjected to various treatments indicate that the induced defences can be both general – being commonly manifested regardless of the type of applied treatment; and specific – providing protection from a certain type of stress [4,5]. In many cases, however, the multidimensional level of network crosstalk

makes it challenging to recognise which of the observed responses are general and which are more stress specific [6,7].

Aphids are one of the world's major insect pests, causing serious economic damage to a range of temperate and tropical crops [8]. Aphids use their mouthparts, formed into a stylet-like structure, to pierce plant tissue in the search for sieve elements (SEs) containing their primary food source: phloem sap [9,10]. Feeding by an aphid causes minimal wounding, as its stylet proceeds mostly intercellulary and is inserted only into selected cells on its way to the phloem tissue [11]. However, the disruption of cell walls and membranes of the pierced cells is likely to be the first factor triggering a plant response. In addition, the salivary secretions lubricating the stylet throughout its pathway through plants tissues and injected into SEs during feeding contain molecular signatures that activate plant defences. Therefore, despite their stealthy feeding, aphids are strong inducers of plant defences against them. Recently Kuśnierczyk *et al.* reported the timing and dynamics of early *Arabidopsis* defence responses [12] to an aphid attack.

P. syringae is a bacterial leaf pathogen that causes extensive chlorosis and necrotic spots [13]. Many strains of *P. syringae* are pathogenic in the model plant *A. thaliana*, and *P. syringae* is therefore widely used to study plant – pathogen interactions under laboratory conditions. *P. syringae* enters host tissues through wounds or natural openings such as stomata, and in susceptible plants, it multiplies to high concentrations in intercellular spaces [14]. The ability of *P. syringae* to multiply endophytically is dependent on its type III secretion pathway enabling the secretion of proteins into the apoplast. These proteins interact with the cell wall and plasma membrane and are directly translocated into the cytoplasm of host cells [15]. Several strains of *P. syringae* produce coronatine, a molecule that mimics endogenous plant jasmonyl-L-isoleucine and an activator of the jasmonic acid signalling pathway [16]. By doing so, the bacteria manipulate host responses, suppressing salicylic acid defences through the activation of jasmonic acid signalling [17,18].

A great number of experiments conducted to assess plant responses to different stresses have made substantial contributions to our understanding of the induced defences of plants. However, the comparison of independent experiments and extraction of meaningful information from such comparisons is complicated and difficult in most cases, mainly due to the lack of common standards regarding how to grow plants, conduct expression profile experiments, and finally, how to evaluate the resulting gene expression data [19]. In recent years, integrated approaches, such as systems biology methods, have been evolving, providing promising tools for studying plant stress responses [20,21]. Scientists intend to go beyond simple functional enrichment analyses to understand the molecular basis of genome-scale microarray experiments. Methods inspired by systems biology utilise lists of differentially expressed genes ranked by biological criteria to search for the distribution of blocks of functionally related genes without imposing any artificial threshold. Such ranked lists of genes can be arranged into functional classes, pathways and biological processes. Co-expression or co-regulation of particular genes can indicate their involvement in similar biological processes, meaning that individual modules of genes can be attributed to specific biological processes. Using this basic concept, modular network topology-based analysis has been proven to be useful in identifying functional modules of genes [22]. In a recent co-expression study, Weston and co-workers showed how a co-expression network-based analysis can be used for understanding population-level adaptive physiological responses of plants to abiotic stress [23].

MicroRNAs (microRNAs) are small, non-coding RNAs that play critical roles in post-transcriptional gene regulation and stress-inducible transcriptional regulation in *Arabidopsis* [24]. In plants, mature microRNAs pair with complementary sites on mRNAs, subsequently leading to the cleavage and degradation of the mRNAs. Many microRNAs target mRNAs that encode transcription factors and, thus, influence the expression of many genes whose regulation is controlled by these transcription factors [25]. The identification, detection, regulation and functional analysis of microRNAs associated with biotic stress remains a great challenge. In contrast, information about plant stress-responsive genes and their transcription factor binding sites is available to some extent in several databases [26,27,28,29,30]. Integration of such publicly available knowledge bases with experimental approaches would provide useful insights in understanding the plant defence responses to different biotic stresses.

In this manuscript, we present such an integrated approach to explore the common (general) and attacker-specific defence responses of *A. thaliana* subjected to two different types of biotic invaders: phloem-feeding aphids (*B. brassicae*) and pathogenic bacteria (*P. syringae*). To allow comparison between the obtained gene expression profiles and the observed regulation of gene pathways involved in defence against the aphid and the bacterium, the same growth and experimental conditions were used in the two simultaneous experimental setups. Transcriptional changes resulting either from infestation with *B. brassicae* or infection with *P. syringae* were assessed with the use of full-genome Arabidopsis microarrays (the data have been deposited in GEO with accession numbers GSE39245 and GSE39246).

Two sets of differentially expressed genes, corresponding to the plant responses to either aphid or bacterial treatment, were created as the outcome of the microarray data analysis. In an attempt to integrate the resulting data with publicly available knowledge extracted from several different databases as well as from published results of other experiments, these two differentially regulated gene sets were subsequently analysed through a set of computational approaches. The following analyses were incorporated into the presented work: an analysis of enriched functional categories or processes; exploration of potential connections between microRNAs and biotic stress-inducible transcriptional regulation during insect and bacterial attack; cross-validation of the aphid- and *Pseudomonas*-regulated genes using a co-expression network constructed from a compendium of 69 other biotic stress microarray datasets complied in the CORNET tool [31] (https://cornet.psb.ugent.be/).

Results and Discussion

Overall Changes in the *Arabidopsis* Transcriptome in Response to Insect and Bacterial Attack

To explore the complexity of the transcriptional changes induced by the different examined *A. thaliana* attackers, we compared the overlap between the obtained gene sets. From the results, it is evident that the transcriptional responses of *A. thaliana* to these very different attackers are massive. Aphid infestation and *P. syringae* infection resulted in significant differential regulation of 4,979 (2,803 up-regulated, 2,176 down-regulated) and 3,199 (1,634 up, 1,565 down) genes, respectively (**Table 1** and **Tables S4, S5**). Although aphids and bacteria exhibit very different modes of action and trigger a highly dissimilar signal signature, a large number of *Arabidopsis* genes were expressed in response to both attackers. There were 1,597 common genes affected after both aphid infestation and *P. syringae* treatment. A total of 3,382 genes (1,963 up, 1,419 down) showed aphid-specific expression,

while 1602 genes (842 up, 760 down) showed *P. syringae*-specific expression (**Table S6**). In the common set of genes, there were a total of 186 genes that showed opposite expression patterns in the two experiments. Of these genes, 117 were up-regulated under aphid and down-regulated under *P. syringae* attack, while 69 genes were down-regulated under aphid and up-regulated under *P. syringae* attack. Out of the 117 genes that were up-regulated in the aphid and down-regulated in the *P. syringae* experiment, 17 have been reported to be transcription factors. Six of these transcription factors are members of the *ERF/AP2* transcription factor family. Among them *ERF104*, which is regulated by *MPK6*, is a key controller of innate immunity and dehydration stress [32].

In total, 303 transcription factors were found to be affected by the aphid treatment, while 191 transcription factors showed altered expression under *P. syringae* infection. The common category (differentially expressed during both aphid infestation and *P. syringae* treatment) included 87 known Arabidopsis transcription factors. The analysis also identified 216 transcription factors that were differentially regulated only during the aphid treatment and 104 transcription factors that were differentially regulated only during *P. syringae* infection. The annotated network of these transcripts showed that some of the differentially expressed transcription factors could be crucial for stress-specific defence responses in *A. thaliana* plants.

Analysis of overrepresented gene ontologies (GO) in *A. thaliana* indicates rigorous reprogramming of several biological processes. As seen from the **Table 1**, a large number of genes were differentially regulated in *A. thaliana* during both the aphid and *P. syringae* experiments, which indicated that intense transcriptional reprogramming took place. A network-based analysis of the corresponding GO terms under the *Biological Process* classification using ClueGO (correction method = Bonferroni, kappa score $\geq$0.3) in the common aphid-specific and *P. syringae*-specific transcript dataset was performed.

When this analysis was applied to the list of 1,597 common genes whose expression was affected during both of the experiments, 17 significantly overrepresented categories were identified (some of these categories are shown in **Figure 1**.) Most of the cellular and metabolic processes were clustered in distinctly separate modules, and there were few highly interconnecting overrepresented processes. More than half of the genes from the common list were involved in central metabolic and cellular processes, such as electron transport and energy pathways located in the plastid. Some of the most significant categories were indole-containing compound metabolic processes, host localised cell death, cellular responses to starvation, downregulation of photosynthesis, responses to jasmonic acid, sulphur compound biosynthetic processes, and negative regulation of cellular processes. Analysis of the modules showed that the majority of the jasmonic acid responsive genes were up-regulated by both treatments, but the number of genes and their degree of induction were markedly higher in the *P. syringae*-treated plants, which may be due to the effects of coronatine (**file S11**). It has been previously reported that *P. syringae* uses the virulence factor coronatine (*COR*) as a mimic of jasmonyl-l-isoleucine (*JA-Ile*) [16,33]. The coronatin-regulated *A. thaliana* genes reported in Thilmony R. et al., 2006 [34] show strong overlap with our *P. syringae* data. More than 450 genes reported as coronatin-regulated by Thilmony R. et al. show highly similar expression patterns in the two datasets (data not shown).

Tryptophan-derived indolic compounds, such as indolic glucosinolates (*iGS*) and indolic-derived phytoalexins, are an important component elicitor-induced responses in Arabidopsis plants [35,36]. The biosynthesis of tryptophan-derived indolic compounds was up-regulated under both treatments but was stronger induced by aphid infestation (**file S4** and **S5**). Two of the affected modules, cellular responses to starvation and sugar-mediated signalling pathways, further indicated that both treatments resulted in cells experiencing a nutrient deficiency. Although we did not analyse cellular nutrient deficiency in the plants during our experiments, the profiles observed here are in agreement with existing information in annotation databases such as TAIR (release 10) and Gene Ontology, which are derived from the published literature.

Localised host programmed cell death is a crucial mechanism through which plants respond to pathogen and insect attack. This phenomenon regulates multiple physiological processes, including terminal differentiation, senescence, and disease resistance [37]. Several of the genes involved in the localised host programmed cell death categories were up-regulated during both treatments. These genes are also known to be induced by senescence and salicylic acid treatment, including the *PR* genes (PATHOGENESIS-RELATED GENE) *PR1, PR2, PR4* and *PR5*.

Visualisation of the networks of GO terms based on the aphid-specific responses (**Figure 2**) and *P. syringae*-specific responses (**Figure 3**) demonstrated the massive transcriptional responses evoked in *A. thaliana*. Most of the significant processes were related to responses to stimuli, biosynthesis of secondary metabolites, and transcriptional and posttranscriptional regulation. Superposition of the two GO term networks generated from the aphid-specific gene list and *P. syringae*-specific gene-list showed significant differences in the overrepresented GO terms. The superimposed network diagram has not been included in this manuscript, but all three networks (.cys file) have been provided as additional files (**files S1, S2,** and **S3**). The interested reader can locally open these files in Cytoscape and conduct interactive exploration. (For local visual-

Table 1. Overall summary of the differentially regulated genes in *A. thaliana* during *Brevicoryne brassicae* (aphid) attack or *P. syringae* (bacteria) infection.

Category	No. of Genes	Up- regulated	Down- regulated	No. of TF
Differentially expressed during Aphid exp.	4979	2803	2176	303
Differentially expressed during *P. syringae* exp.	3199	1634	1565	191
*Common to both exp.	1597	723	688	87
Only Aphid	3382	1963	1419	216
Only *Pseudomonas*	1602	842	760	104

*In the common set of genes, 186 genes showed opposite expression patterns during the two experiments. Among these genes, 117 were up-regulated under aphid and down-regulated under *P. syringae* attack, while 69 genes were down-regulated under aphid and up-regulated under *P. syringae* attack.
doi:10.1371/journal.pone.0058987.t001

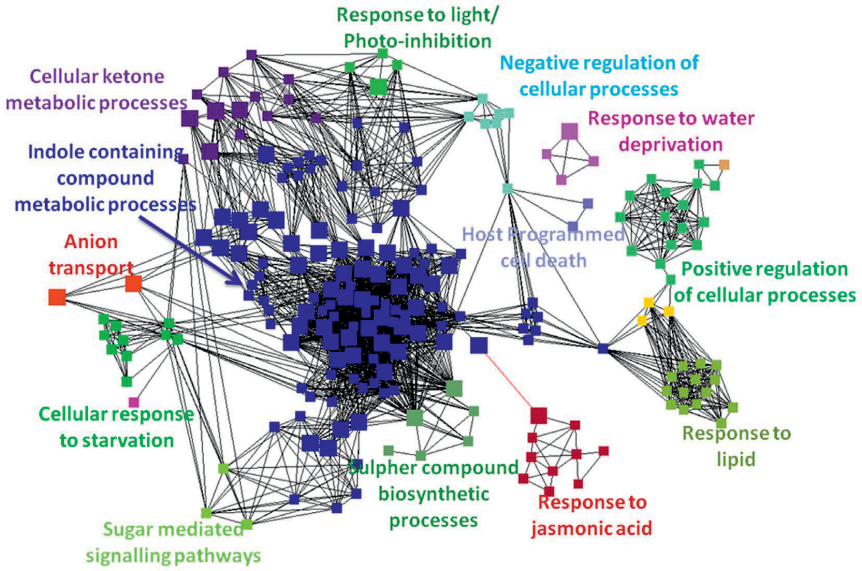


Figure 1. Over-represented GO-categories in the common gene list. Network representations of enriched GO categories among the genes that were differentially regulated during both experiments. Figure generated from the functionally grouped networks of enriched GO categories among genes whose expression is induced by both the aphid and pathogenic bacterium treatments. GO terms are represented as nodes based on their kappa score (≥0.3); only networks with at least three nodes are represented. The node size indicates the significance of the term's enrichment. The edges are related to the relationships between the selected terms, which are defined based on the genes that are shared in a similar way. The label of the most significant term is used as the leading group term. Visualisation was conducted using Cytoscape 2.7.0.
doi:10.1371/journal.pone.0058987.g001

isation, download cytoscape software from http://www.cytoscape.org/, and load the.cys files on the software. Please note that the view of the annotated network presented in this manuscript has been manually simplified for representation purposes.).

Mapping the Insect- and Bacterial-specific Responses on Pathways and Processes

To structure the genes present on the *A. thaliana* whole-genome microarray, they were assigned to functional categories using the pathway analysis program MapMan (http://gabi.rzpd.de/projects/MapMan, version 3.5.0). MapMan is a user-driven tool that displays large datasets such as gene expression data from Arabidopsis microarrays in diagrams of metabolic pathways or other processes. After the normalisation of expression values, differential fold-change values were calculated with statistical tests, as described in the Materials and Methods section. The ratios in the 4 biological replicates were averaged and converted to a log 2 scale, then imported into MapMan as '.xls' files (**files S4, S5**). MapMan converts the values to a false colour scale and displays them in diagrams. Transcripts that increase, decrease or change less than a given threshold are shown in blue, red and white, respectively. Some of the important categories (or functional BINs as per MapMan definition) identified via MapMan analysis are explained below.

Metabolism Overview Map

An overview of the transcriptional responses affecting genes coupled to metabolic processes showed that many genes connected to photosynthesis and energy metabolism were down-regulated after *P. syringae* and aphid attack (**Figure 4**). *P. syringae* infection resulted in leaf senescence and leaf yellowing, which had a major effect on chloroplast function and processes connected to the chloroplast, such as fatty acid biosynthesis, carotenoid production, chlorophyll biosynthesis, carbon fixation and others. Genes related to these processes showed clear down-regulation following *P. syringae* treatment. Secondary metabolism was strongly affected during both treatments, particularly regarding the phenylpropanoid and glucosinolate pathways. The results of *P. syringae* treatment also showed that genes connected to the terpenoid and alkaloid pathways were up-regulated, including *DXPS1*, *TPS10*, *GES/TPS04*, *SS2*, *SQE6* and *LAS1*. In general, the stress associated with the activation and continuation of defence responses is metabolically expensive, and the plant must reallocate a significant amount of the resources that would normally be used in plant growth and reproduction to the production of defence-related compounds [38,39]. However, in a recent work, Foyer et al. [40] explained that the decreases in growth and photosynthesis in response to stress are more likely the result of programmed down-regulation. Our experimental results showed that exposure to two biotic stresses resulted in the down-regulation of genes linked to auxin, gibberelin and cytokinin responses as well as genes coupled to cell wall modifications and cell division. The infected plants might also compensate for the depletion of sugars and amino acids, resulting in increased carbon assimilation and mobilisation of carbon, mannitol and nitrogen reserves. The plants may have degraded proteins/amino acids to generate energy (glycolysis) and re-assimilate nitrogen, through the glutamate dehydrogenase *GDH2* or lysine-ketoglutarate reductase (At4g33150). There were also genes connected to starch degradation/sugar responses induced, indicating that the plants might be

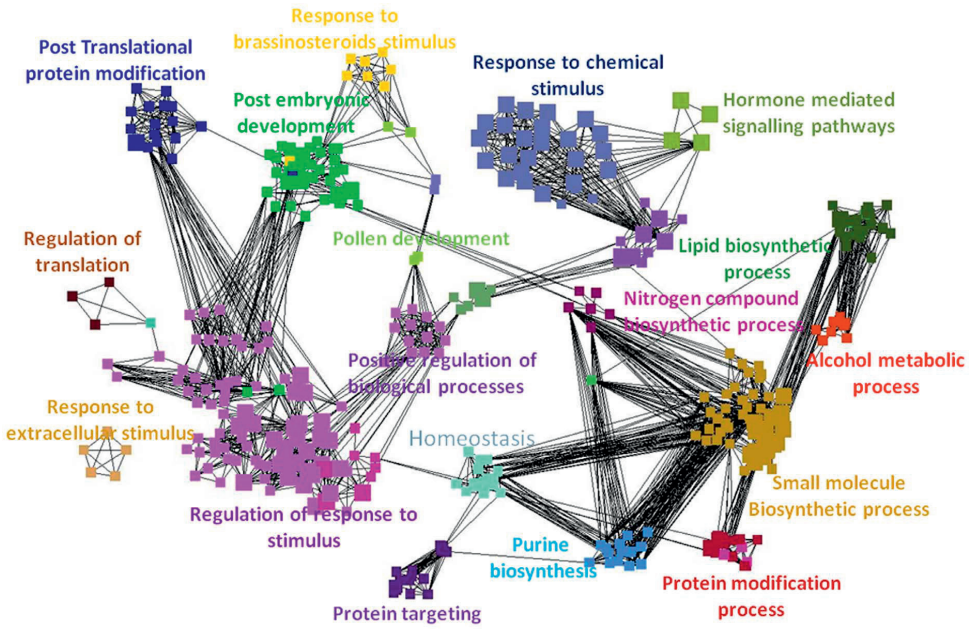


Figure 2. Over-represented GO-categories in the aphid-specific gene list. Network representations of enriched GO categories among genes that were differentially regulated only during the aphid experiment. Figure generated by ClueGO showing functionally grouped networks of enriched GO categories among genes whose expression was induced only in the aphid experiment. GO terms are represented as nodes based on their kappa score (≥0.3); only networks with at least three nodes are represented. The node size represents the significance of the term's enrichment. The edges are related to the relationships between the selected terms, which are defined based on the genes that are shared in a similar way. The label for the most significant term is used as the leading group term. Visualisation was conducted using Cytoscape 2.7.0.
doi:10.1371/journal.pone.0058987.g002

degrading starches, e.g., *BAM5* and *GPT2*, used in glycolysis. Starch biosynthesis genes were generally down-regulated. The degradation of starch and maltose may also generate an osmotic force that balances water losses. During an aphid infestation, plants suffer from osmotic stress as the insect sucks large quantities of liquids from them. To counteract this situation, the transcription of genes involved in the regulation of water balance was observed to be induced, such as the *WRKY40*, *CYP707A3* (ABA-biosynthesis), *ZAT10* and *ZAT7*.

Comparative Overview of the Response to Biotic Stress during the Aphid and *P. syringae* Treatments

A plant's reaction to biotic stress involves several steps: after the initial signal input from the pathogen, which is recognised by the corresponding receptors (putative R genes), transcription of the cascade associated with the plant defence mechanism is triggered, including changes related to oxidative stress. Inside the cell, signals are transmitted and lead to the production of defence molecules (PR proteins, heat shock proteins and secondary metabolites). A large number of signalling genes were activated during both the aphid and *P. syringae* treatments **(Figure 5)**. Most of these genes encode receptor kinases, leucine-rich receptor kinases, MAP kinases, calcium-binding proteins and proteins regulating oxidative stress, such as peroxidases (details in **file S6**). The number of signalling proteins that were differentially expressed during the aphid experiment was more than four times higher compared to the *P. syringae* treatment. There were 278 aphid-specific signalling genes, but only 62 *P. syringae*-specific signalling genes. Thirty-one heat shock proteins were differentially expressed only during the aphid treatment (**file S7**), the majority of which were of the *DnaJ/Hsp40* type chaperones and were induced. Large numbers of proteolytic enzymes were differentially expressed during both the aphid (220) and *P. syringae* (89) treatments. The majority of these enzymes were ubiquitin proteases, F-box proteins, cysteine proteases, serine proteases, *C3HC4-type RING fingers*, and metallo-proteases (**file S8**). Several of the down-regulated proteolytic enzymes were chloroplast localised or predicted to be located in the plastid/chloroplast, while most of the *C3HC4-type RING finger* proteins were induced. Secondary metabolites play a crucial role during plant defences. Sixty-three genes related to secondary metabolic processes were differentially regulated during the aphid and *P. syringae* treatments. Some of these secondary processes include the biosynthesis of isoprenoids, phenylpropanoids, glucosinolates and flavonoids. A detailed analysis of the differentially regulated secondary metabolic processes can be found in **file S9** and in a later section of this article. There were 76 differentially regulated genes connected to cell wall-related processes identified during the aphid treatment, but only 36 in the *P. syringae* experiment. These genes included components involved in cell wall precursor synthesis, cellulose synthases, cell wall structural proteins such as *AGPs* (arabinogalactan protein), *LRR* (leucine-rich repeat) extensin-like proteins, and *HRGPs* (hydroxyproline-rich glycoproteins) (details in **file S10**). In general, aphid attack appeared to affect cell wall-related processes to a greater extent than *P. syringae* infection. In particular, a large number of *APGs* and xyloglucan:xyloglucosyl transferases were observed to be down-

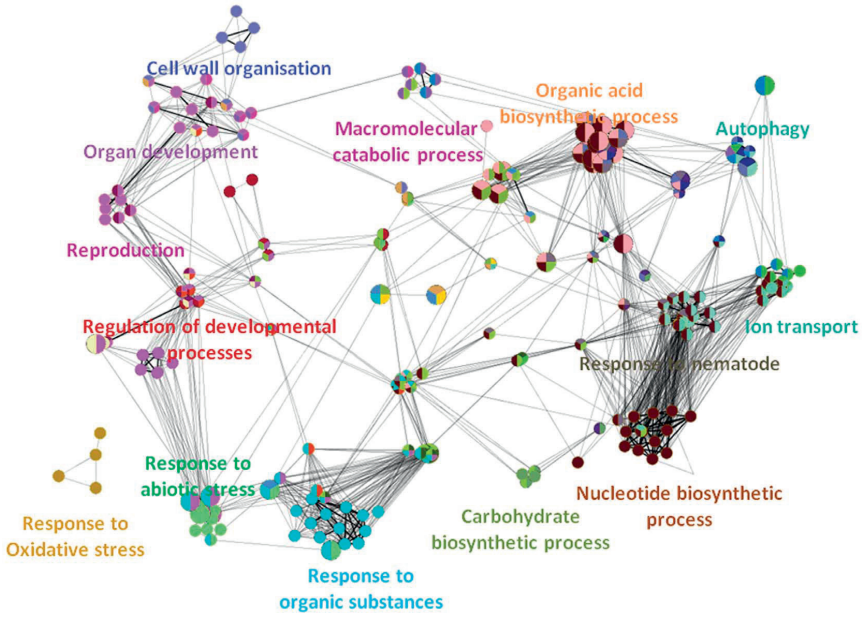


Figure 3. Over-represented GO-categories in the *P. syringae*-specific gene list. Network representations of enriched GO categories among genes that were differentially regulated only during the *P. syringae* experiment. Figure generated by ClueGO showing functionally grouped networks of enriched GO categories among genes whose expression was induced only in the *Pseudomonas* experiment. GO terms are represented as nodes based on their kappa score (≥0.3); only networks with at least three nodes are represented. The node size represents the significance of the term's enrichment. The edges are related to the relationships between the selected terms, which are defined based on the genes that are shared in a similar way. The label for the most significant term is used as the leading group term. Visualisation was conducted using Cytoscape 2.7.0.
doi:10.1371/journal.pone.0058987.g003

regulated. Xyloglucan endotransglucosylases are known to play an important role during cell elongation and cell wall modifications during shade avoidance [41,42]. The effects observed on genes encoding pathogenesis-related proteins (*PR* proteins) showed a clear bias between the treatments: while only 11 *P. syringae*-specific *PR* proteins were affected, 56 genes encoding aphid-specific *PR* proteins showed differential expression. The *PR* proteins include a wide variety of protein types, such as ß-1,3-glucαnases, chitinases, thaumatin-like protein, proteinase inhibitors, plant defensins and others. The PR1 protein, which is often used as a marker for salicylic acid responses, was more than ten-fold higher induced by the aphid attack than by *P. syringae* infection. Another class of proteins that was induced and significantly overrepresented after aphid attack corresponded to a large number of disease resistance proteins belonging to the *TIR-NBS-LRR* (Toll/Interleukin1 receptor–nucleotide binding site–leucine-rich repeat) proteins.

Among the biotic stress-related transcription factors, some *WRKY* and *bZIP* proteins were expressed differentially only during the aphid experiment, while some *MYB* proteins were expressed differentially only during *P. syringae* infection. Other differentially regulated classes of transcription factors included *ERF/AP2, NAC, bHLH* and *DOF*. Details regarding the differentially regulated transcription factors are provided in a separate section of this article. The plant defence responses associated with *P. syringae* and aphid attack induced and repressed various hormonal signalling pathways. The most affected of these pathways during our experiments were the JA, SA, ABA, ethylene and auxin pathways. Among the hormonal signalling pathways, some components of the ethylene, JA, SA, ABA, auxin and brassinosteroid pathways appeared to specifically be regulated during the aphid and *P. syringae* treatments. There were relatively few ethylene responses observed in general, but such effects were clearly stronger after the aphid than the *Pseudomonas* treatment. Examples of ethylene responses included *ACS6*, *ERF11*, which may modulate ABA-regulated ethylene biosynthesis, *ORA59*, which integrates JA and ethylene signals during plant defence, *EFE* (ethylene forming enzyme) and *ATARD3* (methionine recycling during ethylene synthesis). Some proteins involved in the biosynthesis of ethylene were also affected. JA was more strongly induced by *P. syringae*, but the SA response was stronger following aphid attack. The details of the differentially regulated genes involved in hormone-mediated signalling pathways are provided in a separate section of this article.

Regulatory Overview Map

The categories that included most of the induced regulatory genes were TFs, receptor kinases, protein degradation and protein modification. In addition, several genes involved in overrepresented induced biological processes, such as the auxin signalling pathway and autophagy, were included in the regulatory categories (**Figure 6**). A comparative list of the differentially expressed genes (both aphid-specific and *P. syringae*-specific genes) involved in hormonal pathways and their corresponding log2-transformed expression values are provided in **file S11**. The ethylene pathway was up-regulated after the aphid treatment, and genes such as *ACS6, ATARD3, ERS1* and a number of ethylene

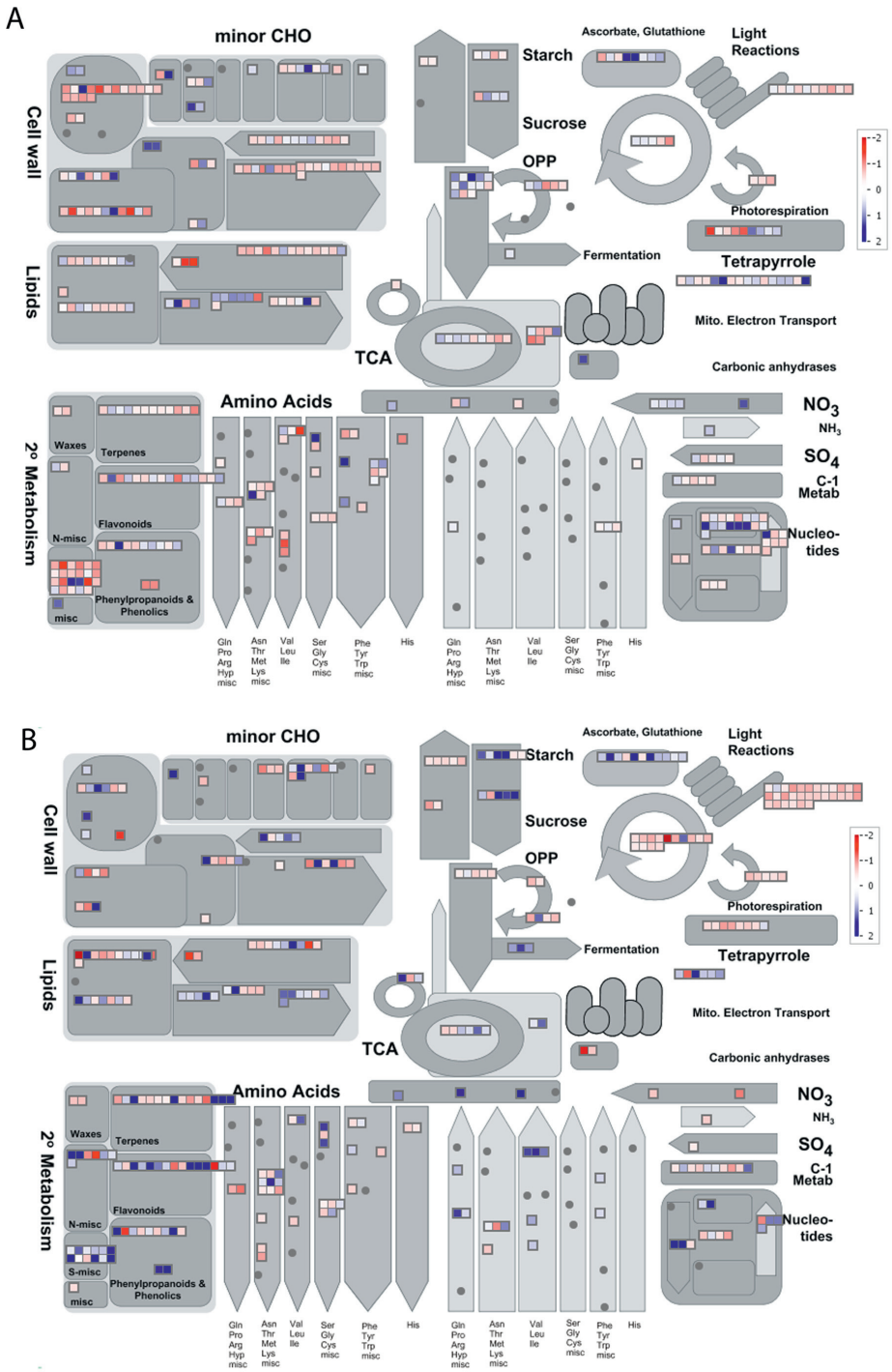

A
minor CHO
Starch
Ascorbate, Glutathione
Light Reactions
Sucrose
OPP
Photorespiration
Fermentation
Tetrapyrrole
Cell wall
Lipids
TCA
Mito. Electron Transport
Carbonic anhydrases
Amino Acids
NO3
NH3
SO4
C-1 Metab
Nucleo-tides
2° Metabolism
Waxes
Terpenes
Flavonoids
N-misc
Phenylpropanoids & Phenolics
misc
Gln Pro Arg Hyp misc
Asn Thr Met Lys misc
Val Leu Ile
Ser Gly Cys misc
Phe Tyr Trp misc
His
B
minor CHO
Starch
Ascorbate, Glutathione
Light Reactions
Sucrose
OPP
Photorespiration
Fermentation
Tetrapyrrole
Cell wall
Lipids
TCA
Mito. Electron Transport
Carbonic anhydrases
Amino Acids
NO3
NH3
SO4
C-1 Metab
Nucleo-tides
2° Metabolism
Waxes
Terpenes
Flavonoids
N-misc
S-misc
Phenylpropanoids & Phenolics
misc
Gln Pro Arg Hyp misc
Asn Thr Met Lys misc
Val Leu Ile
Ser Gly Cys misc
Phe Tyr Trp misc
His

Figure 4. Metabolic overview map. Metabolic pathways associated with the transcriptional changes affecting *A. thaliana* during aphid *and P. syringae* attack. Overview of the expression changes related to metabolic pathways observed in *A. thaliana* plants during the (**A**) aphid and (**B**) *P. syringae* treatments using MapMan software. The represented spots are only for genes showing a significant (P = 0.01) change in expression between the treatment and the untreated control that were attributed to the respective bins by MapMan. Genes whose expression levels were increased are indicated with an increasingly blue colour, while decreasing expression is indicated in red. The graduation can be seen on the scale presented in the top right corner of each subfigure. A change in expression of log2 = 2.0 scale was selected as giving full saturation.
doi:10.1371/journal.pone.0058987.g004

responsive element binding factors (*ERFs*) were induced. Genes belonging to the *ERF/AP2* family are induced by many biotic and abiotic factors, among which ethylene *ERFs* not only control a subset of ET-mediated responses but might also integrate ET with other signalling pathways. Increased ethylene production is a common defence response after herbivore attack and has been reported in several plant species [43]. Both ABA and JA responses were up-regulated by the *P. syringae* treatment, but few known SA-responsive genes were induced. Two categories, receptor kinases and calcium regulation (in **Figure 6**), appeared to be quite highly represented according to the MapMan annotation during the aphid experiment. Nevertheless, two other categories, light signalling and redox control, included fewer transcripts and gene families, respectively. These were some key differences between the aphid and *P. syringae* treatment. Genes encoding receptor kinases and proteins coupled to calcium signalling were overrepresented following the aphid treatment. These genes include a large number of cysteine-rich receptor-like protein kinases, such as, *CRK7, CRK37, CRK36, CRK23, CRK14, CRK11, CRK15, CRK6, CRK28* and calmodulin-like proteins (*CML40, CML47, CML11, TCH3, TCH2/CML24, CML44, CML45, CML30, CML37, CML38* and others) as well as several calmodulin-binding IQ-domain proteins. Together with the MAP kinases (*MPK11, MKK9, ATMPK3, MEKK3, MEK1, MEKK1, MKK2, MKK4, MPK4* and others), they constitute a large network that activates various plant defence responses, resulting in the activation of key transcription factors.

Differences Observed in the Jasmonic Acid (JA) Biosynthesis Pathway during the Aphid and P. syringae Treatments

The jasmonic acid signalling pathway is a highly conserved, powerful regulator of plant defence signalling that is activated during infection by various pathogenic microorganisms as well as upon insect attack [44]. Kuśnierczyk et al. reported that more than 200 genes are dependent on the plant's jasmonate status, irrespective of external stimuli, and that the aphid-induced response of more than 800 transcripts is regulated by jasmonate signalling [45]. The release of linolenic acid from membrane lipids initiates a series of enzymatic reactions known as the octadecanoid pathway, leading to accumulation of JA and related compounds. Additionally, 12-oxophytodieonic acid (*OPDA*) is a biosynthetic precursor of JA signalling molecules, which activate the expression of related-related genes. The selection of transcripts induced by JA and *OPDA* varies to some extent. This difference can be attributed to the electrophilic activities of the cyclopentanone ring of JA [46]. A number of enzymes coupled to oxylipin/JA biosynthesis, such as *AOC3, OPR3, OPCL1, LOX2* and *LOX3*, were up-regulated by both treatments, while *AOS, AOC1, AOC2, AOC4, ACX1, ACX5* and *LOX1* were mainly induced by *Pseudomonas*. Two *OPR*-related genes, At1g18020 and At1g17990, as well as the lipoxygenases *LOX4, LOX5* and *LOX6* were only induced by aphid attack. Almost none of the genes encoding proteins potentially linked to oxylipin biosynthesis were down-regulated, with the exception of *OPR1*, which was down-regulated by *P. syringae* infection.

SA Regulates the Expression of Aphid-specific Defence Proteins, and Methyl Salicylate Activates *P. syringae*-specific Defence Proteins

Salicylic acid is another stimulator of plant defence responses and is an important trigger of systemic acquired resistance (SAR), resulting in increased defence against a variety of pathogens. Methyl salicylate (MeSA) has been identified as one of the mobile signals required for SAR. MeSA is translocated from the site of infection through the vascular system to distal (systemic) tissues, where it activates specific defence responses. The SAR response results in a complex chain of events and is regulated by various transcription factors. In higher plants, SA can be synthesised from phenylalanine via cinamic acid or from isochorismate. During pathogen attack, SA signalling leads to accumulation of various pathogenesis-related proteins (*PR* proteins), which can possess antimicrobial and anti-insect activities. Interestingly, MeSA released by the attacked plants can be detected by insects and changes their plant preferences [47]. In our analysis, expression of a methyltransferase gene (At3g21950) related to salicylate O-methyltransferases was down-regulated during aphid treatment. At3g21950 encodes a S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase related to BSMT1 that may convert SA to MeSA. In contrast, other methyltransferase genes BSMT1 (At3g11480), which converts SA to MeSA, and UGT74E2 (At1g05680) were up-regulated during *P. syringae* treatment (**file S11**). UGT74E2 is hydrogen peroxide responsive and may be involved in water stress responses. There were relatively few known genes coupled to the biosynthesis of SA found in both datasets. However, BSMT1 might be a key enzyme.

Although relatively few genes connected to the biosynthesis of SA and MeSA were found in the obtained datasets, several genes induced by SA were identified. Additionally, isochorismate synthase 1 (*ICS1*) and one of its transcriptional regulators, *WRKY46*, were induced by aphid infestation. Another gene induced after aphid treatment that may be under the regulation of *WRKY46* is *PBS3*. *PBS3* most likely encodes an enzyme producing SA-glucoside, a putative storage form of SA, and *pbs3* mutant plants exhibit impaired activation of defence genes such as *PR1*. The PR1 gene, a common marker for SA-induced genes, was strongly up-regulated by the aphid treatment (log2 = 5.5) and slightly less induced by *P. syringae* treatment (log2 = 1.7). The *WRKY53* gene, which is known to be up-regulated by SA [48], was only induced in the aphid treatment. A number of genes, such as *ALD1* and *BAP1*, coupled to systemic defence responses were uniquely induced by aphids.

Overview of Differences in Secondary Metabolism

Plants have evolved many secondary metabolites involved in plant defence, which are collectively known as antiherbivory compounds and can be classified into three sub-groups: nitrogen compounds (including alkaloids, cyanogenic glycosides and glucosinolates), terpenoids, and phenolics [49]. In addition to the three larger groups of substances mentioned above, fatty acid derivatives, amino acids and even peptides are used in defence. The terpene synthase genes *GES* (geranyllinalool synthase, At1g61120), *LAS1* (Lanosterol synthase, At3g45130), and *TPS10*

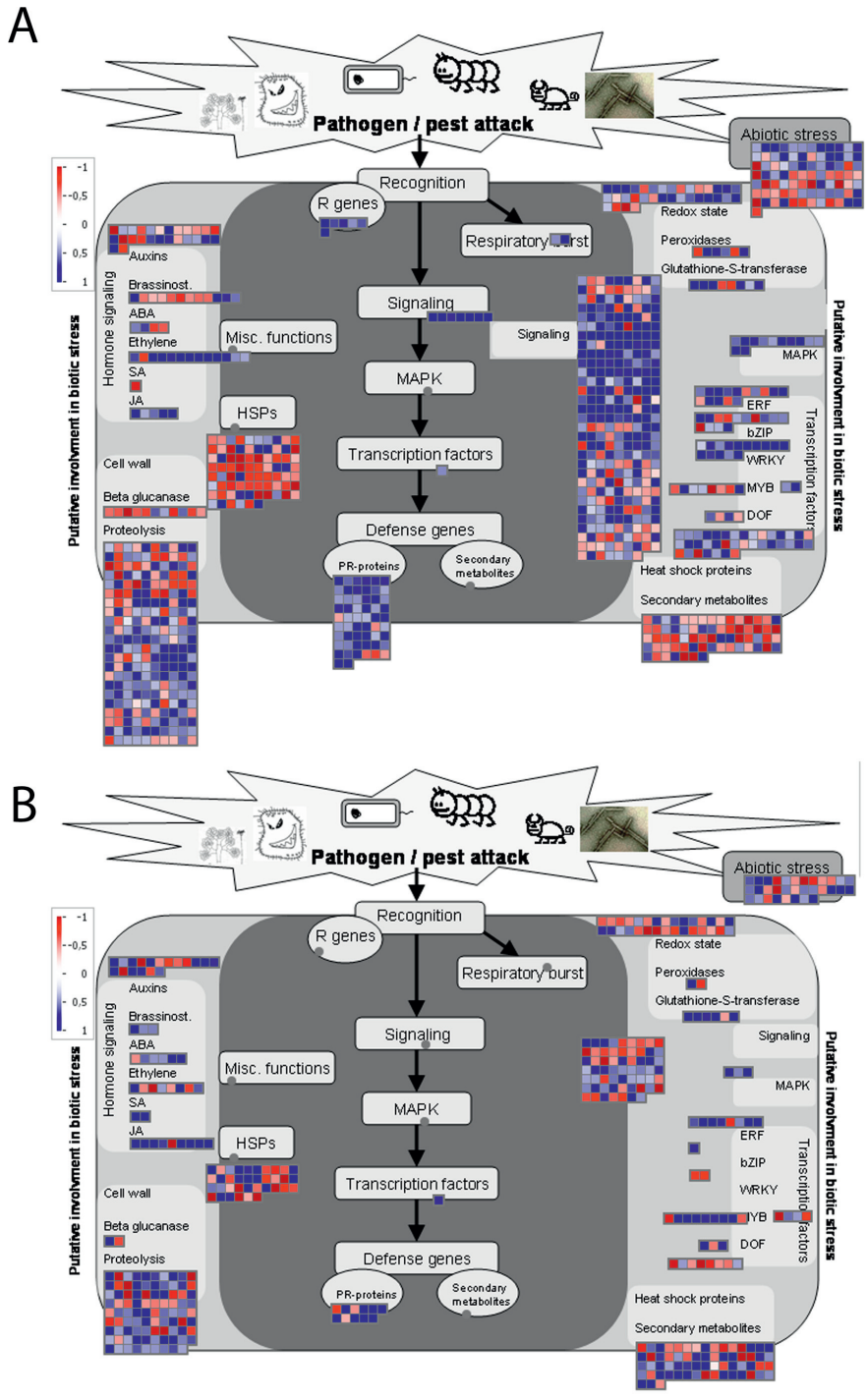

A
Pathogen / pest attack
Abiotic stress
Recognition
R genes
Respiratory burst
Signaling
MAPK
Transcription factors
Defense genes
PR-proteins
Secondary metabolites
Misc. functions
HSPs
Hormone signaling
Auxins
Brassinost.
ABA
Ethylene
SA
JA
Cell wall
Beta glucanase
Proteolysis
Redox state
Peroxidases
Glutathione-S-transferase
Signaling
MAPK
ERF
bZIP
WRKY
MYB
DOF
Transcription factors
Heat shock proteins
Secondary metabolites
Putative involvment in biotic stress
B
Pathogen / pest attack
Abiotic stress
Recognition
R genes
Respiratory burst
Signaling
MAPK
Transcription factors
Defense genes
PR-proteins
Secondary metabolites
Misc. functions
HSPs
Hormone signaling
Auxins
Brassinost.
ABA
Ethylene
SA
JA
Cell wall
Beta glucanase
Proteolysis
Redox state
Peroxidases
Glutathione-S-transferase
Signaling
MAPK
ERF
bZIP
WRKY
DOF
Transcription factors
Heat shock proteins
Secondary metabolites
Putative involvment in biotic stress

Figure 5. Biotic stress response overview map. This figure shows the changes in the expression of biotic stress-responsive genes in *A. thaliana* plants during the response to the aphid and *P. syringae* treatments. Genes that have been experimentally indicated to be involved in biotic stress are collected in the main panel (coloured with dark grey), while genes and pathways that are putatively involved in biotic stress pathways are shown on the left and right sides (coloured in light grey). (**A**) Aphid infestation. (**B**) *P. syringae* infection. In both cases, the signal after infection is expressed as a ratio relative to the signal in uninfected controls, which was converted to a log2 scale and displayed. The scale is shown in the figures. Only the genes showing a significant (P = 0.01) change in expression between the treatment and the untreated control that were attributed to the respective bins by MapMan are shown. Genes whose expression was increased are indicated with increasingly intense blue and red colours. The gradation can be seen in the scale presented in the top right corner of each subfigure.
doi:10.1371/journal.pone.0058987.g005

(Terpene Synthase 10, At2g24210) were highly up-regulated during P. syringae treatment. The elicitor-activated gene *CAD-B2* (At4g37990), belonging to the phenylpropanoid metabolism category, was strongly up-regulated in a *P. syringae*-specific manner. In the alkaloid-like compound biosynthesis category, strictosidine synthase genes (At1g74010, At1g74020) were highly up-regulated in the *P. syringae* experiment. In the flavonoids category, two genes *SRG1*(Senescence-Related Gene 1; At1g17020) and 2-oxoacid-dependent oxidas (At3g50210), were also up-regulated in a *P. syringae*-specific manner. Significant differences were observed in genes coupled to glucosinolate biosynthesis and hydrolysis. Glucosinolates (*GS*) are secondary metabolites typical of the order *Brassicales* [50]. Most of the aphid-specifically expressed aliphatic *GS* genes were repressed, whereas most of the *Pseudomonas*-specifically expressed genes were positively regulated. The lists of these genes are provided in **Tables 2** and **3.** Following *P. syringae* treatment, two myrosinase-associated proteins (At1g52040, At1g54020) and a nitrile-specific protein *AtNSP5* (At5g48180) were highly up-regulated. It was reported by Kissen *et al.*, that the nitrile specifier proteins involved in glucosinolate hydrolysis in *Arabidopsis thaliana* and products generated after hydrolysis, such as isothiocyanates, play multiple roles in growth regulation and defence [51].

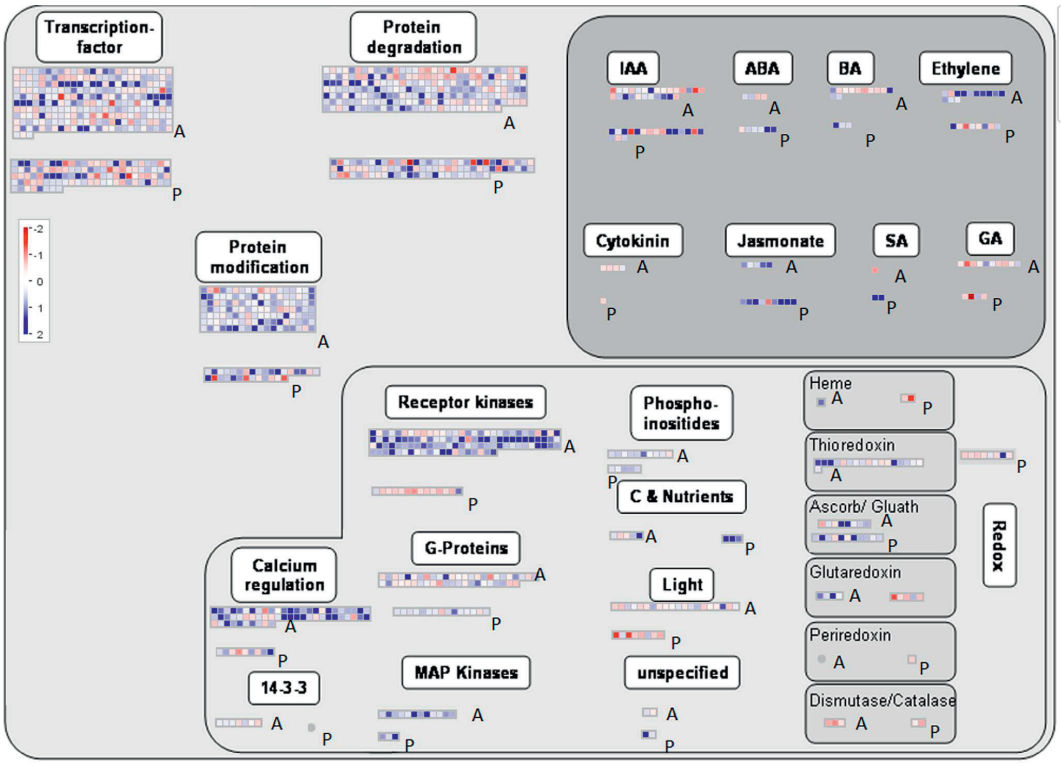


Figure 6. Regulatory overview map. MapMan regulatory overview map showing differences in transcript levels between aphid-specific and *P. syringae*-specific genes. Aphid-specific and *P. syringae*-specific bins are marked as 'A', and *P. syringae*-specific bins are marked as 'P'. In the colour scale, blue represents higher gene expression, and red represents lower gene expression. IAA, Indole-3-acetic acid; ABA, abscisic acid; BA, brassinosteroid; SA, salicylic acid; MAP, mitogen-activated protein.
doi:10.1371/journal.pone.0058987.g006

Table 2. Genes involved in glucosinolate metabolism affected by aphid infestation.

Gene ID	Log2	Description
At1g52040	4.185	*MBP1*
At1g54020	3.094	myrosinase-associated protein, putative
At5g48180	2.293	*NSP5*
At3g19710	1.231	*BCAT4* (branched-chain aminotransferase4)
At1g16400	0.72	*CYP79F2*
At1g62540	0.696	*FMO GS-OX2*
At5g25980	0.487	*TGG2, BGLU37*
At4g13430	0.448	*IIL1*
At2g44490	−0.675	*PEN2, BGLU26* (penetration 2)
At1g54010	−0.2	myrosinase-associated protein, putative
At1g62570	1.56	*FMO GS-OX4*

doi:10.1371/journal.pone.0058987.t002

Table 3. Genes involved in glucosinolate metabolism affected by *P. syringae* infection, with log2 fold-change values.

Gene ID	Log2	Description
At4g03070	−0.869	*AOP1, AOP, AOP1.1*
At3g49680	−1.529	*ATBCAT-3*
At3g58990	−0.648	aconitase C-terminal domain-containing protein
At2g43100	−1.12	aconitase C-terminal domain-containing protein
At1g80560	−0.7	3-isopropylmalate dehydrogenase
At1g31180	−1.045	3-isopropylmalate dehydrogenase
At4g13770	−0.415	*CYP83A1, REF2*
At2g31790	−0.813	UDP-glucoronosyl/UDP-glucosyl transferase family protein
At1g18590	−0.679	*SOT17, ATSOT17, ATST5C*
At1g74090	−0.901	*SOT18, ATSOT18*
At1g12140	0.313	*FMO GS*
At1g65860	−0.97	*FMO GS-OX1*
At1g62560	−0.679	*FMO GS*
At4g03060	−1.248	*AOP2* (alkenyl hydroxalkyl producing 2)
At5g57220	1.527	*CYP81F2*
At4g31500	1.352	*CYP83B1, SUR2, RNT1, RED1, ATR4*
At5g07690	−1.511	*MYB29, ATMYB29, PMG2*
At5g61420	−0.707	*MYB28, HAG1*
At2g33070	0.524	*NSP2*
At4g12030	−0.882	sodium symporter family protein

doi:10.1371/journal.pone.0058987.t003

A Large Number of Transcription Factors are Differentially Regulated, Many of which are Unique to Insect or Bacterial Stress

Transcription factors are the key regulators of gene expression changes and, thus, represent important part of a complex regulatory network allowing plants to adjust to changes in their environment [52]. Members of several Arabidopsis transcription factor families have been linked to plant stress responses, and a significant overlap in the expression profiles of many of these genes corresponding to a range of stress conditions has been reported. TFs are often induced by signalling phytohormones such as JA, SA or ET. The TFs that were differentially expressed during the aphid and *Pseudomonas* treatments are reported in **Table 1**, and their names are given in **file S12**. Additionally, pictorial representations of the aphid-specific and *P. syringae*-specific TFs produced using MapMan software are shown in **Figure 7**. There were 16 WRKY TFs that were up-regulated in an aphid-specific manner (*WRKY20, WRKY22, WRKY39, WRKY21, WRKY40, WRKY26, WRKY50, WRKY25, WRKY38, WRKY51, WRKY53, WRKY47, WRKY46, WRKY69, WRKY33, WRKY16*). *WRKY* TFs can act as both positive and negative regulators of plant defence pathways. The mechanisms activating *WRKY* TFs can involve the MAP kinase cascade and calcium signalling. It has been demonstrated that a subgroup of *WRKY* TFs can act as calcium concentration sensors, being activated by the increase in the Ca^{2+} concentration that occurs under inducer attack [53]. Mechanical penetration of cells by aphid stylets changes the plasma membrane potential and increases in intracellular $Ca2^{+}$ concentrations. Fluctuations in the cytosolic Ca^{2+} concentration resulting from the opening of membrane-bound calcium channels are further decoded by several Ca^{2+}-binding proteins, including the *WRKY* TFs. The up-regulated aphid-specific TFs also include C2H2 zinc finger proteins.

The MYB family, which is another large family of TFs characterised by a conserved MYB DNA-binding domain, bind to a variety of different DNA sequences. Among the *P. syringae*-specific TFs, there are 9 *MYBs* (*MYB95, MYB112, MYB90, MYB102, MYB32, MYB114, MYB59, MYB60, MYB20*), of which the first 7 are significantly up-regulated. *MYB90* is also known as *PAP2* (PRODUCTION OF ANTHOCYANIN PIGMENT 2), suggesting that some of these MYBs are likely to be involved in anthocyanin biosynthesis. A few members of this family specifically activate genes related to tryptophan aliphatic glucosinolate and indoyl glucosinolate synthesis [54].

Integrated Information from Available Public Domains Reveals a Pattern of Potential MicroRNA-mediated Post-transcriptional Regulation during Insect and Bacterial Attack

We constructed a genetic network of the differentially regulated gene lists using the *Gene network* tool in *VirtualPlant*. First, individual genes belonging to the common category were grouped into a "super node" based on shared functional properties, such as GO terms, KEGG pathways, Gene families and even similar annotations. The functional annotations were categorised in a hierarchical manner, where the functional terms and pathways were themselves grouped into higher, more generic categories (details are given in the Materials and Methods). During this analysis, we used post-transcriptional regulation, protein-protein interactions, and transcriptional regulation information from both experimental and predicted databases. The 'Regulated Edges' are predicted interactions based on the presence of known transcription factor cis-acting binding sites located in the 3 kbp upstream region of annotated transcripts. Interestingly, some of the key stress-regulated transcription factors are reported in publications, or have been computationally predicted to be regulated by different microRNAs. Thus, we were able to hypothesise that the activation of microRNA genes under biotic stresses would lead to the repression of many downstream protein-coding genes and affect physiological responses. This analysis indicates a new direction for conducting large-scale experiments and subsequent

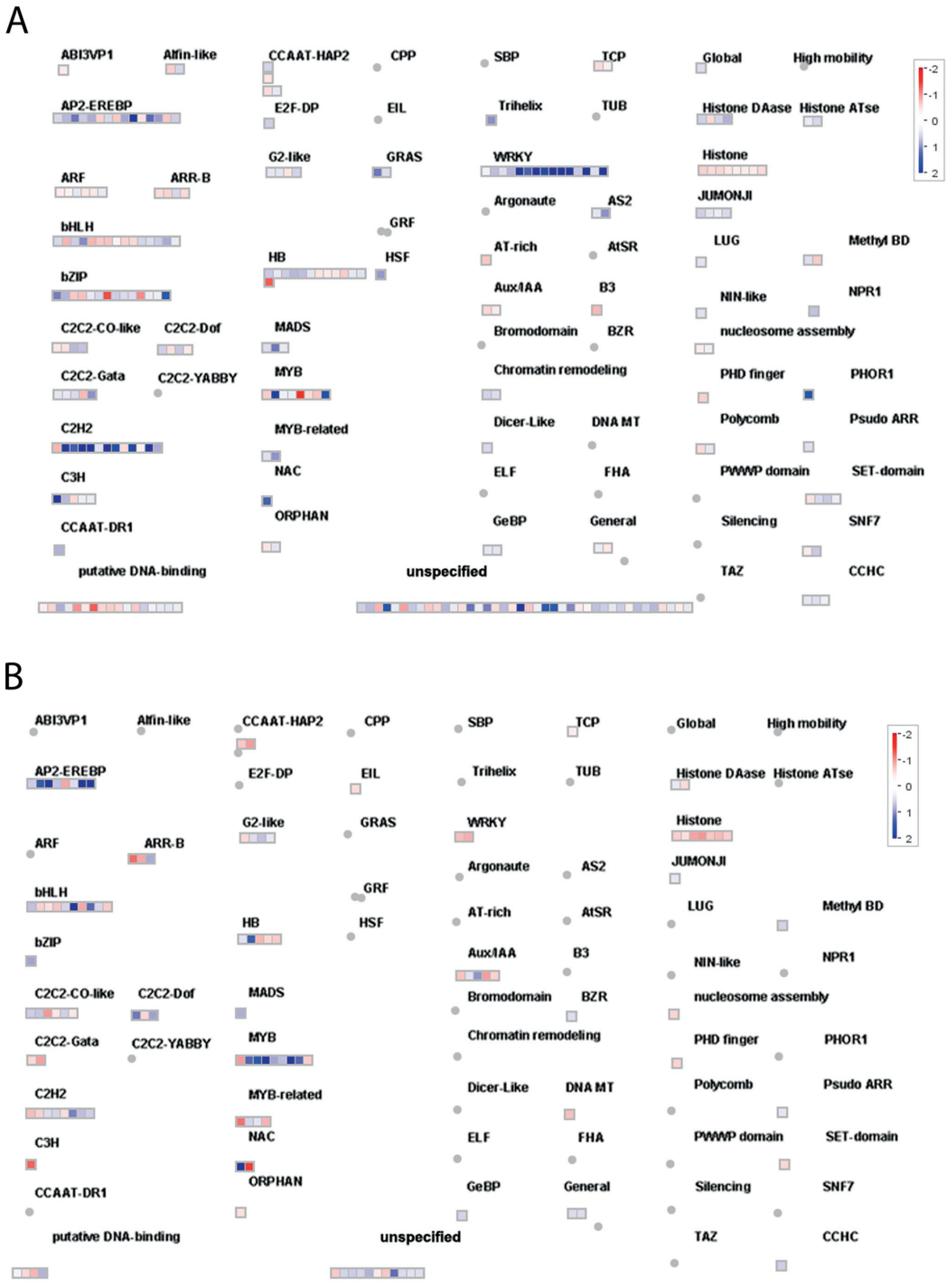


Figure 7. Transcription overview map. (A) Aphid specific; (B) *P. syringae* specific.
doi:10.1371/journal.pone.0058987.g007

bioinformatics analyses to explore the regulatory links between biotic stress and microRNAs in *A. thaliana*.

Connection of microRNAs to Genes from the Common Category

Supernode analysis of the differentially expressed common genes using the *VirtualPlant* tool revealed a supernode, or cluster, of 66 genes known to show connections with 27 microRNAs (**Figure 8A,** marked as a blue-coloured cluster**).** Further analysis on this cluster of 66 genes identified 9 genes (**Table 4**) with experimentally validated microRNA-binding sites (**Figure 8B**). Six of these genes encode known Arabidopsis transcription factors. Manually retrieved related literature references for each of this microRNA are provided in **Table 5**. We then studied all of the microRNA genes curated in the microRNA Registry database (microrna.sanger.ac.uk/sequences/). Out of the 66 genes in this cluster (**Figure 8A,** supernode annotated as 'none'), At1g20510 (*OPCL1*) and At4g05160 (putative 4-coumarate-CoA ligase/4-coumaroyl-CoA synthase) are known to be involved in jasmonic acid biosynthetic processes. Two genes, At1g50670 and At5g53160 (*SPL4*), showed the maximum number of connections to microRNAs. *RD26* (RESPONSIVE TO DESSICATION 26, At4g27410), *BZIP25* (BASIC LEUCINE ZIPPER 25, At3g54620), *JIN1* (JASMONATE INSENSITIVE 1, At1G32640) and *BGLU11* (BETA GLUCOSIDASE 11, At1g02850) contain putative microRNA binding sites, though they have not yet been verified experimentally. Out of these 13 genes, 6 are known to be TFs. Details are given in **Tables 4 and 5.**

Connection of microRNAs to Genes Showing Aphid-specific Responses

Among the 3,382 transcripts showing an aphid-specific response, the GO enrichment category 'response to stimuli (biotic and abiotic stress)' included 242 stress-responsive genes. Among these genes, 42 are known to exhibit transcription factor activity **(Figure 9A).** Additionally, out of these 242 stress-regulated genes, 21 genes have been reported to be associated with microRNAs based on literature and database searches, as described in the Materials and Methods section (**Table 6).** The reported target gene families for these microRNAs were retrieved through a manual literature search and are listed in **Table 7**. Many of the genes that were differentially regulated by aphid attack belong to these reported gene families.

Connection of microRNAs to Genes Showing *P. syringae*-specific Responses

Among the 1602 transcripts showing *P. syringae*-specific responses, the GO enrichment category 'response to stimuli (biotic and abiotic stress)' included 146 genes. Out of these 146 stress-responsive genes, 24 are known to exhibit transcription factor activity (**Figure 9B**), and 6 have been reported to be associated with microRNAs based on literature and database searches **(Table 8).** The reported target gene families for these microRNAs were retrieved through a manual literature search and are listed in **Table 9**. Many of the genes that were differentially regulated by *Pseudomonas* attack belong to these reported gene families.

Cross-validation of Differentially Regulated Aphid and *Pseudomonas*-specific Transcription Factors via Co-expression Analysis of the Multiple Biotic Stress Dataset

The differentially regulated gene sets included many signature transcription factors known for their involvement in stress responses. A co-expression analysis based on a compendium of 69 ATH1 biotic stress experiments, generated using the COR-NET tool, showed that many of these TFs have been found to be strongly co-expressed during various biotic stress experiments. From the 66-gene supernode cluster in the common group, the co-expression analysis produced a network of 26 nodes with 25 edges (**Figure 10A**). One module consisted of 9 genes *CPK6, TCH3, BZIP25, AOX1D, RD26, ERD2, MPK1, GDH2* and *HSF4* that were strongly co-expressed. The extended module contained 16 genes, 5 of which are involved in calcium-mediated signalling: *CPK6, TCH2, TCH3* and two EF-hand proteins. Functional annotation revealed that these genes are known to be involved in several different biotic and abiotic stress responsive processes. The calcium-dependent protein kinase (*CPK6*) is a positive regulator of methyl jasmonate signalling in guard cells and represents an important gene involved in methyl jasmonate signalling and signal crosstalk between methyl jasmonate and abscisic acid in guard cells [55]. *TCH3* is a calmodulin-like protein that is up-regulated in response to various environmental stimuli, including mechanical stimuli [56]. Responsive to desiccation 26 (*RD26)* encodes an *NAC* transcription factor that may be coupled to an ABA-dependent stress-signalling pathway [57], while the heat shock protein-70 cognate protein Early-responsive to dehydration (*ERD2*) which is induced by heat and dehydration is a key element in defence

Table 4. The 9 genes in the common set known to be regulated by biotic stresses and their association with stress-inducible microRNAs (Refer to **Figure 8 B**).

Gene ID	microRNA
At1g53160	*mir156f, mir156d, mir156h, mir157a, mir157b, mir156c, mir156a, mir156g, mir156e, mir157c, mir156b*
At5g50670	*mir156f, mir156d, mir156h, mir157a, mir157b, mir156c, mir156a, mir156g, mir156e, mir157c, mir156b*
At3g44860	*mir163a*
At3g44870	*mir163a*
At1g56010	*mir164b, mir164c,mir164a*
At5g43780	*mir395a, mir395b, mir395c, mir395d, mir395e, mir395f*
At2g33770	*mir399a, mir399b, mir399c, mir399e*
At1g24793	*mir859a*
At1g25054	*mir859a*

References from the literature related to each of the reported microRNA families are provided in **Table 5.**
doi:10.1371/journal.pone.0058987.t004

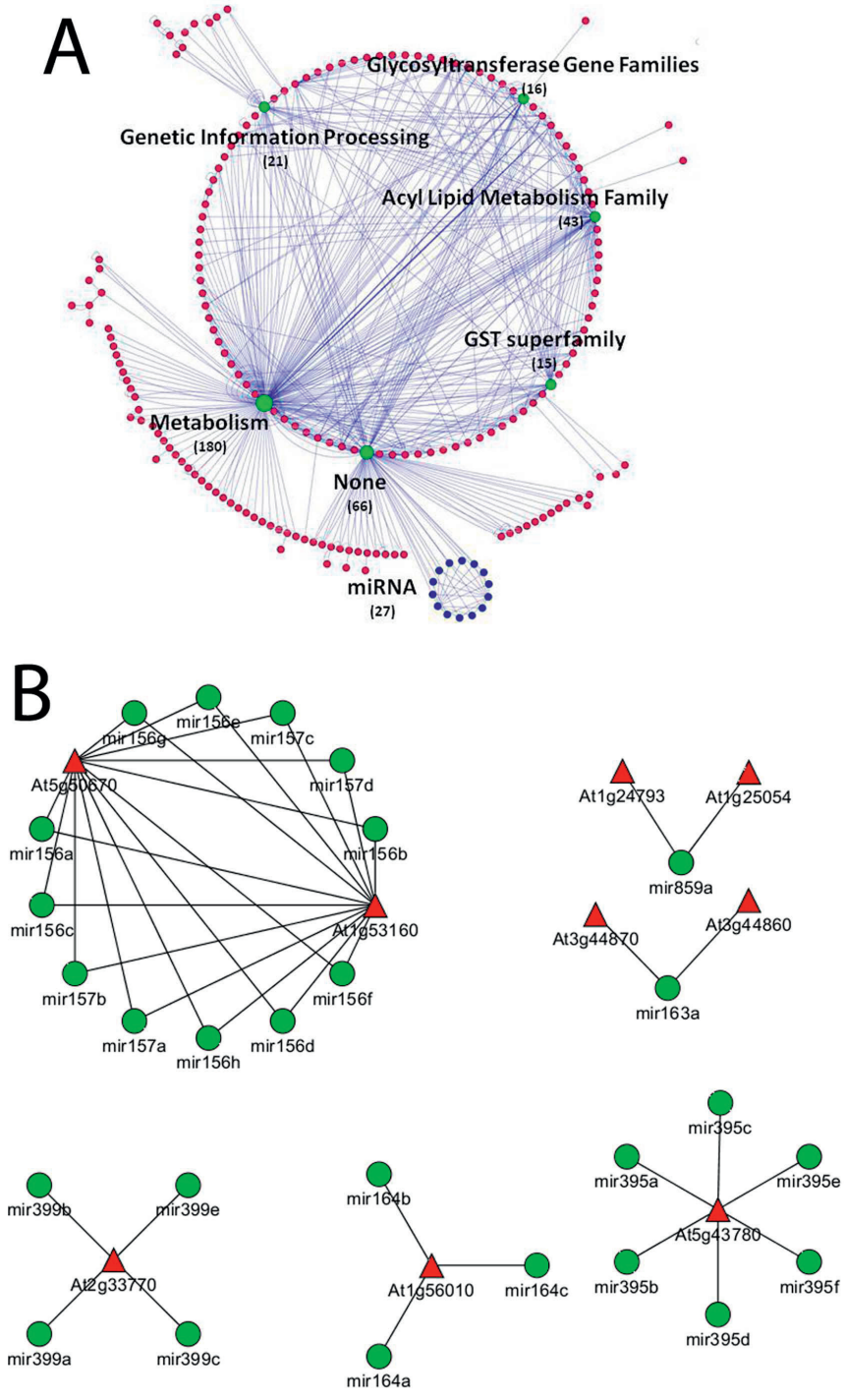

A
Glycosyltransferase Gene Families
(16)
Genetic Information Processing
(21)
Acyl Lipid Metabolism Family
(43)
GST superfamily
(15)
Metabolism
(180)
None
(66)
miRNA
(27)
B
mir156g
mir156e
mir157c
At5g60670
mir157d
mir156a
mir156b
mir156c
At1g53160
mir157b
mir156f
mir157a
mir156h
mir156d
At1g24793
At1g25054
mir859a
At3g44870
At3g44860
mir163a
mir399b
mir399e
At2g33770
mir399a
mir399c
mir164b
At1g56010
mir164c
mir164a
mir395c
mir395a
mir395e
At5g43780
mir395b
mir395f
mir395d

Figure 8. Retrieved micro-RNA connections of the common genes. A) Super node analysis using the *Gene networks* tool in *VirtualPlant*, visualised with Cytoscape 2.7.0. Individual genes in the common category were grouped into a supernode (red-coloured nodes) based on shared functional properties, such as GO terms, KEGG pathways, gene families and even similar annotations. Each supernode size corresponds to the number of genes present in that category. The edges represent connections among different functionally grouped supernodes. The top 6 most highly populated supernodes are filled with green colour. A supernode consisting of 66 genes known to show connections with 27 microRNAs (cluster of blue-coloured nodes). microRNA binding sites have been reported in existing literature for 9 of these genes, and 6 of them are known transcription factors. **B)** Details of the 9 genes mentioned above, which are known to be regulated by 27 microRNAs. MicroRNAs are shown as green-coloured circles, whereas target genes are depicted as red-coloured triangles. Edges represent the interactions between microRNAs and their target genes. Please also refer to **Table 4** and **Table 5** for detailed information and related evidence in the literature.
doi:10.1371/journal.pone.0058987.g008

response signalling pathways [58]. The MAP-kinase gene *MPK1* participates in pathogen signalling, and its kinase activity increases in response to mechanical injury [59]. Glutamate dehydrogenase 2 (*GDH2*), the alpha-subunit of glutamate dehydrogenase, is a mitochondrial protein that has been reported to be responsive to diverse environmental stresses [60]. Arabidopsis heat shock factor (*HSF4*) regulates the expression of heat shock proteins [61]. The genes in the aphid-specific and pseudomonas-specific co-expression module have been discussed in previous sections.

Conclusions

We generated and analysed data from two different biotic stress experiments conducted in *Arabidopsis thaliana* in which the plants were challenged with the aphid *Brevicoryne brassicae* and the bacterium *P. syringae syringae*. Our data showed that the transcriptional response of *Arabidopsis* to these very different attackers resulted in the differential regulation of a diverse range of biological processes. Transcriptional responses and networks unique to insect or bacterial stress conditions were identified, as were sets of genes showing similar a response under both stresses. By examining the responding genes and the functional network characteristics of each stress response, we found that a significant number of the transcripts encode transcription factors. Most of these transcription factors have shown to be involved in stress responses and regulatory processes. Some *WRKY* and *bZIP* genes were expressed differentially only during the aphid experiment, whereas some *MYB* genes were expressed differentially only during *P. syringae* infection. A Gene Ontology-based overrepresentation analysis revealed that half of the genes from the common list were involved in central metabolic and cellular processes, such as electron transport and energy pathways localised to the plastid. Secondary metabolism was strongly affected during both treatments, particularly the phenylpropanoid and glucosinolate pathways. Processes connected to the chloroplast, such as fatty acid biosynthesis, carotenoid production, chlorophyll biosynthesis, carbon fixation and others were down-regulated following *P. syringae* treatment. Starch biosynthesis genes were generally down-regulated, and an indication was found that the plants were degrading starch, which could help the plants to maintain the osmotic balance. Components of the ethylene, JA, SA, ABA, auxin and brassinosteroid pathways appeared to be specifically regulated during the aphid and *P. syringae* treatments. Ethylene responses were clearly induced during aphid feeding, while JA was more strongly induced by *Pseudomonas*. The number of signalling proteins that were differentially expressed during the aphid experiment was more than four times higher compared to the *P. syringae* treatment. By integrating secondary information from most available public sources, we further explored the regulatory links between biotic stress and microRNAs associated with aphid-and *P. syringae* - specific differentially regulated processes in *A. thaliana*, and the corresponding genes are briefly summarised in **Table 10.**

This study therefore demonstrates that the integration of heterogeneous publicly available information from multiple databases with experimental results can help plant biologists develop a better understanding of stress-associated processes in plants. Due to logistics and costs we examined only a single time point during the *A. thaliana* (Col-0) - *P. syringae* treatment. We are fully aware that comparing single time point restricts some analyses and is a potential limiting factor as demonstrated by Bricchi et al (2012) [62]. Although several datasets reporting temporal responses of *A.thaliana* to *P. syringae* infection were available from previously published independent studies [33,63,64,65], we decided not to combine them in the current analysis while making comparisons with our own *B. brassicae* data [66] to maintain the homogeneity of the comparisons. The analysis presented here will therefore not explain the comparative temporal dynamics of *A. thaliana* – *B. brassicae* and *A. thaliana* – *P. syringae* interactions.

Table 5. Functional targets of the microRNA families in the common set of genes (retrieved from literature searches).

microRNA	Target family
mir156 [86]	*SPL* family members, including *SPL3, SPL4*, and *SPL5*. By regulating the expression of *SPL3* (and probably also *SPL4* and *SPL5)*, this microRNA regulates vegetative phase change.
mir157 [87] [88]	*SPL* family members, including *SPL3, SPL4*, and *SPL5*.
mir163 [89]	*SAMT* family members. *mir163*, is highly expressed in *A. thaliana* diploids but down regulated in *A. thaliana* autotetraploids and repressed in *A. arenosa* and *A. suecica*.
mir164 [90]	*NAC* domains including *NAC1* and *ORE1*. Over expression leads to decreased *NAC1* mRNA and reduced lateral roots. Loss of function mutants have increased *NAC1* and increased number of lateral roots. Also targets *CUC2* and modulates the extent of leaf margin serration. Also targets *ORE1* to negatively regulate the timing of leaf senescence.
mir395 [91]	*APS* and *AST* family members.
mir399 [87,88]	*PHO2*, an *E2-UBC* that negatively affects shoot phosphate content.
mir859 [92]	*F-box* family members.

doi:10.1371/journal.pone.0058987.t005

Table 6. The 21 genes in the aphid-specific gene set known to be regulated by biotic stresses and their association with stress-inducible microRNAs.

Gene ID	microRNA
At3g15270	*mir156b, mir156f, mir156g, mir156a, mir156e, mir156d, mir156c, mir157d*
At2g33810	*mir156d, mir156c, mir156b, mir156f, mir156g, mir156a*
At5g43270	*mir157c, mir156h, mir157b, mir157a, mir157d, mir156c, mir156d, mir156b, mir156f, mir156e, mir156a, mir156g*
At5g50570	*mir157c, mir156h, mir157a, mir157d, mir156c, mir156d, mir156b, mir156f, mir156e, mir156a, mir156g*
At5g06100	*mir159b, mir159c*
At4g30080	*mir160a, mir160b, mir160c*
At1g66700	*mir163a*
At1g52150	*mir166a, mir166b, mir166c, mir165a, mir166d, mir166f, mir166e, mir166g, mir165b*
At5g37020	*mir167a, mir167b, mir167c, mir167d*
At1g72830	*mir169a, mir166b, mir16c, mir169m, mir169h, mir169l, mir169j, mir169k, mir169n, mir169i*
At5g67180	*mir172a, mir172b, mir172c, mir172d, mir172e*
At3g15030	*mir319a, mir319b, mir319c*
At4g18390	*mir319a, mir319b, mir319c*
At3g22890	*mir395a, mir395b, mir395c, mir395d, mir395e, mir395f*
At5g53660	*mir396a, mir396b*
At5g60020	*mir397a, mir397b*
At1g31280	*mir403a*
At1g12210	*mir472a*
At5g63020	*mir472a*
At1g53290	*mir775, 775a*
At5g42460	*mir859a*

Data retrieved from searches of the published literature and databases. (Refer to **Figure 9A**). References from the literature related to each of the reported microRNA families are provided in **Table 7.**
doi:10.1371/journal.pone.0058987.t006

Table 7. Functional targets of the microRNA families in the aphid-specific set of genes (retrieved from the existing literature).

Micro-RNA	Target Gene family
mir156 [93]	*SPL2, SPL3, SPL4, SPL10*
mir157 [93]	*SPL* family members, including *SPL3,4,* and *5*
mir159 [94,95]	*MYB* 107, *MYB* 116, *MYB33, MYB65, TCP2, TCP3, TCP4, TCP10, TCP24*
mir160 [93]	*ARF* family members (*ARF10, ARF16, ARF17*)
mir163 [96]	*SAMT* family members. *miR163*, is highly expressed in A. thaliana diploids but down-regulated in *A. thaliana* autotetraploids and repressed in *A. arenosa* and *A. suecica.*
mir165 [97]	*HD-ZIPIII* family members including *PHV, PHB, REV, ATHB-8, and ATHB-15*
mir166 [97]	*HD-ZIPIII* family members including *PHV, PHB, REV, ATHB-8, and ATHB-15*
mir167 [93]	*ARF* family members *ARF6* and *ARF8.*
mir169 [93]	*HAP2* family members
mir172 [94]	several genes containing *AP2* domains
mir319 [97,98]	*TCP* family members.
mir395 [97]	*APS* and *AST* family members.
mir397 [97,99]	targets several Laccase family members
mir403 [99]	*AGO2* and *AGO3*
mir472 [100]	Several *CC-NBS-LRR* family members.
mir859 [92]	Several *F-box* family members.

Most, but not all were affected by the aphid treatment.
doi:10.1371/journal.pone.0058987.t007

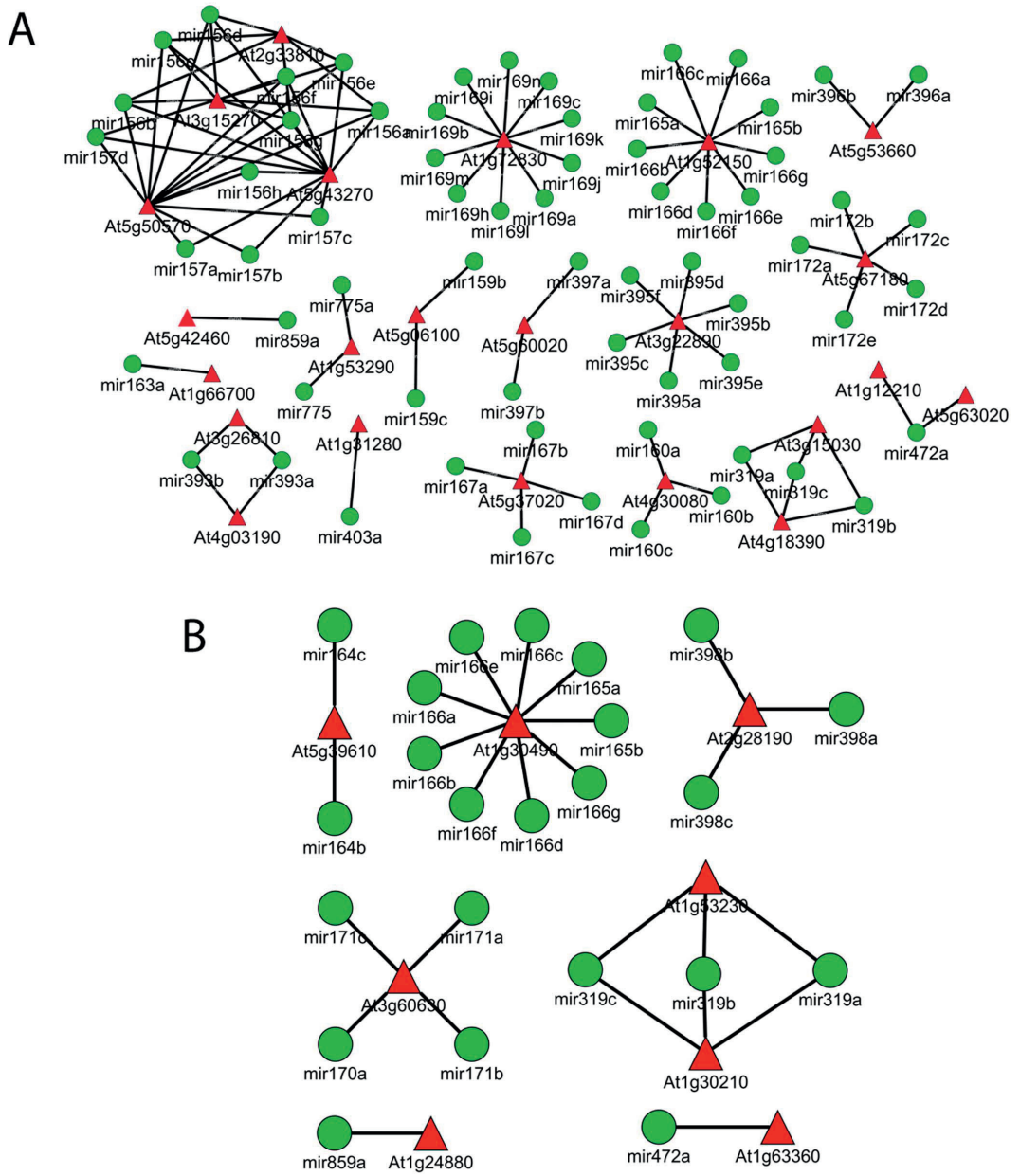


Figure 9. Retrieved micro-RNA connections of aphid-specific and *P. syringae*-specific genes. A red triangle represents a target gene, and a green circle represents a microRNA. A) Among the transcripts showing aphid-specific responses, 42 genes are known to contain microRNA binding sites. Please also refer to Table 6 and Table 7 for detailed information and related evidence from the literature. B) Among the transcripts showing *P. syringae*-specific responses, 9 genes are known to contain validated microRNA binding sites. We were able to find related references in the literature for the reported 23 microRNAs. Please also refer to Table 8 and Table 9 for detailed information and related evidence from the literature.
doi:10.1371/journal.pone.0058987.g009

Table 8. The 6 genes in the *Pseudomonas*-specific gene set known to be regulated by biotic stresses and their association with stress-inducible microRNAs.

Gene ID	microRNA
At1g30490	*mir165a, mir165b, mir166a, mir166b, mir166c, mir166d, mir166e, mir166f, mir166g*
At1g30210	*mir319a, mir319b. mir319c*
At1g53230	*mir319a, mir319b. Mir319c*
At2g28190	*mir398a, mir398b, mir398c*
At1g63360	*mir472a*
At1g24880	*mir859a*

Data retrieved from searches of the published literature and databases. (Refer to **Figure 9B**). References from the literature related to each of the reported microRNA families are provided in **Table 9**.
doi:10.1371/journal.pone.0058987.t008

Materials and Methods

To overcome the problem of the incompatibility of independent microarray experiments, a genome-wide expression analysis involving 2 different biotic stresses was conducted, in which *Arabidopsis thaliana* plants were infested with aphids (*Brevicoryne brassicae*) [45] or infected with *P. syringae* bacteria (4 biological replicates, and an untreated control were used for each comparison). The microarray data from the aphid experiment was part of a larger plant-insect study [45]. The Pseudomonas data were generated for the present study using the same technology platform to reduce experimental variation. All data have been deposited in GEO (**GSE39245** and **GSE39246**). A systems biology approach was followed to understand common and specific responses in terms of different pathways and processes in *Arabidopsis* during insect and bacterial attack. A simplified flow chart diagram of the applied methodology is provided in **Figure 11**.

Plant Material and Cultivation

The *Arabidopsis thaliana* Columbia-0 ecotype (Col-0) plants used in the experiment were derived from seeds produced by Lehle Seeds (Round Rock, USA; Catalog No. WT-2-8, Seed Lot No. GH195-1). The seeds were sterilised according to standard procedures and grown on agar medium containing an MS basal salt mixture (Sigma), 3% (v/w) sucrose, and 0.7% (v/w) agar (pH 5.7) to assure uniform germination. After 15 days, the seedlings were transferred to 6 cm diameter pots (3 seedlings per pot) filled with a sterile soil mix (1.0 part soil and 0.5 parts horticultural perlite). The plants were kept in Vötsch VB 1514 growth chambers (Vötsch Industrietechnik GmbH, Germany) under the following conditions: 8 h/16 h (light/dark) photoperiod, $22^{\circ}C/18^{\circ}C$, 40%/70% relative humidity, and 70/0 µmol $m^{-2}s^{-1}$ light intensity. A short day length was applied to prevent the plants from bolting.

Infestation Experiments

At 32 days of age (17 days after being transferred to soil), the plants had 8 fully developed leaves. Each plant was infested with 32 wingless aphids (4 per leaf), which were transferred to the leaves with a fine paintbrush. Infested plants and aphid-free controls were maintained in Plexi-glass cylinders, as described previously [66]. The plants were harvested 72 h after infestation between the 6th and 8th hours of the light photoperiod. Four biological replicates were produced from the control and infested plants, with each being sampled from 15 individual plants. Whole rosettes were cut at the hypocotyl, and aphids were removed by washing with Milli-Q-filtered water. The harvested material was immediately frozen in liquid nitrogen.

A *P. syringae* culture was grown overnight in 10 ml of Kings B solution (King *et al.*, 1954) supplemented with the antibiotics rifampicin (50 μg ml^{-1}) and kanamycin (25 μg ml^{-1}). The overnight culture was washed once in 10 mM $MgCl_2$, and the final cell densities were adjusted to an OD of approximately 0.20 at 600 nm (approximately 1.5×10^8 cfu ml^{-1}) in 10 mM $MgCl_2$. Plants were grown as described in the Plant material and cultivation section. Then, 30-dayold plants were mock-challenged with 10 mM $MgCl_2$ or inoculated with the DC3000 strain of *P. syringae* by infiltrating 3–4 leaves on the abaxial surface with a needleless 1 ml syringe. Four biological replicates of infested leaves and leaves obtained from control plants grown under identical conditions were harvested after 3 days (between the 6th and 8th hours of the light photoperiod). The leaf material was immediately

Table 9. Functional targets of the microRNA families in the *P. syringae*-specific set of genes (retrieved from the existing literature).

microRNA	Target Gene family
mir165 [51]	*HD-ZIPIII* family members including *PHV, PHB, REV, ATHB-8,* and *ATHB-15*
mir166 [97]	*HD-ZIPIII* family members including *PHV, PHB, REV, ATHB-8,* and *ATHB-15*
mir319 [51,52]	*TCP* family members.
mir398 [97,101]	*CSD* and CytC oxidase family members.
mir472 [53]	Several *CC-NBS-LRR* family members.
mir859 [46]	Several *F-box* family members.

Most, but not all were affected by the *Pseudomonas* treatment. Corresponding AtIDs are provided in **Table S6.**
doi:10.1371/journal.pone.0058987.t009

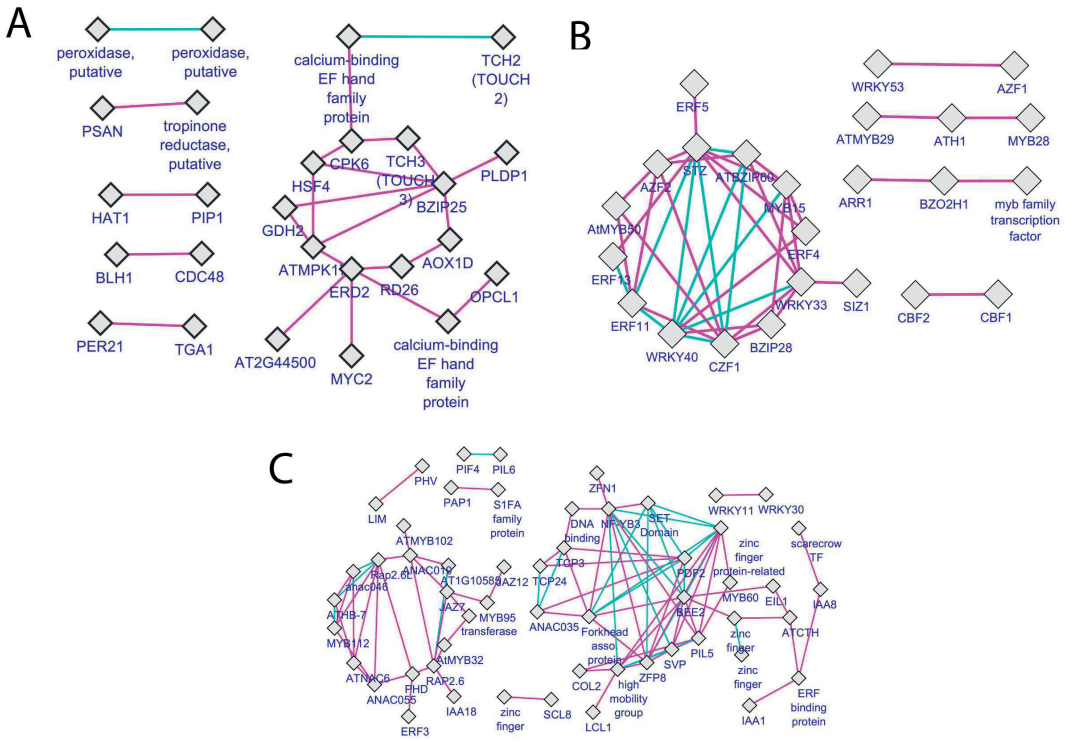


Figure 10. Co-expression network. Co-expression networks generated by CORNET using AtGenExpress biotic stress compendia based on a Pearson's correlation coefficient threshold ≥0.7. The networks were visualised using Cytoscape2.7.0. Pink-coloured edges represent a strong correlation of ≥0.9, and cyan-coloured edges represent a correlation of ≥0.7 to 0.9. **A)** Co-expression network analysis among the 66-supernode cluster in the common group resulted in a network of 26 nodes 25 edges. **B)** Co-expression network analysis of the aphid-specific TFs resulted in a network of 24 tightly co-expressed TF modules. **C)** Co-expression network analysis among 104 Pseudomonas-specific TFs resulted in a tightly co-expressed modular network consisting of 55 nodes and 94 edges.
doi:10.1371/journal.pone.0058987.g010

frozen in liquid nitrogen. Leaves from 15 plants were included in each replicate.

RNA Isolation, cDNA Synthesis, Labelling and Hybridisation

Total RNA was isolated from cauline leaf tissue from plants from each experiment. Each experiment consisted of four infested samples and four control samples. Total RNA was extracted from 100 mg of cauline leaf material using the RNeasy Plant Minikit (Qiagen, Hilden, Germany) and eluted in 2×50 µl of RNAse-free water. Any residual DNA in the RNA samples was removed by on-column treatment with RNAse-free DNase. The eluted RNA was concentrated to 10–20 µl using a 30 kDa cut-off Microcon spin filter unit (Amicon, Bedford, USA). To protect the RNA from degradation, the RNasin Plus RNase inhibitor (Promega, Madison, USA) was added to a final concentration of 1 unit µl-1. The purity and quantity of the obtained RNA was determined using a Nanodrop ND 1000 instrument (Nanodrop Technologies, Wilmington, DE, USA). RNA integrity was analysed via formaldehyde agarose gel electrophoresis. First-strand cDNA was generated from total RNA (15 µg) using the Superscript III Reverse Transcriptase (Invitrogen, Carlsbad, USA) and oligo dT primers with a 3DNA capture sequence from the 3DNA Array 350TM kit (Genisphere, Hatfield, USA). RNA samples were labelled with either the Cy5-capture primer or Cy3-capture primers (sample dye-swapping). The cDNAs were hybridised to the microarray slides at 58°C using a Slide Booster Hybridisation Station (Advalytix, Brunnthal, Germany) together with Cy3- and Cy5-labelled dendrimers from Genisphere. The slides were washed according to the manufacturers' instructions (Genisphere and Advalytix).

Microarrays

The microarray slides contained 31811 unique 70-mer oligos with a C6-amino linker, corresponding to a total of 33696 spots, covering 26624 genes. Of these oligos, 29110 were from the Qiagen-Operon *Arabidopsis* Genome Array Ready Oligo Set (AROS), Version 3.0, while the others were custom made and produced by Operon (Alameda, CA, USA) or MWG (Ebersberg, Germany). The sequences of all of the custom-made probes on the chip have been deposited in GEO and are available under accession **GPL15699**. The oligonucleotides were dissolved in MQ grade water and 50% DMSO (20 pmol/µl) and spotted on aminosilane-coated UltraGaps slides (Corning, NY, USA) using a BioRobotics MicroGrid II robot (Genomic Solutions, MI, USA).

Table 10. Summary of aphid-specific and *P. syringae*-specific genes associated with differentially regulated processes during both of the treatments.

Categories	Aphid specific	*Pseudomonas* specific
Biotic stress signaling processes (up)	*FRK1, ATMPK11, ATRABA1e, PBP1, CRK11, EDA39, CRK6,CRT3, RLK5, LECRK1, ACA2, WAKL2, GLR2.7, ATMKK9, CML39, AP4.3A, CPK10, AtRABH1c, B120, MKK4, CPK29, ACA11, XLG2, CPK32, WAK2, RPK1, CPK7, ATSERK5, AGG1, NMAPKK, CPN1, RLK, CPK5, EP1, WAK1, ATHRGP1, WAKL22, MEKK1, CPK4, PHYD, CAM2, MKK2, AGG2, MPK4, ARK2, ELK4, FRS2,ATVPS34, RAN1, AtRABH1a, ARK3, RHA1, FRS5, CPK1, CAM3, CPK3, EFR, GRF1, MKK5, MSS3, MPK1, SOS3, ATGB1, AHP5, CRT1, CAM9, ATRABA1D, ATMPK15, MEE62, ATGDI1, PAT1, PIP5K9, ATPERK1, GLR1, APKKK5, WAKL6, GRF6, ATGDI2, BON2, GRF5, JAB1, GRF10, RAB6A, ARG, SIRANBP, ATG5, TIC, RAN3*	*AHP1, ATCP1, ATPH1, ATRABA1A, AtRABA1g, AtRABA5d, ATRABC2B, ATRABD2B, AtRABE1a, ATSARA1A, CCL, DRP1A, GLR1.2, GLR1.3, IQD14, MAPKKK18, MAPKKK3, MPK7 MSL4, PHYB, PLC1, RABF1, RD20, SAC9, SMG1*
Biotic stress signaling processes (down)	*AtRABA1f, AtRABA2d, AtRABA5b, ATRABE1C, ATRABE1D, AtRABG3d, CAM7, CPK8, ECT1, FRS12, GLR3.6, GRF2, GRF4, HSL1, iqd21, IQD31, LRR1, LSH1, MSL6 NIK1, NPGR1, NPGR2,NPY1, PAP2, PHYC, PHYE, RALFL22, RALFL23, RCI1, ROPGEF1, RPT1, SCABP8, SnRK1.2, SPA1, TOC33, VAN3, VAR3*	*ACA4, ATCAMBP25, ATRABC2A, BAM1, iqd2, IQD3, NIK3, PKS1, QRP1, RABG3B, RALFL32, SRL2, TMK1*
HSPs (up)	*ATJ1, ATJ2, ATJ3, BIP1, BIP2, BIP3, HSF A4A, HSP70, HSP70-1, HSP81-2, HSP81-3, HSP83, HSP91, J8, KAM2, MTHSC70-2, SHD*	–
HSPs(down)	*ARL1*	–
Proteolyitc enzymes (up)	*ATAPG9, AtATG18d, AtATG18f,ATL2, ATL6, ATL8, AtMC2, AtMC3, AtMC4, AtPNG1, AtPP2-B10, ATTLP9, BCS1, bt5, DA1, EBF2, FUS9, MCP1B, mos5, NHL8, PAC1, PAE2, PAF1, PBC2, PBG1, RHA1A, RHC1A, RHF2A, RKP, RMA1, scpl46, SKIP4, SUMO3, UBC15, UBC18, UBC23, UBC25, UBC33, UBC35, UBC9, UBP22, UBP3, UBP4, UBP5, UBP9, UBQ11, UEV1B, UPL3, UPL6, XERICO*	*AIR3, ATAPM1, ATG5, ATG8A, ATG8F, ATG8I, ATGGH1, ATGGH2, ATGGH3, AtTLP7, BPM2, HSP93-V, NSF, PAA1, PBB2, PBE1, PUX3, RGLG2, RHA2A, RIN2, ROC1, RPN10, RPT3, RPT5B, SAG12, scpl49, SKP2A, SKP2B, UBC2, UBC28, UBQ3, UBQ9, UCH3, XBCP3*
Proteolyitc enzymes (down)	*ATL3, ATL5, ATRBL2, EGY1, EMB2083, emb2458, FKF1, ftsh9, GRH1, MUB5, PIP, SBT1.3, scpl10, scpl2, scpl20, scpl25, scpl42, SKP1B, SLP2, SLP3, SLY2, SNG1, UBC20, UBC29, UBC7, UBP24, V*	*DEGP8, FTSH1, FTSH11, nClpP6, PT, RUB1, UBC8*
Secondary metabolic (up)	*CYP73A5, CYP81F2, FAH1, pal1, PGGT-I, SUR2, UGT72E1*	*4CL5, ALDH10A8, ALDH10A9,ATCPISCA,BCAT4, CYP79F2, DXPS1, ELI3-2, LAS1, MBP1, NIC2, SIAA1, SRG1, SS2, TGG2, TPS04, TPS10, TT3, VTE2*
Secondary metabolic (down)	*ABC4, AOP1.1, AOP2, BCAT3, CYP706A5, FPS2, GGPS1, IPP2, ISPH, KCS5, LAC11, LAC17, LUP1, MVA1, PMG1, PMG2, REF2, TT4, YRE,*	*CAC3, CAD4, DXS, FLS, HCT, KCS10, LUT2, PAL3, PDE277, PEN2, POP1, PSY, SPS2, TT5, VTE3*
Cell wall (up)	*AGP5, ATHRGP1, ATPME3, BXL1, CSLE1, EXP16, FUT4, FUT7, GER1, GER2, MUR_1, UXS4, XTH22, XTR4*	*ATAGP1, ATAGP10, AtAGP24, CSLA01, CSLG1, DIN9, ISA1, KING1, MEE31, PGAZAT, PGIP2, PMEPCRA, RGP1, UGE3*
Cell wall (down)	*AGP7, AGP9, ATAGP12, ATAGP18, ATAGP19, AtAGP21, ATAGP22, ATAGP26, ATAGP4, ATFUC1, ATFXG1, AtGH9B5, AtGH9B8, AtkdsA1, COB, EXPB1, EXPL2, EXPR, EXT, FLA10, FLA11, FLA12, FLA17, FLA18, FLA9, FLR1, LEW2, PMR6, QUA1, UER1, UGE2, UXS3, XTH9*	*ATAGP16, ATAGP25, BGAL2, CSLA03, CSLA7, CSLB03, EXP1, EXP15, LGT1, PRP4, ROL1, SOS5*
TFs (up)	*HSS, HYH, KNAT4, KNAT6S, LBD37, LBD39, LD, LUH, MBD4, MYB15, MYB33, NIMIN-2, NIMIN-3, ORA47, RAP2.4, RAV1, SAI1, SDG15, SIZ1, SNF7.1, SPL, TGA3, TGA5, TOC1, WRKY20, WRKY21, WRKY22, WRKY25, WRKY26, WRKY33, WRKY38, WRKY39, WRKY40, WRKY46, WRKY47, WRKY50, WRKY51, WRKY53, WRKY69, ZAT10, ZAT6, ZAT7, ZCW32, ZFAR1*	*RAP2.6,AGD5, ARR2,ATHB7, AtIDD11, ATMYB102, AtMYB32, ATNAC3, ATRBP45C, CDC5, GL19, IAA18, LCL1, LHW, LOL2, LZF1, MBD11, MYB112, MYB114, MYB59, MYB95, PAP1, PAP2, PHV, PMZ, PRR2, PUR, RAP2.12, RAP2.3, Rap2.6L, RPD3A, TOM1, ZFP7, ZFP8*
TFs(down)	*ARF11, ARF22, ARF8, ARR12, ATH1, ATRR3, BLH6, BZO2H2, CIA2, ETT, GBF5, HDT2, HMGB6, IAA16, ICU4, MBD10, MEE47, MS1, MYB124, NGA2, OBP4, PCNA2, pde191, PMG1, PMG2, PTAC1 RAP2.2, RR16, SAW2, SDG26, SHY1, STH, TCP4, TINY2, UNE10, VPS46.1, VRN2, WHY3, WLIM1, WOX4, FHD2*	*TCP24, anac061, ARR7, ATCTH, ATHB5, BEE2, COL3, EIL1, HAT2, hda14, IAA1, IAA8, IBC6, METI, MFP1, MSG2, MYB20, MYB60, PDF2, PIL5, PIL6, PTAC4, TRY, WRKY11, WRKY30, ZFN1, ZFP4*
Ethylene (up)	*ERF11, ERF5, ERF2, ERF-6-6, ERF13, ORA59, RAP2.5, ERS1, ERF7,atpdx1.2,*	*MBF1B, ERF3*
Ethylene (down)	*2-oxoglutarate-dependent dioxygenase*	*ACS10*
ABA (up)	*STO1, AAO3*	*SIR3, FIP1*
ABA (down)	*ATHVA22C, HVA22D*	*HVA22H*
JA (UP)	*LOX5*	*LOX1, CYP74A, AOC2, AOC4, JMT, JR1*
IAA(UP)	*WIN3, AIR12, AXR1, TIR5, ILL1, ARG1*	*TGG2, GH3-10, WES1, YDK1, ILR1, GH3.6, ILL5*
IAA(Down)	*AFB2, COV1, MES17*	*SAUR_AC1*
SA (UP)	–	BSMT1, UDP-glucoronosyl
SA(Down)	methyltransferase	

*Only those genes with an alias (short annotation name present in TAIR) have been included in this summary table. A complete list of aphid-specific and *P. syringae*-specific genes and their corresponding At IDs have been provided in **Table S6.**

doi:10.1371/journal.pone.0058987.t010

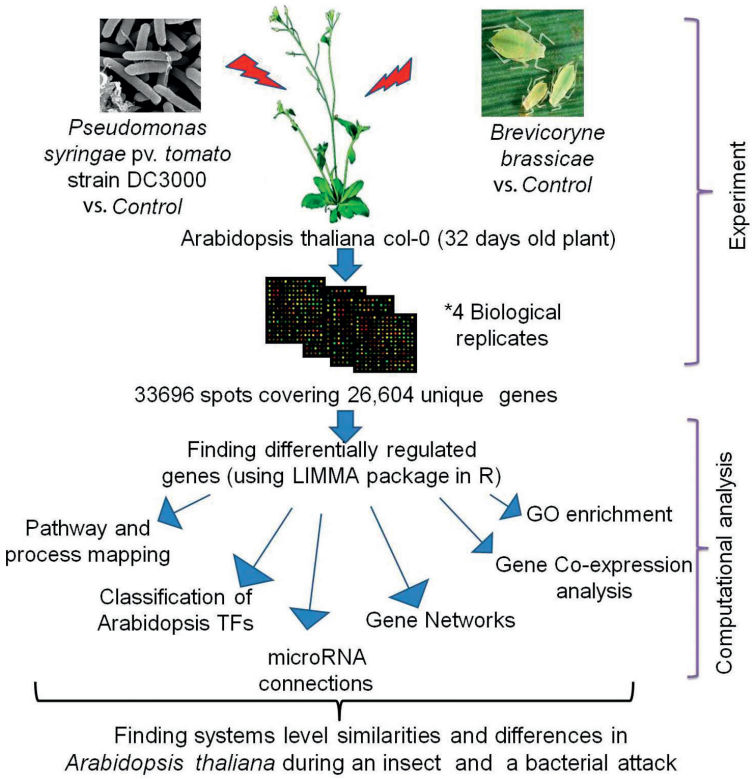


Figure 11. Flow chart of the methodology.
doi:10.1371/journal.pone.0058987.g011

Printing of the microarray slides was performed at the Norwegian Microarray Consortium (Trondheim, Norway). Hybridisations were conducted using a Slide Booster Hybridization Station (Advalytix, Brunnthal, Germany), and the slides were washed according to the manufacturer's instructions (Genisphere and Advalytix). The slides were scanned at a 10 µm resolution on a G2505B Agilent DNA microarray scanner (Agilent Technologies). The resulting images were processed using GenePix 5.1 software (Axon Instruments, Union City, USA).

Statistical Analysis of the Microarray Data

Each dataset obtained from the aphid and *Pseudomonas* treatments corresponded to 4 microarray slides, where the controls and treated samples were alternately labelled with Cy5 and Cy3. The GenePix-processed data were filtered to remove spots that had been flagged as 'Absent', 'Not Found' or 'Bad', or exhibited median foreground intensity below the local median background intensity. The R statistical program (version 2.10.1) was used for all statistical analyses [67]. No background subtraction was performed. The data from each array were log-transformed and normalised using the printtip-loess approach (Yang et al. 2001). Within-array replicated measurements for the same gene were merged by taking the average over the replicates. The data were then scaled so that all array datasets presented the same median absolute deviation. Genes showing dye-biased responses due to Cy5 and Cy3 labelling were identified and excluded. During data processing, we focused on genes for which at least 3 out of 4 biological replicates for the examined time points passed the quality control criteria suggested by Jørstad et al. [68,69]. To make statistical inferences about differentially regulated genes, the Limma package [70] was used. The Limma approach is based on fitting a linear model to the expression data from each probe on a microarray. Genes showing an adjusted p-value of less than 0.05 were considered to be significantly differentially expressed. All of the genes discussed in this paper were found to be significantly differentially expressed in one of the two treatments (aphid or *Pseudomonas*).

GO Enrichment Analysis of Common Genes

We employed a simple set theory-based operation in R to find common and specific transcriptional responses that occurred in both experiments. To conduct automated GO [71], TAIR [72] annotations, we simultaneously used three programs: ClueGO [73], BiNGO [74] and *VirtualPlant* [75]. Only the ClueGO results were included in this manuscript. Transcription factors were classified according to the 'The Database of Arabidopsis Transcription Factors' [76]. In ClueGO, to calculate enrichment values for terms and groups, we used two-sided (enrichment/depletion) tests based on the hypergeometric distribution to calculate doubling for two-sided tests to address discreetness and conserva-

tism effects, as suggested by Rivals et al. [77]. To correct the P-values for multiple testing, the Bonferroni method was used to control the type I error (false positive) rate [78]. ClueGO employs a new kappa statistic. To link the terms in the network, ClueGO first creates a binary gene-term matrix with the selected terms and their associated genes. Based on this matrix, a term–term similarity matrix is calculated using chance-corrected kappa statistics to determine the strength of the associations between the terms. Because the term–term matrix is of categorical origin, using a kappa statistic was found to be the most suitable method. Finally, the created network represents the terms as nodes, which are linked based on a predefined kappa score level. The kappa score threshold can initially be adjusted on a positive scale from 0 to 1 to restrict the network connectivity in a customised way. In our analysis, we used a kappaScore threshold of 0.3. The size of the nodes reflects the enrichment significance of the terms. The functional groups represented by their most significant (leading) term are visualised in the network, providing an insightful view of their interrelationships. Furthermore, other ways of selecting the group-leading term, e.g., based on the number or percentage of genes per term, are also provided.

VirtualPlant [75] integrates genome-wide data regarding the known and predicted relationships among genes, proteins, and molecules as well as genome-scale experimental measurements. This warehouse includes descriptions of molecular entities (e.g., gene annotations and functional classifications), molecular interactions (metabolic associations, regulatory interactions, and other interaction data from public databases), and publicly available microarray data (including more than 1,800 gene chip hybridisations from the ATH1 Affymetrix platform obtained from the European Arabidopsis Stock Center [NASC] using the Affywatch subscription service). VirtualPlant also provides visualisation techniques that render multivariate information in visual formats that facilitate the extraction of biological concepts.

Co-expression Analysis of Common Genes using CORNET

The construction of co-expression networks for multiple input genes was conducted using the CORNET tool [31]. The co-expression tool calculates the correlation between gene expression profiles using one or more precompiled expression datasets and, as such, identifies possible functional associations between genes. Out of all of the available expression data, we selected the subgroup consisting of 69 ATH1 AtGenExpress biotic stress compendium expression data. All the expression data were processed using RMA from the R BioConductor package and making use of the CDF described in Casneuf et al. [79]. Pearson's correlation coefficients were calculated between the given genes. Correlation coefficients higher and lower than a certain value are reported. Pearson's correlation coefficient (PCC) was used at a cut off ≥ 0.7. Networks and associated evidence were visualised in Cytoscape 2.7.0.

Gene Networks, microRNAs and Connections to Post-transcriptional Gene Regulation

The *Gene networks* tool in *VirtualPlant* groups individual genes into a supernode based on shared functional properties, such as GO terms, KEGG pathways, gene families and even similar annotations. Edges were drawn between two supernodes when at least one gene or gene product in each supernode showed a molecular interaction. To improve the regulatory interaction predictions, we filtered the transcription factor:target gene predictions to include only the transcription factor and target pairs whose expression values were correlated in the microarray experiment [80]. The selected statistic for the calculation of correlations in this analysis was the Pearson's correlation coefficient, with a cut-off value of less than or equal to 0.7. The results were then cross-compared with all of the microRNA genes curated in the microRNA Registry (microrna.sanger.ac.uk/sequences) [81] and in the Arabidopsis Small RNA Project (ASRP) Database [82]. In certain cases, we also compared the results with microRNAs and precursor candidates predicted for the *A. thaliana* genome by the algorithm *findMicroRNA* [83]. We followed specific criteria required for the annotation of plant microRNAs, including experimental and computational data as well as refinements of standard nomenclature, as described in [84] [85].

Supporting Information

File S1 GO-annotation cytoscape network file (.cys) for common genes.
(CYS)

File S2 GO-annotation cytoscape network file (.cys) for aphid-specific genes.
(CYS)

File S3 GO-annotation cytoscape network file (.cys) for *P. syringae*-specific genes.
(CYS)

File S4 MapMan input file for all aphid-responsive genes and corresponding log2-transformed expression values.
(XLS)

File S5 MapMan input file for all *P. syringae*-responsive genes and corresponding log2-transformed expression values.
(XLS)

File S6 Lists of aphid-specific and *P. syringae*-specific genes and their log2-transformed expression values related to biotic stress signalling processes.
(XLS)

File S7 List of the 31 heat shock protein (HSP) genes that were differentially expressed only during aphid treatment and their log2-transformed expression values.
(XLS)

File S8 List of the proteolyitc enzymes differentially expressed during the aphid and *P. syringae* treatments.
(XLS)

File S9 List of genes related to secondary metabolic processes that were differentially regulated during the aphid and *P. syringae* treatments.
(XLS)

File S10 List of genes involved in cell wall precursor synthesis that were differentially regulated during the aphid and *P. syringae* treatments.
(XLS)

File S11 A comparative list of the differentially expressed genes (aphid-specific and *P. syringae*-specific genes) involved in hormonal pathways with their corresponding log2-transformed expression values.
(XLS)

File S12 Details of all of the TFs that were differentially expressed during the aphid and Pseudmonas treatments.
(XLS)

Acknowledgments

The authors thank Torfinn Sparstad for excellent technical assistance.

Author Contributions

Coordinated the study: AMB. Conceived and designed the experiments: PB AMB. Performed the experiments: PB AK DHT. Analyzed the data: PB PW. Contributed reagents/materials/analysis tools: PW AMB. Wrote the paper: PB PW AK AMB.

References

1. Ahuja I, de Vos RC, Bones AM, Hall RD (2010) Plant molecular stress responses face climate change. Trends Plant Sci 15: 664–674.
2. Koornneef M, Meinke D (2010) The development of Arabidopsis as a model plant. Plant Journal 61: 909–921.
3. Pitzschke A, Hirt H (2010) Bioinformatic and systems biology tools to generate testable models of signaling pathways and their targets. Plant Physiol 152: 460–469.
4. Reymond P, Bodenhausen N, Van Poecke RM, Krishnamurthy V, Dicke M, et al. (2004) A conserved transcript pattern in response to a specialist and a generalist herbivore. Plant Cell 16: 3132–3147.
5. Kreps JA, Wu Y, Chang HS, Zhu T, Wang X, et al. (2002) Transcriptome changes for Arabidopsis in response to salt, osmotic, and cold stress. Plant Physiol 130: 2129–2141.
6. Dicke M, van Loon JJ, Soler R (2009) Chemical complexity of volatiles from plants induced by multiple attack. Nat Chem Biol 5: 317–324.
7. Knight H, Knight MR (2001) Abiotic stress signalling pathways: specificity and cross-talk. Trends Plant Sci 6: 262–267.
8. Van Emden HF, Harrington R (2007) Aphids as crop pests/edited by Helmut F. van Emden and Richard Harrington. Wallingford, UK; Cambridge, MA: CABI. xxviii, 717 p., 716 p. of plates p.
9. Goggin FL (2007) Plant-aphid interactions: molecular and ecological perspectives. Curr Opin Plant Biol 10: 399–408.
10. Broekgaarden C, Poelman EH, Steenhuis G, Voorrips RE, Dicke M, et al. (2008) Responses of Brassica oleracea cultivars to infestation by the aphid Brevicoryne brassicae: an ecological and molecular approach. Plant Cell Environ 31: 1592–1605.
11. Spiller NJ, Kimmins FM, Llewellyn M (1985) Fine-Structure of Aphid Stylet Pathways and Its Use in Host Plant-Resistance Studies. Entomologia Experimentalis Et Applicata 38: 293–295.
12. Kusnierczyk A, Winge P, Midelfart H, Armbruster WS, Rossiter JT, et al. (2007) Transcriptional responses of Arabidopsis thaliana ecotypes with different glucosinolate profiles after attack by polyphagous Myzus persicae and oligophagous Brevicoryne brassicae. J Exp Bot 58: 2537–2552.
13. Whalen MC, Innes RW, Bent AF, Staskawicz BJ (1991) Identification of Pseudomonas syringae pathogens of Arabidopsis and a bacterial locus determining avirulence on both Arabidopsis and soybean. Plant Cell 3: 49–59.
14. Fumiaki Katagiri RT, Sheng Yang He (2002) The Arabidopsis Thaliana-Pseudomonas Syringae Interaction. The Arabidopsis Book: American Society of Plant Biologists. pp. e012.
15. Preston GM (2000) Pseudomonas syringae pv. tomato: the right pathogen, of the right plant, at the right time. Mol Plant Pathol 1: 263–275.
16. Weiler EW, Kutchan TM, Gorba T, Brodschelm W, Niesel U, et al. (1994) The Pseudomonas Phytotoxin Coronatine Mimics Octadecanoid Signaling Molecules of Higher-Plants (Vol 345, Pg 9, 1994). Febs Letters 349: 317–317.
17. Brooks DM, Bender CL, Kunkel BN (2005) The Pseudomonas syringae phytotoxin coronatine promotes virulence by overcoming salicylic acid-dependent defences in Arabidopsis thaliana. Mol Plant Pathol 6: 629–639.
18. Melotto M, Underwood W, Koczan J, Nomura K, He SY (2006) Plant stomata function in innate immunity against bacterial invasion. Cell 126: 969–980.
19. Moreau Y, Aerts S, De Moor B, De Strooper B, Dabrowski M (2003) Comparison and meta-analysis of microarray data: from the bench to the computer desk. Trends Genet 19: 570–577.
20. Cramer GR, Urano K, Delrot S, Pezzotti M, Shinozaki K (2011) Effects of abiotic stress on plants: a systems biology perspective. BMC Plant Biol 11: 163.
21. Konika Chawla PB, Kuiper M, Bones AM (2010) Systems Biology: a promising tool to study abiotic stress responses. Omics and Plant Abiotic Stress Tolerance. USA Bentham Publishers.
22. Williams EJ, Bowles DJ (2004) Coexpression of neighboring genes in the genome of Arabidopsis thaliana. Genome Res 14: 1060–1067.
23. Weston DJ, Gunter LE, Rogers A, Wullschleger SD (2008) Connecting genes, coexpression modules, and molecular signatures to environmental stress phenotypes in plants. BMC Syst Biol 2: 16.
24. Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116: 281–297.
25. Zhou X, Wang G, Sutoh K, Zhu JK, Zhang W (2008) Identification of cold-inducible microRNAs in plants by transcriptome analysis. Biochim Biophys Acta 1779: 780–788.
26. Davuluri RV, Sun H, Palaniswamy SK, Matthews N, Molina C, et al. (2003) AGRIS: Arabidopsis gene regulatory information server, an information resource of Arabidopsis cis-regulatory elements and transcription factors. BMC Bioinformatics 4: 25.
27. Bulow L, Bolivar JC, Ruhe J, Brill Y, Hehl R (2012) 'MicroRNA Targets', a new AthaMap web-tool for genome-wide identification of miRNA targets in Arabidopsis thaliana. BioData Min 5: 7.
28. Bulow L, Engelmann S, Schindler M, Hehl R (2009) AthaMap, integrating transcriptional and post-transcriptional data. Nucleic Acids Res 37: D983–986.
29. Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res 27: 297–300.
30. Katari MS, Nowicki SD, Aceituno FF, Nero D, Kelfer J, et al. (2010) VirtualPlant: a software platform to support systems biology research. Plant Physiol 152: 500–515.
31. De Bodt S, Carvajal D, Hollunder J, Van den Cruyce J, Movahedi S, et al. (2010) CORNET: a user-friendly tool for data mining and integration. Plant Physiol 152: 1167–1179.
32. Bethke G, Scheel D, Lee J (2009) Sometimes new results raise new questions: the question marks between mitogen-activated protein kinase and ethylene signaling. Plant Signal Behav 4: 672–674.
33. Lin W, Katagiri F, Glazebrook J (2009) Arabidopsis defense response against Pseudomonas syringae - Effects of major regulatory genes and the impact of coronatine; 2009 17–21 May. 1–4.
34. Thilmony R, Underwood W, He SY (2006) Genome-wide transcriptional analysis of the Arabidopsis thaliana interaction with the plant pathogen Pseudomonas syringae pv. tomato DC3000 and the human pathogen Escherichia coli O157 : H7. Plant Journal 46: 34–53.
35. Ahuja I, Kissen R, Bones AM (2012) Phytoalexins in defense against pathogens. Trends Plant Sci 17: 73–90.
36. Burow M, Zhang ZY, Ober JA, Lambrix VM, Wittstock U, et al. (2008) ESP and ESM1 mediate indol-3-acetonitrile production from indol-3-ylmethyl glucosinolate in Arabidopsis. Phytochemistry 69: 663–671.
37. Pajerowska-Mukhtar K, Dong X (2009) A kiss of death–proteasome-mediated membrane fusion and programmed cell death in plant defense against bacterial infection. Genes Dev 23: 2449–2454.
38. Theis N, Lerdau M (2003) The evolution of function in plant secondary metabolites. International Journal of Plant Sciences 164: S93–S102.
39. Strauss SY, Rudgers JA, Lau JA, Irwin RE (2002) Direct and ecological costs of resistance to herbivory. Trends in Ecology & Evolution 17: 278–285.
40. Foyer CH, Noctor G, van Emden HF (2007) An evaluation of the costs of making specific secondary metabolites: Does the yield penalty incurred by host plant resistance to insects result from competition for resources? International Journal of Pest Management 53: 175–182.
41. Sasidharan R, Chinnappa CC, Staal M, Elzenga JT, Yokoyama R, et al. (2010) Light quality-mediated petiole elongation in Arabidopsis during shade avoidance involves cell wall modification by xyloglucan endotransglucosylase/hydrolases. Plant Physiol 154: 978–990.
42. Rose JK, Braam J, Fry SC, Nishitani K (2002) The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: current perspectives and a new unifying nomenclature. Plant Cell Physiol 43: 1421–1435.
43. De Vos M, Van Oosten VR, Van Poecke RM, Van Pelt JA, Pozo MJ, et al. (2005) Signal signature and transcriptome changes of Arabidopsis during pathogen and insect attack. Mol Plant Microbe Interact 18: 923–937.
44. Browse J, Howe GA (2008) New weapons and a rapid response against insect attack. Plant Physiol 146: 832–838.
45. Kusnierczyk A, Tran DH, Winge P, Jorstad TS, Reese JC, et al. (2011) Testing the importance of jasmonate signalling in induction of plant defences upon cabbage aphid (Brevicoryne brassicae) attack. BMC Genomics 12: 423.
46. Farmer EE, Ryan CA (1992) Octadecanoid Precursors of Jasmonic Acid Activate the Synthesis of Wound-Inducible Proteinase Inhibitors. Plant Cell 4: 129–134.
47. Ulland S, Ian E, Mozuraitis R, Borg-Karlson AK, Meadow R, et al. (2008) Methyl salicylate, identified as primary odorant of a specific receptor neuron type, inhibits oviposition by the moth Mamestra brassicae L. (Lepidoptera, Noctuidae). Chemical Senses 33: 35–46.
48. Luna E, Bruce TJ, Roberts MR, Flors V, Ton J (2011) Next Generation Systemic Acquired Resistance. Plant Physiol.
49. Grant M, Truman B, Zabala MD, Bennett M, Turnbull C (2006) Towards a systems approach to plant defense responses. Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology 143: S137–S137.
50. Kissen R, Rossiter JT, Bones AM (2009) The 'mustard oil bomb': not so easy to assemble?! Localization, expression and distribution of the components of the myrosinase enzyme system. Phytochemistry Reviews 8: 69–86.
51. Kissen R, Bones AM (2009) Nitrile-specifier proteins involved in glucosinolate hydrolysis in Arabidopsis thaliana. J Biol Chem 284: 12057–12070.
52. Eulgem T (2005) Regulation of the Arabidopsis defense transcriptome. Trends Plant Sci 10: 71–78.
53. Reddy AS, Ali GS, Celesnik H, Day IS (2011) Coping with stresses: roles of calcium- and calcium/calmodulin-regulated gene expression. Plant Cell 23: 2010–2032.
54. Gigolashvili T, Berger B, Mock HP, Muller C, Weisshaar B, et al. (2007) The transcription factor HIG1/MYB51 regulates indolic glucosinolate biosynthesis in Arabidopsis thaliana. Plant J 50: 886–901.

55. Munemasa S, Hossain MA, Nakamura Y, Mori IC, Murata Y (2011) The Arabidopsis calcium-dependent protein kinase, CPK6, functions as a positive regulator of methyl jasmonate signaling in guard cells. Plant Physiol 155: 553–561.
56. Wright AJ, Knight H, Knight MR (2002) Mechanically stimulated TCH3 gene expression in Arabidopsis involves protein phosphorylation and EIN6 downstream of calcium. Plant Physiol 128: 1402–1409.
57. Fujita M, Fujita Y, Maruyama K, Seki M, Hiratsu K, et al. (2004) A dehydration-induced NAC protein, RD26, is involved in a novel ABA-dependent stress-signaling pathway. Plant J 39: 863–876.
58. Tamura K, Fukao Y, Iwamoto M, Haraguchi T, Hara-Nishimura I (2010) Identification and characterization of nuclear pore complex components in Arabidopsis thaliana. Plant Cell 22: 4084–4097.
59. Doczi R, Brader G, Pettko-Szandtner A, Rajh I, Djamei A, et al. (2007) The Arabidopsis mitogen-activated protein kinase kinase MKK3 is upstream of group C mitogen-activated protein kinases and participates in pathogen signaling. Plant Cell 19: 3266–3279.
60. Tarasenko VI, Garnik EY, Shmakov VN, Konstantinov YM (2009) Induction of Arabidopsis gdh2 gene expression during changes in redox state of the mitochondrial respiratory chain. Biochemistry (Mosc) 74: 47–53.
61. Prandl R, Hinderhofer K, Eggers-Schumacher G, Schoffl F (1998) HSF3, a new heat shock factor from Arabidopsis thaliana, derepresses the heat shock response and confers thermotolerance when overexpressed in transgenic plants. Mol Gen Genet 258: 269–278.
62. Bricchi I, Bertea CM, Occhipinti A, Paponov IA, Maffei ME (2012) Dynamics of membrane potential variation and gene expression induced by Spodoptera littoralis, Myzus persicae, and Pseudomonas syringae in Arabidopsis. Plos One 7: e46673.
63. Sato M, Tsuda K, Wang L, Coller J, Watanabe Y, et al. (2010) Network modeling reveals prevalent negative regulatory relationships between signaling sectors in Arabidopsis immune signaling. PLoS Pathog 6: e1001011.
64. Bhardwaj V, Meier S, Petersen LN, Ingle RA, Roden LC (2011) Defence responses of Arabidopsis thaliana to infection by Pseudomonas syringae are regulated by the circadian clock. PLoS One 6: e26968.
65. Thilmony R, Underwood W, He SY (2006) Genome-wide transcriptional analysis of the Arabidopsis thaliana interaction with the plant pathogen Pseudomonas syringae pv. tomato DC3000 and the human pathogen Escherichia coli O157: H7. Plant J 46: 34–53.
66. Kusnierczyk A, Winge P, Jorstad TS, Troczynska J, Rossiter JT, et al. (2008) Towards global understanding of plant defence against aphids–timing and dynamics of early Arabidopsis defence responses to cabbage aphid (Brevicoryne brassicae) attack. Plant Cell Environ 31: 1097–1115.
67. Hornik K (2011) The R FAQ.
68. Jorstad TS, Midelfart H, Bones AM (2008) A mixture model approach to sample size estimation in two-sample comparative microarray experiments. BMC Bioinformatics 9: 117.
69. Jorstad TS, Langaas M, Bones AM (2007) Understanding sample size: what determines the required number of microarrays for an experiment? Trends Plant Sci 12: 46–50.
70. Smyth GK (2005) Limma: linear models for microarray data. In: Gentleman VC, Dudoit S, Irizarry R, Huber W, editors. Bioinformatics and Computational Biology Solutions using R and Bioconductor. New York: Springer.
71. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25: 25–29.
72. Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, et al. (2012) The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res 40: D1202–1210.
73. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, et al. (2009) ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 25: 1091–1093.
74. Maere S, Heymans K, Kuiper M (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics 21: 3448–3449.
75. Tardieu F (2003) Virtual plants: modelling as a tool for the genomics of tolerance to water deficit. Trends Plant Sci 8: 9–14.
76. Guo A, He K, Liu D, Bai S, Gu X, et al. (2005) DATF: a database of Arabidopsis transcription factors. Bioinformatics 21: 2568–2569.
77. Rivals I, Personnaz L, Taing L, Potier MC (2007) Enrichment or depletion of a GO category within a class of genes: which test? Bioinformatics 23: 401–407.
78. Ge Y, Dudoit S., Speed T.P. (2003) Resampling-based multiple testing for microarray data analysis. UC Berkeley Dept. of Statistics. 633 p.
79. Casneuf T, Van de Peer Y, Huber W (2007) In situ analysis of cross-hybridisation on microarrays and the inference of expression correlation. BMC Bioinformatics 8: 461.
80. Gutierrez RA, Lejay LV, Dean A, Chiaromonte F, Shasha DE, et al. (2007) Qualitative network models and genome-wide expression data define carbon/nitrogen-responsive molecular machines in Arabidopsis. Genome Biol 8: R7.
81. Griffiths-Jones S (2004) The microRNA Registry. Nucleic Acids Res 32: D109–111.
82. Backman TW, Sullivan CM, Cumbie JS, Miller ZA, Chapman EJ, et al. (2008) Update of ASRP: the Arabidopsis Small RNA Project database. Nucleic Acids Res 36: D982–985.
83. Adai A, Johnson C, Mlotshwa S, Archer-Evans S, Manocha V, et al. (2005) Computational prediction of miRNAs in Arabidopsis thaliana. Genome Res 15: 78–91.
84. Ambros V, Bartel B, Bartel DP, Burge CB, Carrington JC, et al. (2003) A uniform system for microRNA annotation. RNA 9: 277–279.
85. Meyers BC, Axtell MJ, Bartel B, Bartel DP, Baulcombe D, et al. (2008) Criteria for annotation of plant MicroRNAs. Plant Cell 20: 3186–3190.
86. Hsieh LC, Lin SI, Shih AC, Chen JW, Lin WY, et al. (2009) Uncovering small RNA-mediated responses to phosphate deficiency in Arabidopsis by deep sequencing. Plant Physiol 151: 2120–2132.
87. Allen E, Xie Z, Gustafson AM, Carrington JC (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell 121: 207–221.
88. Jones-Rhoades MW, Bartel DP (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol Cell 14: 787–799.
89. Kurihara Y, Watanabe Y (2004) Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. Proc Natl Acad Sci U S A 101: 12753–12758.
90. Sieber P, Wellmer F, Gheyselinck J, Riechmann JL, Meyerowitz EM (2007) Redundancy and specialization among plant microRNAs: role of the MIR164 family in developmental robustness. Development 134: 1051–1060.
91. Nikovics K, Blein T, Peaucelle A, Ishida T, Morin H, et al. (2006) The balance between the MIR164A and CUC2 genes controls leaf margin serration in Arabidopsis. Plant Cell 18: 2929–2945.
92. Merchan F, Boualem A, Crespi M, Frugier F (2009) Plant polycistronic precursors containing non-homologous microRNAs target transcripts encoding functionally related proteins. Genome Biol 10: R136.
93. Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, et al. (2002) Prediction of plant microRNA targets. Cell 110: 513–520.
94. Park W, Li J, Song R, Messing J, Chen X (2002) CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. Curr Biol 12: 1484–1495.
95. Palatnik JF, Allen E, Wu X, Schommer C, Schwab R, et al. (2003) Control of leaf morphogenesis by microRNAs. Nature 425: 257–263.
96. Kurihara Y, Watanabe Y (2004) Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. Proceedings of the National Academy of Sciences of the United States of America 101: 12753–12758.
97. Jones-Rhoades MW, Bartel DP (2004) Computational identification of plant micro-RNAs and their targets, including a stress-induced miRNA. Mol Cell 14: 787–799.
98. Nag A, King S, Jack T (2009) miR319a targeting of TCP4 is critical for petal growth and development in Arabidopsis. Proc Natl Acad Sci U S A 106: 22534–22539.
99. Sunkar R, Zhu JK (2004) Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. Plant Cell 16: 2001–2019.
100. Cuperus JT, Carbonell A, Fahlgren N, Garcia-Ruiz H, Burke RT, et al. (2010) Unique functionality of 22-nt miRNAs in triggering RDR6-dependent siRNA biogenesis from target transcripts in Arabidopsis. Nat Struct Mol Biol 17: 997–1003.
101. Jagadeeswaran G, Saini A, Sunkar R (2009) Biotic and abiotic stress down-regulate miR398 expression in Arabidopsis. Planta 229: 1009–1014.

CHAPTER 4

(Paper III)

Combined stresses in *Arabidopsis thaliana*

Correspondence:
John Mundy
Address: Dept. of Biology, Ole Maaloes Vej 5, 2200 Copenhagen, Denmark DK 2200
Telephone number: (+45) 28754278
Email: mundy@bio.ku.dk

Journal area:
Environmental Stress and Adaptation

Plant Physiology February 2013 pp.112.210773
Published online before print February 2013,
doi: http://dx.doi.org/10.1104/pp.112.210773

Transcriptome responses to combinations of stresses in *Arabidopsis thaliana*

Simon Rasmussen, Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark DK-2800

Pankaj Barah, Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway N-7491

Maria Cristina Suarez-Rodriguez, Department of Biology, University of Copenhagen, Copenhagen, Denmark, DK-2200

Simon Bressendorff, Department of Biology, University of Copenhagen, Copenhagen, Denmark DK-2200

Pia Friis, DMAC, Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark DK-2800

Paolo Costantino, Dept. of Genetics and Molecular Biology, University of Rome La Sapienza, P.le A. Moro, 00185 Rome, Italy

Atle M Bones, Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway N-7491

Henrik Bjørn Nielsen, Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark DK-2800

John Mundy, Department of Biology, University of Copenhagen, Copenhagen, Denmark DK-2200

Financial sources:
ERA-NET Plant Genomics (http://www.erapg.org/) programme, Biotechnology and Functional genomics (FUGE) programmes of the Norwegian Research Council. The microarray facility was supported by the Technical University of Denmark.

Present address Maria Cristina Suarez-Rodriguez: Department of Plant & Microbial Biology, University of California, Berkeley, CA 94720-3100
Present address Pia Friis: Department of Biological Sciences, University of Denver, Denver, CO 80208

Corresponding author:
John Mundy
Address: Dept. of Biology, Ole Maaloes Vej 5, 2200 Copenhagen, Denmark DK 2200
Telephone number: (+45) 28754278
Email: mundy@bio.ku.dk

Abstract

Biotic and abiotic stresses limit agricultural yields, and plants are often simultaneously exposed to multiple stresses. Combinations of stresses such as heat and drought or cold and high light intensity have profound effects on crop performance and yields. Thus, delineation of the regulatory networks and metabolic pathways responding to single and multiple concurrent stresses is required for breeding and engineering crop stress tolerance. Many studies have described transcriptome changes in response to single stresses. However, exposure of plants to a combination of stress factors may require agonistic or antagonistic responses or responses potentially unrelated to responses to the corresponding single stresses. To analyze such responses, we initially compared transcriptome changes in ten *Arabidopsis thaliana* ecotypes using cold, heat, high light, salt and flagellin treatments as single stress factors, as well as their double combinations. This revealed that some 61% of the transcriptome changes in response to double stresses were not predictable from the responses to single stress treatments. It also showed that plants prioritized between potentially antagonistic responses for only 5 to 10% of the responding transcripts. This indicates that plants have evolved to cope with combinations of stresses, and may therefore be bred to endure them. In addition, using a subset of this data from the *Columbia* (Col) and *Landsberg erecta* (L*er)* ecotypes, we have delineated co-expression network modules responding to single and combined stresses.

Keywords: Arabidopsis, stress, multiple stresses, biological networks, microarray, co-expression

Introduction

Plants are often simultaneously exposed to various biotic and abiotic stresses in their natural or agronomic habitats (Ahuja et al., 2010). Roughly 300 cellular stress genes are conserved in all organisms to defend or repair vital macromolecules against environmental factors (Kultz, 2005). However, stress response genes also evolve rapidly as organisms adapt to changing environments. Thus, antifreeze proteins evolved separately in different phyla (Cheng, 1998), and roughly half of the osmoresponsive genes in the model plant *Arabidopsis* are plant-specific (Rabbani et al., 2003). Because biotic and abiotic stresses reduce harvest yields, considerable research has aimed to understand the responses of model plants and crops to single stresses (reviewed in (Hirayama and Shinozaki, 2010; Chew and Halliday, 2011). This work has identified sets of canonical response genes induced by heat, cold, osmotic or high light stresses (Kreps et al., 2002; Seki et al., 2002; Rizhsky et al., 2004; Oono et al., 2006; Kleine et al., 2007; Hannah et al., 2010; Gonzalez-Perez et al., 2011) and in response to pathogen infection and exposure to pathogen associated molecular patterns (Navarro et al., 2004; Nielsen et al., 2007). It has also revealed that plant responses to different stresses are coordinated by complex and often interconnected signaling pathways regulating numerous metabolic networks (Nakashima et al., 2009). Nonetheless, apart from a notable study on the effects of simultaneous drought and heat stress (Rizhsky et al., 2004), the effects of stress combinations have been little studied (Mittler, 2006; Atkinson and Urwin, 2012). Further work is therefore needed if we wish to understand the full complement of stress responses by comparing data on single stresses with data on multiple stress responses. Such data will be relevant to agronomy (Oerke, 1994), and provide tools to answer basic questions about signaling 'crosstalk' in systems biology (Mundy et al., 2006).

Whole genome expression profiling with microarrays is a useful tool to monitor changes in transcript levels and thereby gene expression in response to stresses and other factors (Seki et al., 2009). Most such studies have used *Arabidopsis thaliana* as a model because of its amenability to subsequent forward and reverse genetic analyses (Somerville and Koornneef, 2002). Robust algorithms have been developed for high throughput microarray data to decipher global biological processes and to generate testable biological hypotheses (Harr and Schlotterer, 2006; Chawla, 2011). For example, lists of transcripts differentially responding to different stresses can be generated and ranked by biological criteria (Dudoit et al., 2002; Nielsen et al., 2007). Network based algorithms can successfully deal with some of the complexities of biological and other

physical systems (Barabasi and Oltvai, 2004). Different methods and algorithms have been developed (Chawla, 2011) to construct major types of biological networks including Gene-Metabolite, Protein-Protein Interaction (PPI), Transcriptional Regulatory, Gene Regulatory, and Co-expression networks (Yuan et al., 2008). Co-expression or co-regulation of genes may implicate them in similar biological processes, such that individual modules of genes can be attributed to specific processes. A primary assumption is that, in such networks, strongly co-regulated or co-expressed group of genes participate in similar biological processes such as a signalling or metabolic pathways (Williams and Bowles, 2004). For example, a study by Weston and co-workers showed that a co-expression network-based analysis could delineate population level, adaptive physiological responses of plants to abiotic stress (Weston et al., 2008). In addition, by meta-analysis of microarray data and other publicly available information (Mentzen and Wurtele, 2008), modular co-expression based analysis can dissect regulon organization in the *Arabidopsis* genome by identifying functional modules that share a similar expression profile across multiple spatial, temporal, environmental and genetic conditions.

We conducted a large-scale microarray experiment to analyse plant responses to multiple, concurrent stresses, and to identify the level and functions of stress regulatory networks. To this end, 10 ecotypes of the model *Arabidopsis* were subjected to 5 individual stress treatments and 6 combinations of these stress treatments under the same growth and experimental conditions. Here we present and analyse this homogeneous dataset of ecotype responses to single and combined stresses. Importantly, our analysis shows that, when two stresses were combined, 61% of the transcripts responded on average in modes that could not be predicted from individual single stress treatments. In addition, only a minor fraction (6%) of the transcripts exhibited antagonistic responses to stress combinations under which the plants apparently must prioritize between the responses. Given the novelty of the responses we uncovered, we explored the modular organisation of transcription networks using Weighted Gene Co-expression Networks (WGCN) (Zhang and Horvath, 2005). This permitted us to identify stress responsive modules and potentially key regulatory genes to further understand plant responses to multiple stresses.

Results

Transcript profiling of stress treatments

To investigate the effects of 5 single (cold, high light, salt, heat and flagellin (FLG)) and 6 combined stress treatments (cold and high light, salt and heat, salt and high light, heat and high light, heat and FLG, cold and FLG) on global transcript levels in *Arabidopsis thaliana*, labelled RNA was hybridized in triplicates to *Arabidopsis thaliana* NimbleGen ATH6 microarrays (Table S1). A total of 210 arrays covering stress experiments from 10 different ecotypes (Col, L*er*, C24, Cvi, Kas1, An1, Sha, Kyo2, Eri and Kond) were hybridized, and 3 arrays were identified as outliers (data not shown) and removed to achieve a total of 207 arrays. The array contained probes for 30380 transcripts for which significant changes in levels were determined by comparing single or double stress treatments to ecotype matched controls and used in further analyses.

Responses to single stresses and intraspecific variation

We initially compared the Col and L*er* transcript responses of the single stress treatments to a benchmark set of responses to similar stress treatments. Despite the fact that the benchmark sets are composed from previously published studies using various ecotypes and experimental setups, there was a good overall overlap (Table S2) including key stress responsive transcripts (Table S3). These results indicate that the individual stress treatments we applied had effects similar to those described previously in other analyses, and that our individual stress treatments were appropriate.

We then investigated how the ecotypes responded to single stress treatments to identify ecotype differences (Figure S1). In general, the responses of the ten ecotypes were highly correlated, except for *Columbia* (Col) which behaved as an outlier for responses to heat, salt or high light. Such intraspecific variation in responses to environmental stimuli have been well documented in a number of plant species including *Arabidopsis* ecotypes (Koornneef et al., 2004). As Col and L*er* are widely used for *Arabidopsis* research, we focused on these two ecotypes to obtain consensus results. Using the Col and L*er* transcript sets, we compared transcript overlap and congruency between each of the responses. This revealed low overlap between abiotic and biotic transcript stress responses. Among the single abiotic stresses, cold and high light were most

similar (33% transcript overlap) with very congruent responses such that 87% of the transcripts responded in the same direction (increased or reduced levels; Figure S2). In contrast, considerable dissimilarity was observed between responses to salt and cold or salt and heat single stresses, while responses to salt and high light were more congruent.

Responses to combined stresses

In principle, when plants are exposed to combined stresses, their responses to the single stresses must be modulated to produce a combined response. The responses of a given transcript to two single stresses may be neutral, agonistic, antagonistic or un-related, and the response to the combined stresses may be a combination of such responses. However, as described below, the response of transcripts to combined stresses is not easily predictable. To describe these responses, we clustered significantly responding transcripts (Table S4) from the two single stresses and the combined stress treatments to predefined expression profiles (see Methods) and defined five transcript responses behaviours or modes: *combinatorial*, *cancelled*, *prioritized*, *independent* and *similar* (Figure 1). Transcripts in *combinatorial* mode have similar responses to the two single stresses and different responses to the combined stresses. Transcripts in *cancelled* mode respond differently to the single stresses and are returned to control levels in response to the combined stress treatment. Transcripts in *prioritized* mode respond differently to the single stresses and remain at one of these levels in response to the combined stresses. Importantly, these three modes represent transcript responses or regulatory modes between two stress factors that cannot be predicted from single stress experiments. In contrast, transcripts in *independent* mode, whose regulation pattern do not respond to the addition of a second stress, as well as transcripts in *similar* mode for the two single and the combined stresses, might be more readily identified from single stress experiments.

Comparison of the responses of single versus combined stresses revealed that on average 61% of the transcripts responded in a mode (*combinatorial*, *cancelled* or *prioritized*) in which the two single stress responses interact such that the response to the combined stresses cannot be predicted from the single stress experiments alone (Figure 2a). The extents of these interacting modes were dependent on the particular stresses applied and ranged from 49.3% of the transcripts in the heat and FLG experiment to 73.8% in the salt and heat double stress experiments (Figure 2b). The majority of the interacting transcript responses were *cancelled* or *combinatorial*, with

averages of 29.3% and 24.7% respectively. However, on average only 6.8% of the transcripts responded in the *prioritized* mode in which the plant must decide between antagonistic responses. This indicates that responses to multiple stresses involve relatively few, oppositely responding transcripts. The experiment with the smallest fraction of transcripts responding in *prioritized* mode is the cold and high light experiment (3.0%, Figure 2b). This is in good agreement with the responses to the two single stress treatments, which share a large overlap in transcripts and thus exhibit very congruent responses (Figure S2). In contrast, the combined salt and heat treatments show the highest level of prioritized transcripts (12.1%).

The *independent* and *similar* response modes, in which stress responses apparently do not interact upon combined stress exposure, comprised on average 39% of the transcript responses, and the number of transcripts regulated in the *independent* mode generally was the larger fraction of these (28%). The response to combined salt and heat treatment was the most oppositely directed in that it had the highest level of unpredictable responses (combinatorial, cancellation or prioritized), and with the fewest number of transcripts in *similar* mode (1.5%), whereas the combined cold and high light treatments had the largest number of transcripts in *similar* mode (18.6%). This corresponds well to the congruency and similarity of the single stress treatments and the levels of prioritized transcripts found above for these double stress combinations.

We also searched for transcripts that consistently behaved in an interactive or in a non-interactive mode across all double stress experiments. This identified two transcripts regulated in combinatorial mode in more than four of the experiments: Arabinogalactan protein 10 (*AGP10*, At4g09030) and a non-coding retrotransposon *AT5TE39795* (At5g28913). In addition, 11 transcripts were found to be regulated in more than 4 experiments in cancellation mode: *AOR* (alkenal/one oxidoreductase, At1g23740, Gene Ontology (GO) process oxidative stress), *GBSS1* (UDP-glycosyltransferase, At1g32900, light and low temp), *G-TMT* (γ-tocopherol methyltransferase, At1g64970, oxidative stress), *FKF1* (At1g68050, blue light), At1g73325 (Kunitz trypsin inhibitor), *ATUTR1* (UDP-glycosyl transferase, At2g02810, protein folding), At2g36220 (oxidative stress, high light), *ARC5* (dynamin-like GTPase, At3g19720, circadian clock), *SEX4* (glucan phosphatase, At3g52180, circadian rhythm), *PIF6* (At3g62090, light), *CIPK20* (calcineurin B-like kinase, At5g45820, abiotic stress, abscisic acid).

We then investigated whether transcripts of the particular response modes could be associated to biological functions via their corresponding, significant GO terms (Figure S3, Table S5).

Several patterns appear clear, as three of the response modes are associated with sets of GO terms specific to them. The independent mode is associated with chloroplastic or photosynthetic terms including thylakoid membrane organization, and responses to cyclopentenones and high light intensity. This indicates that the effects of high light treatment may contain a specific set of transcripts that are not influenced by the other stress treatments and may solely be associated to light treatment. The cancellation mode is primarily associated with terms related to secondary metabolism (anthocyanin, indoleacetic acid, phenylpropanoid, etc.) and growth regulation (ethylene and auxin responses). This may indicate that different stresses promote different secondary metabolite pathways and differentially affect growth in response to auxin and ethylene. The combinatorial mode is primarily associated with defense terms (systemic acquired resistance, programmed cell death, salicylate biosynthesis, etc.) which may reflect the intersecting pathways regulating defense against varied pathogens as well as alterations in defense-related programs, such as endoplasmatic stress, in response to abiotic stresses. Not surprisingly, other, more general GO terms such as response to hormones and water deficit, are shared among five and four of the response modes.

We note that the transcript response mode assignments were stable using different transcript significance thresholds (Figure S4) and verified the profiles and assignments for six transcripts by qPCR (Figure S5).

To further investigate the dominance in a pair of stress treatments, we quantified the extent of the influence regulation that each stress imposed on each other in a stress combination (Figure 3). Interestingly, this indicated that the single stress responses were not always dominated by another stress or vice versa. For example, transcripts that significantly responded to the single high light treatment were regulated to a lesser extent when combined with cold or salt than when high light was combined with heat. However, transcripts responding to single heat treatment were regulated less in this combination (HL and heat) than when combined with a different stress than high light. In contrast, the transcript response to FLG treatment alone was in both cases regulated to a lesser extent than the abiotic transcript responses (heat and cold), whereas the cold and salt single stress transcript responses were regulated to a greater extent in their stress combination experiments.

Weighted gene co-expression network analysis

To group these stress responding transcripts into stress regulatory modules, we performed a Weighted Gene Co-expression Network Analysis on the 2236 most differential regulated transcripts with a thresholding power of 7 which was the lowest power for a good fit of the scale-free topology index (see Methods and Supplementary, Figure S6. This identified 9 significant co-expression modules (Table 1 and Table S6). These describe transcript sets that have similar response profiles throughout the sample series of Col and L*er* ecotypes. These modules were associated to stress treatments by singular value decomposition (Langfelder and Horvath, 2007) which significantly associated modules 2, 6, 7 and 8 to abiotic, cold and high light, cold, and biotic stress treatments, respectively (Figure 4). In addition, Gene Ontology (GO) enrichment of transcripts in each module (Table 1) was performed. Only 104 transcripts could not be placed in any of the modules and were assigned to module 10. Furthermore, the network node degree distribution of the WGCNA showed scale-free behaviour and could identify highly connected transcripts inside the modules (Figure S7 and Table S6). Some of these are known to be central stress regulators, in particular a number of transcription factors. Four of the most significant, functionally associated modules are described below.

Biotic stress response module (module 8)

Module 8 showed significant association to biotic stress (FLG) and to combinations of biotic and temperature stresses. It did not show any association to single cold, heat, salt or high ligh. This module included 72 annotated transcription factors (TFs) (Guo et al., 2005). Some of these TFs with higher connectivity within the module were *WRKY6* (At1g62300, (Robatzek and Somssich, 2002)), *WRKY11* (At4g31550, (Journot-Catalino et al., 2006)), *WRKY17* (At2g24570, (Journot-Catalino et al., 2006)), *WRKY22* (At4g01250, (Asai et al., 2002)), *WRKY25* (At2g30250, (Zheng et al., 2007)), *WRKY28* (At4g1817, (van Verk et al., 2011)), *WRKY29* (At4g23550, (Asai et al., 2002)), *WRKY33* (At4g23810, (Zheng et al., 2006), *WRKY40* (At1g80840, (Pandey et al., 2010)), *WRKY55* (At2g40740), *JAZ10* (At5g13220, (Chung and Howe, 2009)), as well as *MYB15* (At3g23250, (Zhou et al., 2011)), and *ANAC13* and *53* (At1g32870, At3g10500). Significantly, all but the last three (*MYB15* and *ANACs 13* & *53*) have been functionally linked to responses to pathogen infection or to the phytohormones ethylene, jasmonate and/or salicylate which

coordinate immune responses. This clearly indicates the relevance of this module to further analyses of how plants orchestrate responses to pathogens.

Cold stress response module (module 7)

Module 7 showed significant association to cold, high light, and to combined cold and high light, and combined cold and FLG. It had very little association to combined salt and high light, and none to single FLG treatment. The latter indicates that combined biotic stress (FLG) and cold generates a unique stress response pattern different from the single stress treatments. Based on their connectivity within the module, two TFs apparently important in cold temperature responses are *PRR7* (At5g02810, (Salome et al., 2010)) and *HMGB2* (At4g23800, (Kwak et al., 2007)). *PRR7* has been implicated as a morning loop component in temperature compensation, while *HMG2* has been shown to be induced by cold treatment. Other highly connected transcripts in this module with functional associations to cold stress include *COR47* (At1g20440), *PGM* (At5g51820), *RSA4* (At5g01410), *LTI30* (At3g50970), *HVA22D* (At4g24960) and *ERD7* (At2g17840) (Oono et al., 2006).

Cold and high light stress response module (module 6)

Module 6 showed significant association to two independent abiotic treatments, cold and high light. The gene enrichment results clearly reflect these associations (Table 1). The presence of some previously reported cold response regulators such as *CBF1* (At4g25490), *CBF2* (At4g25470) and *DREB1A* (At4g25480) clearly implicate this module in responses to cold. The module includes 17 other TFs, and some of these that are higly connected are *IAA19* (At3g15540), At2g46670, *APRR9* (At2g46790), *APRR5* (At5g24470), *ATHB-2* (At4g16780), *CCA1* (At2g46830), *HFR1* (At1g02340), *PIL1* (At2g46970), as well as a *MYB* and homeodomain-like protein (At3g10113), a basic helix-loop-helix protein (At3g21330), and three zinc finger proteins (At1g73870, At5g48250, At5g44260). Most of these TFs have clear functional connections to temperature and light-dependent developmental programs. For example, *APRR5* and *9*, *CCA1*, and *PIL1* are involved in temperature compensation in circadian rhythms (Salome et al., 2010), *HFR1* regulates a phytochrome A dependent photomorphogenesis pathway (Yang et al., 2009), *ATHB-2* regulates photomorphogenesis and shade avoidance (Steindler et al., 1999) in part by modulating auxin responsive growth mediated by *IAA19*

(Tatematsu et al., 2004), and At5g48250 appears to be a target of FLC-independent effects of the autonomous floral pathway. The module also has some association to combined cold and FLG, but no association to single FLG treatment. This again highlights the fact that the interaction of a biotic stress factor (FLG) with cold was unique.

Abiotic stress response module (module 2)

Module 2 exhibited significant association to both single and combined abiotic stresses such as single heat, high light, salt, and combined salt and high light or salt and heat, and slight association to cold and high-light. The module did not exhibit significant association to FLG treatments. Functional enrichment analysis (Table 1) showed significantly enriched categories associated with various abiotic stress responses. This abiotic stress responsive module has 24 TFs including *IAA1*, *5* and *17* (At4g14560, At1g15580, At1g04250), *RGL1* (At1g66350, (Wen and Chang, 2002)), *MYB59*, *73* and *86* (At5g59780, At4g37260, At5g26660, (Mu et al., 2009)), *TCP3* and *14* (At1g53230, At3g47620, (Kieffer et al., 2011)), *HSFB4* (At1g46264, (Begum et al., 2012)), *HB31* (At1g14440, (Torti et al., 2012)), *RVE2* (At5g37260, (Zhang et al., 2007)). These TFs have been implicated in developmental processes such as root, leaf and internode growth which are regulated by auxin and giberellin. Little is known of the functions of other TFs in this module including *DREBA-4* (At5g52020), *SRS6* (At3g54430), *emb2746* (At5g63420), *HMG1/2-like* (At1g76110), *DOF* (At5g65590), *ZN-finger* (At2g37430), *SMAD/FHA* (At2g21530), and two *BLH* (At2g16400, At5g5091).

Discussion

This study was designed to investigate combinations of stresses that mimic harsh environmental conditions that occur in the field. We therefore examined high light intensity that may occur due to diminished ozone layer protection, in combination with low or high temperature, as well as saline irrigation combined with high temperature. Furthermore, we were interested in the interaction of abiotic stresses such as low or high temperatures combined with biotic stresses such as the immune response to pathogens induced by the flagellin elicitor. Here we provide a benchmarked set of transcript profiles responding to such single and combined stresses. We

initially determined that the transcript responses we detected to five single stresses were similar to those of a benchmark set of responses to similar stresses described by others. This meant that the stresses we applied were comparable to other studies, and permitted us to then analyse the transcript profiles responding to combinations of these stresses. We note, however, two potential limitations to the stress applications that others and we have used. First, due to logistics and costs, we harvested rosettes after stress applications at a single time point based on previous studies. The assessment will therefore not detect the temporal dynamics of single stress responses or of the interactions between combinations of stresses. Nonetheless, we chose the sampling time point before the onset of visual stress symptoms to attempt to detect responses caused specifically by the environmental insult and not systemic responses that occur 'downstream' or as indirect consequences of the specific stress combinations. Second, the assessment does not detect the relative intensities of single stresses. Despite these caveats, we expect that our profiles, as the largest robust such dataset at present, can be productively mined by other researchers.

The second aim of this study was to analyse the response profiles to identify the behaviours or modes of regulation of sets or modules of transcripts in response to combined stresses. To these ends we used two types of analyses to attempt to capture the range of responses displayed in the data. Our analysis of transcript behaviours or modes was based on clustering the top 500 most significantly responding transcripts for each single stress and for the stress combinations. Five transcript response modes (combinatorial, cancelled, prioritized, independent and similar) were identified which describe potential transcript regulatory complexity and assign predictabilities to the responses of individual transcripts. Importantly, the combinatorial, cancelled and prioritized response modes on average comprise 61% of the total transcripts, and these modes cannot be predicted from the corresponding single stress experiments. That the majority of the transcript responses are not predictable when two stresses are combined points to the limitations of attempts to delineate common stress responses or points of 'cross-talk' between signaling pathways during multiple stresses by simply identifying overlapping sets of genes that are regulated by both stresses (Kreps et al., 2002; Mentzen and Wurtele, 2008; Carrera et al., 2009). Additionally unpredictable are transcripts which respond only to the combined stresses and not to either individual stress. For example, using the 500 most significant transcripts from each stress experiment, we found that 55.8% to 79.6% of the transcripts regulated in the double stress experiment were not among the most significant transcripts in the corresponding single stress

experiments. Therefore, these potentially novel stress related transcripts will be absent in analyses using only single stress response experiments.

Overall, the most abundant transcript response modes are cancelled, independent and combinatorial which together on average include 85% of the total transcripts. The combinatorial response (27.1%) of transcripts with similar responses to two single stresses and different responses to the combined stresses is indicative of the level of interaction between responses to different stresses. For example, combined heat and flagellin treatment (FLG) has the lowest level of combinatorial transcripts (7.1%) and the highest level of independent transcripts (39.7%) along with a relatively high level of prioritized transcripts (9.1%). The low overlap in transcripts between the abiotic and the biotic single stress experiments (~5%) (Figure S2) could indicate that the early transcriptional responses to an innate immune elicitor versus an extreme physical change in the environment target different transcript sets.

The independent mode (28.2%) contains transcripts that are regulated in either of the single stresses and whose regulation is maintained without interference from the other stress. These transcripts define the proportion of the responses to combined stresses that are not shared by, and do not interfere with, the responses to the single stresses. For example, responses to combined cold and high light have the lowest level of independent (18.5%) and prioritized (3.0%) transcripts. This is in line with the highest (33%) and most congruent (0.87) transcript overlap between responses to cold or high light singly, and the highest similar response mode of combined cold and high light (18.6%). In addition, 60% of the transcripts respond to combined cold and high light in cancellation or combinatorial mode, indicative of strong regulatory interactions between these two stresses. This may mirror the ecology of the temperate *Arabidopsis* ecotypes used here in that combined cold and high light stress is common in temperate regions (Ivanov et al., 2012).

The most common mode is cancelled (30.6%) in which transcripts responding oppositely to single stresses return to control levels in response to the combined stresses. For example, responses to salt or heat stress are significantly dissimilar (0.31) and the transcript response modes to combined salt and heat have the highest level of prioritized transcripts (12.1%) of all the combined stress experiments. In addition, combined salt and heat have the lowest level of similar transcripts (1%) compared to an average of 11.2% for all combined stress experiments, and very high fold changes compared to the single salt or heat stresses (Figure 3). The latter

indicates that, under the conditions we applied, responses to heat stress largely dominate responses to salt stress. Despite the caveats noted above concerning stress durations and intensities, this result suggests that adaptation to combined salt and heat stresses is more difficult than adaptation to the other combined stresses assessed here. Nonetheless, such a conclusion may only apply to temperate plants with similar ecologies to *Arabidopsis* given the extent of adaptive diversity in related plants such as the halophyte *Thellungiella salsuginea* (Wu et al., 2012).

Our description of the transcript sets or modules responding to combined stresses employed weighted gene co-expression analysis (Langfelder and Horvath, 2008) to hierarchically cluster modules with similar transcript profiles. The resulting weighted transcriptional networks exhibited scale-free, modular topology and the clustering resulted into 9 significant modules. Eigengene significance based analysis showed that among the 9 modules, 4 have significant association to different biotic and abiotic stresses. Module 2 appears to be associated to abiotic stress, while module 6 exhibits cold and high light associated response signatures. Modules 7 and 8 have significant association to biotic stress (FLG) and to combinations of biotic and temperature stresses. Transcripts in modules 7 and 8 may be useful for addressing agronomic problems such as reduced crop productivity due to new pathogen invasions coupled with temperature stress in predicted scenarios of global climate change. Using networks connectivity and gene significance measures, supported by existing information from *Arabidopsis* transcription factor databases, we have identified a number of transcription factors as targets for future translational experiments to engineer increased stress resistance in crops.

Materials and methods

Plant stress treatments

Arabidopsis thaliana plants of ecotypes (Col, L*er*, C24, Cvi, Kas1, An1, Sha, Kyo2, Eri and Kond) were subjected to the following stress treatments: Salt, Cold, Heat, High Light (HL), Salt+Heat, Salt+HL, Cold+HL, Heat+HL, as well as FLG (Flagellin, flg22 peptide), Cold+FLG, Heat+FLG. In addition, FLG treated plants were grown with two control conditions, ‘control’ and ‘control + Silwet’ (control for the effect of silwet detergent used for flagellin application). Stress treatments were selected from previous studies (Kreps et al., 2002; Seki et al., 2002; Kilian et al.,

2007) and microarray experiments compiled at www.weigelworld.org/resources/microarray/AtGenExpress. Combinations of high light (800 µm photons $m^{-2}s^{-1}$), cold (10°C), heat (38°C), high salinity (100 mM NaCl), and foliar spray application of a bacterial elicitor (20 µM flagellin peptide flg22) were performed in environmentally controlled chambers (RISØ DTU National Laboratory, Roskilde, Denmark). A pilot study using Col and L*er* ecotypes was performed to identify sub-lethal doses of combined stress treatments. This identified an optimal period of three hours before the onset of visible phenotypic responses such as wilting. To ensure independence between biological replicates the stress treatments and plant growth were done in three independent batches. Each stress treatment lasted three hours and was done on three weeks old plants. The high salinity treatments were performed by soil irrigation with 100mM NaCl solution. In order to saturate the soil, irrigation with the saline solution started at the end of the light period the night before collection and was refilled at the onset of the combined treatment. For the Cold+HL treatment, heat from 3 sodium lamps was displaced by circulating fans and a plexiglass shield and ambient plant temperature maintained by ice trays and monitored at 10°C with an infrared thermometer (ThermaTwin TN408LC). To reduce effects of circadian rhythmicity, treatments were performed 5 hours after chamber dawn. After stress treatments, leaves were collected and frozen in liquid nitrogen.

RNA isolation and microarray hybridizations

Total RNA samples were isolated (RiboPure kit, Ambion, Austin, TX, USA) and reverse transcription of messenger RNA was performed from total RNA using a double stranded cDNA synthesis kit (Superscript, Invitrogen). The cDNA obtained was subsequently labeled with Cy3 and the product precipitated using NimbleGen kits according to the NimbleGen Gene Expression protocol for microarrays. 4 micrograms of the labeled products were loaded onto microarrays, hybridized overnight, and washed in the NimbleGen Wash Buffer Kit following the NimbleGen protocol. Scanning was performed on a Roche 2-microns scanner and the images analyzed with the NimbleScan software. The microarray used was the *Arabidopsis thaliana* NimbleGen 12-plex chips using the ATH6 build (GEO: GPL16226) in a Latin Square design with 4 independent probes per transcript. A total of 210 arrays were hybridized.

Microarray data pre-processing

Data was imported into *R* (R Core Team, 2012) using the *oligo* (Carvalho and Irizarry, 2010) and *pdInfoBuilder* packages (Falcon, 2012) using the AgilentAT6 build. If more than one scan was available for an array the best scan was selected using Singular Value Decomposition (SVD) as the array with the lowest residuals. The data was normalized using quantiles (Bolstad et al., 2003) and 3 outliers were removed by comparing the arrays using Pearson correlation coefficient and SVD plots, giving a total of 207 arrays for the analysis. Expression indexes were calculated using RMA (Irizarry et al., 2003). All statistical comparisons between experiments were performed by student´s t-test using the normalized log2 transcript expression indexes. All treatments were compared to the control experiments except treatments including FLG, which were compared to control and silwet samples. Transcript annotation was acquired from TAIR10 (Lamesch et al., 2012) using *biomaRt* data mining tool (Guberman et al., 2011). The microarray data is available at the Gene Expression Omnibus (GEO) with the record GSE41935.

Benchmarking

Because the benchmarking gene sets are derived from various experiments, ecotypes and sources, we used all transcripts from the Col and L*er* single stress experiments with a p-value ≤ 0.01 as input for the benchmarks. Additionally we used the top 500 most significant transcripts from each treatment from comparisons using all ecotypes as a single group. The benchmarking gene sets were derived from (Ashburner et al., 2000; Huala et al., 2001; Navarro et al., 2004; Thimm et al., 2004; Kilian et al., 2007; Kleine et al., 2007; Kant et al., 2008; Papdi et al., 2008; Shameer et al., 2009; Gonzalez-Perez et al., 2011; Less et al., 2011; Avin-Wittenberg et al., 2012; Kilian et al., 2012) and are listed in Table S2.

Transcriptional response modes

For each stress combination, transcript sets were created by the union of the top 500 most significantly responding transcripts for each single stress and for the combination of the two stresses. Hereafter the transcript sets were clustered using Pearson Correlation Coefficient (PCC) to 20 pre-defined expression profiles each categorizing a potential expression pattern that may occur upon multiple stress application. The pre-defined expression profile with the highest PCC was selected for each transcript. The transcriptional response modes (combinatorial, cancelled,

prioritized, independent and similar) were a multiset of several pre-defined expression profiles. Association of transcriptional response modes to Gene Ontology (GO) terms was performed using GO-slim from TAIR as of 12/12/2012 (Berardini et al., 2004) using Fisher's exact test with a significance threshold of 10^{-5} after Bonferroni correction. To simplify the network we for GO-terms with identical connectivity only kept the most specific term.

qPCR verification of microarray transcripts

Quantitative PCR was performed using Brilliant II SYBR green one step kit (Agilent Technologies) with 10 pmol of each primer and 12.5 ng total cDNA in 10 µl and the reactions were run on a CFX 96 Thermocycler (BioRad). For this verification, biological triplicate *Col* samples were used and relative log2 expression determined using ACT2 (AT3G18780), which was determined to be highly expressed with minimal variation across the different treatments in the microarray data. All primer efficiencies were within 100+/-2% and expression levels were calculated assuming 100% efficiency. Primers used and an agarose gel of PCR products matching the expected product sizes is shown in Table S7.

Weighted Gene Co-expression Analysis (WGCNA)

A Weighted Gene Co-expression network was constructed using the R-package *WGCNA* (Langfelder and Horvath, 2008) using a united list of significant transcripts ($p \leq 0.01$) generated from t-tests between control and treated plants in the Col and L*er* ecotypes. A total of 2236 transcripts that responded to at least two of the stress treatments were used to construct the weighted network from the normalized expression data by transforming the pair wise gene correlation matrix to a weighted matrix with a scaling factor beta ($\beta=7$) and using the assumption that biological networks are scale-free. Here weight represents the connection strength between gene pairs (Zhang and Horvath, 2005). To minimize the effects of intrinsic noise in high-throughput transcriptomic data, we transformed the adjacency into a topological overlap matrix (Yip and Horvath, 2007). The tree was created using hierarchical clustering and the dynamic tree cut algorithm was used to identify modules with similar expression patterns. A minimum module size of 30 and a height cut of 0.25 corresponding to a correlation of 0.8 were used to merge similar transcripts. The module eigengenes were used to define measures of module membership (at the significance level $p \leq 0.001$), intramodular connectivity and gene significance (Langfelder

and Horvath, 2007). Intramodular connectivity of transcripts were used to identify hubs in the modules and was measured by computing whole network connectivity kTotal, the within module connectivity kWithin and the outgoing connectivity kOut=kTotal-kWithin. The Biological Networks Gene Ontology tool (BiNGO) (Maere et al., 2005), an open-source Java tool, was used to determine which Gene Ontology (GO) terms that were significantly overrepresented in our module transcript lists (p-values were Bonferroni corrected).

Supplemental Material

The following materials are available in the online version of this article.

Supplemental Figure S1. Comparison of ecotype responses to the single abiotic stresses (salt, cold, heat and high light).

Supplemental Figure S2. Overlap congruence between responses to individual stress treatments based on comparisons between top 500 transcripts of each single treatment.

Supplemental Figure S3. Network of transcriptional response modes and their associated GO terms.

Supplemental Figure S4. Comparison using different thresholds for generating the transcriptional response modes shows consistent profiles.

Supplemental Figure S5. Micorarray and qPCR transcript profiles for six transcripts verifying the transcriptional data and profile assignments.

Supplemental Figure S6. Hierarchical clustering, soft threshold and clustering dendrogram of transcripts of the WGCNA network.

Supplemental Figure S7. Scale-free behaviour of network node degree distribution of the WGCNA network.

Supplemental Table S1. Overview of the experimental setup with regard to ecotypes, stress treatments and biological replicates.

Supplemental Table S2. Benchmark of single stress experiment data versus previous single stress experiments showing overlap, including the referene gene sets used for benchmarking.

Supplemental Table S3. Selection of the key single stress benchmark genes identified in benchmarking.

Supplemental Table S4. Top 500 regulated transcripts for each of the stress treatments from the combined analysis of Col and L*er* ecotypes.

Supplemental Table S5. The p-values of the GO terms and transcriptional response modes for overrepresentation used to build the network.

Supplemental Table S6. Table with WGCNA module membership and connectivity for each of the included transcripts.

Supplemental Table S7. Table of primers used for qPCR and agarose gel showing PCR products with expected sizes.

Acknowledgements

The authors acknowledge Lene Holberg Blicher, Luma Odish and Suksawad Vongvisuttikun for technical assistance.

Literature cited

Ahuja I, de Vos RC, Bones AM, Hall RD (2010) Plant molecular stress responses face climate change. Trends Plant Sci **15:** 664-674

Asai T, Tena G, Plotnikova J, Willmann MR, Chiu WL, Gomez-Gomez L, Boller T, Ausubel FM, Sheen J (2002) MAP kinase signalling cascade in Arabidopsis innate immunity. Nature **415:** 977-983

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet **25:** 25-29

Atkinson NJ, Urwin PE (2012) The interaction of plant biotic and abiotic stresses: from genes to the field. J Exp Bot **63:** 3523-3543

Avin-Wittenberg T, Tzin V, Angelovici R, Less H, Galili G (2012) Deciphering energy-associated gene networks operating in the response of Arabidopsis plants to stress and nutritional cues. Plant J **70:** 954-966

Barabasi AL, Oltvai ZN (2004) Network biology: Understanding the cell's functional organization. Nature Reviews Genetics **5:** 101-U115

Begum T, Reuter R, Schoffl F (2012) Overexpression of AtHsfB4 induces specific effects on root development of Arabidopsis. Mech Dev

Berardini TZ, Mundodi S, Reiser L, Huala E, Garcia-Hernandez M, Zhang P, Mueller LA, Yoon J, Doyle A, Lander G, Moseyko N, Yoo D, Xu I, Zoeckler B, Montoya M, Miller N, Weems D, Rhee SY (2004) Functional annotation of the Arabidopsis genome using controlled vocabularies. Plant Physiol **135:** 745-755

Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics **19:** 185-193

Carrera J, Rodrigo G, Jaramillo A, Elena SF (2009) Reverse-engineering the Arabidopsis thaliana transcriptional network under changing environmental conditions. Genome Biol **10:** R96

Carvalho BS, Irizarry RA (2010) A framework for oligonucleotide microarray preprocessing. Bioinformatics **26:** 2363-2367

Chawla KB, P.; Kuiper, M.; Bones, A. M. (2011) Systems Biology: a promising tool to study abiotic stress responses. *In* N Tuteka, Gill, S., S., Tuteja, R., ed, Omics and Plant Abiotic Stress Tolerence. Betham eBooks: International Centre for Genetic Engineering and Biotechnology, India

Cheng CH (1998) Evolution of the diverse antifreeze proteins. Curr Opin Genet Dev **8:** 715-720

Chew YH, Halliday KJ (2011) A stress-free walk from Arabidopsis to crops. Curr Opin Biotechnol **22:** 281-286

Chung HS, Howe GA (2009) A critical role for the TIFY motif in repression of jasmonate signaling by a stabilized splice variant of the JASMONATE ZIM-domain protein JAZ10 in Arabidopsis. Plant Cell **21:** 131-145

Dudoit S, Yang YH, Callow MJ, Speed TP (2002) Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. Statistica Sinica **12:** 111-139

Falcon SC, B.; Carey, V.; Settles, M. (2012) pdInfoBuilder: Platform Design Information Package Builder. *In*,

Gonzalez-Perez S, Gutierrez J, Garcia-Garcia F, Osuna D, Dopazo J, Lorenzo O, Revuelta JL, Arellano JB (2011) Early transcriptional defense responses in Arabidopsis cell suspension culture under high-light conditions. Plant Physiol **156:** 1439-1456

Guberman JM, Ai J, Arnaiz O, Baran J, Blake A, Baldock R, Chelala C, Croft D, Cros A, Cutts RJ, Di Genova A, Forbes S, Fujisawa T, Gadaleta E, Goodstein DM, Gundem G, Haggarty B, Haider S, Hall M, Harris T, Haw R, Hu S, Hubbard S, Hsu J, Iyer V, Jones P, Katayama T, Kinsella R, Kong L, Lawson D, Liang Y, Lopez-Bigas N, Luo J, Lush M, Mason J, Moreews F, Ndegwa N, Oakley D, Perez-Llamas C, Primig M, Rivkin E, Rosanoff S, Shepherd R, Simon R, Skarnes B, Smedley D, Sperling L, Spooner W, Stevenson P, Stone K, Teague J, Wang J, Whitty B, Wong DT, Wong-Erasmus M, Yao L, Youens-Clark K, Yung C, Zhang J, Kasprzyk A (2011) BioMart Central Portal: an open database network for the biological community. Database (Oxford) **2011:** bar041

Guo A, He K, Liu D, Bai S, Gu X, Wei L, Luo J (2005) DATF: a database of Arabidopsis transcription factors. Bioinformatics **21:** 2568-2569

Hannah MA, Caldana C, Steinhauser D, Balbo I, Fernie AR, Willmitzer L (2010) Combined transcript and metabolite profiling of Arabidopsis grown under widely variant growth conditions facilitates the identification of novel metabolite-mediated regulation of gene expression. Plant Physiol **152:** 2120-2129

Harr B, Schlotterer C (2006) Comparison of algorithms for the analysis of Affymetrix microarray data as evaluated by co-expression of genes in known operons. Nucleic Acids Res **34:** e8

Hirayama T, Shinozaki K (2010) Research on plant abiotic stress responses in the post-genome era: past, present and future. Plant J **61:** 1041-1052

Huala E, Dickerman AW, Garcia-Hernandez M, Weems D, Reiser L, LaFond F, Hanley D, Kiphart D, Zhuang M, Huang W, Mueller LA, Bhattacharyya D, Bhaya D, Sobral BW, Beavis W, Meinke DW, Town CD, Somerville C, Rhee SY (2001) The Arabidopsis Information Resource (TAIR): a comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. Nucleic Acids Res **29:** 102-105

Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP (2003) Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res **31:** e15

Ivanov AG, Rosso D, Savitch LV, Stachula P, Rosembert M, Oquist G, Hurry V, Huner NP (2012) Implications of alternative electron sinks in increased resistance of PSII and PSI photochemistry to high light stress in cold-acclimated Arabidopsis thaliana. Photosynth Res **113:** 191-206

Journot-Catalino N, Somssich IE, Roby D, Kroj T (2006) The transcription factors WRKY11 and WRKY17 act as negative regulators of basal resistance in Arabidopsis thaliana. Plant Cell **18:** 3289-3302

Kant P, Gordon M, Kant S, Zolla G, Davydov O, Heimer YM, Chalifa-Caspi V, Shaked R, Barak S (2008) Functional-genomics-based identification of genes that regulate Arabidopsis responses to multiple abiotic stresses. Plant Cell Environ **31:** 697-714

Kieffer M, Master V, Waites R, Davies B (2011) TCP14 and TCP15 affect internode length and leaf shape in Arabidopsis. Plant J **68:** 147-158

Kilian J, Peschke F, Berendzen KW, Harter K, Wanke D (2012) Prerequisites, performance and profits of transcriptional profiling the abiotic stress response. Biochim Biophys Acta **1819:** 166-175

Kilian J, Whitehead D, Horak J, Wanke D, Weinl S, Batistic O, D'Angelo C, Bornberg-Bauer E, Kudla J, Harter K (2007) The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. Plant J **50:** 347-363

Kleine T, Kindgren P, Benedict C, Hendrickson L, Strand A (2007) Genome-wide gene expression analysis reveals a critical role for CRYPTOCHROME1 in the response of Arabidopsis to high irradiance. Plant Physiol **144:** 1391-1406

Koornneef M, Alonso-Blanco C, Vreugdenhil D (2004) Naturally occurring genetic variation in Arabidopsis thaliana. Annu Rev Plant Biol **55:** 141-172

Kreps JA, Wu Y, Chang HS, Zhu T, Wang X, Harper JF (2002) Transcriptome changes for Arabidopsis in response to salt, osmotic, and cold stress. Plant Physiol **130:** 2129-2141

Kultz D (2005) Molecular and evolutionary basis of the cellular stress response. Annu Rev Physiol **67:** 225-257

Kwak KJ, Kim JY, Kim YO, Kang H (2007) Characterization of transgenic Arabidopsis plants overexpressing high mobility group B proteins under high salinity, drought or cold stress. Plant Cell Physiol **48:** 221-231

Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander DL, Garcia-Hernandez M, Karthikeyan AS, Lee CH, Nelson WD, Ploetz L, Singh S, Wensel A, Huala E (2012) The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res **40:** D1202-1210

Langfelder P, Horvath S (2007) Eigengene networks for studying the relationships between co-expression modules. BMC Syst Biol **1:** 54

Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics **9:** 559

Less H, Angelovici R, Tzin V, Galili G (2011) Coordinated gene networks regulating Arabidopsis plant metabolism in response to various stresses and nutritional cues. Plant Cell **23:** 1264-1271

Maere S, Heymans K, Kuiper M (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics **21:** 3448-3449

Mentzen WI, Wurtele ES (2008) Regulon organization of Arabidopsis. BMC Plant Biol **8:** 99

Mittler R (2006) Abiotic stress, the field environment and stress combination. Trends Plant Sci **11:** 15-19

Mu RL, Cao YR, Liu YF, Lei G, Zou HF, Liao Y, Wang HW, Zhang WK, Ma B, Du JZ, Yuan M, Zhang JS, Chen SY (2009) An R2R3-type transcription factor gene AtMYB59 regulates root growth and cell cycle progression in Arabidopsis. Cell Res **19:** 1291-1304

Mundy J, Nielsen HB, Brodersen P (2006) Crosstalk. Trends Plant Sci **11:** 63-64

Nakashima K, Ito Y, Yamaguchi-Shinozaki K (2009) Transcriptional regulatory networks in response to abiotic stresses in Arabidopsis and grasses. Plant Physiol **149:** 88-95

Navarro L, Zipfel C, Rowland O, Keller I, Robatzek S, Boller T, Jones JD (2004) The transcriptional innate immune response to flg22. Interplay and overlap with Avr gene-dependent defense responses and bacterial pathogenesis. Plant Physiol **135:** 1113-1128

Nielsen HB, Mundy J, Willenbrock H (2007) Functional Associations by Response Overlap (FARO), a functional genomics approach matching gene expression phenotypes. PLoS One **2:** e676

Oerke ECD, H. W.; Schönbeck, F.; Weber, A. (1994) Crop production and crop protection: estimated losses in major food and cash crops. Elsevier, New York

Oono Y, Seki M, Satou M, Iida K, Akiyama K, Sakurai T, Fujita M, Yamaguchi-Shinozaki K, Shinozaki K (2006) Monitoring expression profiles of Arabidopsis genes during cold acclimation and deacclimation using DNA microarrays. Funct Integr Genomics **6:** 212-234

Pandey SP, Roccaro M, Schon M, Logemann E, Somssich IE (2010) Transcriptional reprogramming regulated by WRKY18 and WRKY40 facilitates powdery mildew infection of Arabidopsis. Plant J **64:** 912-923

Papdi C, Abraham E, Joseph MP, Popescu C, Koncz C, Szabados L (2008) Functional identification of Arabidopsis stress regulatory genes using the controlled cDNA overexpression system. Plant Physiol **147:** 528-542

R Core Team (2012) R: A Language and Environment for Statistical Computing. *In*. R Foundation for Statistical Computing, Vienna, Austria

Rabbani MA, Maruyama K, Abe H, Khan MA, Katsura K, Ito Y, Yoshiwara K, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. Plant Physiol **133:** 1755-1767

Rizhsky L, Liang H, Shuman J, Shulaev V, Davletova S, Mittler R (2004) When defense pathways collide. The response of Arabidopsis to a combination of drought and heat stress. Plant Physiol **134:** 1683-1696

Robatzek S, Somssich IE (2002) Targets of AtWRKY6 regulation during plant senescence and pathogen defense. Genes Dev **16:** 1139-1149

Salome PA, Weigel D, McClung CR (2010) The role of the Arabidopsis morning loop components CCA1, LHY, PRR7, and PRR9 in temperature compensation. Plant Cell **22:** 3650-3661

Seki M, Narusaka M, Ishida J, Nanjo T, Fujita M, Oono Y, Kamiya A, Nakajima M, Enju A, Sakurai T, Satou M, Akiyama K, Taji T, Yamaguchi-Shinozaki K, Carninci P, Kawai J, Hayashizaki Y, Shinozaki K (2002) Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. Plant J **31:** 279-292

Seki M, Okamoto M, Matsui A, Kim J-M, Kurihara Y, Ishida J, Morosawa T, Kawashima M, To T, Shinozaki K (2009) Microarray Analysis for Studying the Abiotic Stress Responses in Plants. *In* SM Jain, DS Brar, eds, Molecular Techniques in Crop Improvement. Springer Netherlands, pp 333-355

Shameer K, Ambika S, Varghese SM, Karaba N, Udayakumar M, Sowdhamini R (2009) STIFDB-Arabidopsis Stress Responsive Transcription Factor DataBase. Int J Plant Genomics **2009:** 583429

Somerville C, Koornneef M (2002) Timeline - A fortunate choice: the history of Arabidopsis as a model plant. Nature Reviews Genetics **3:** 883-889

Steindler C, Matteucci A, Sessa G, Weimar T, Ohgishi M, Aoyama T, Morelli G, Ruberti I (1999) Shade avoidance responses are mediated by the ATHB-2 HD-zip protein, a negative regulator of gene expression. Development **126:** 4235-4245

Tatematsu K, Kumagai S, Muto H, Sato A, Watahiki MK, Harper RM, Liscum E, Yamamoto KT (2004) MASSUGU2 encodes Aux/IAA19, an auxin-regulated protein that functions together with the transcriptional activator NPH4/ARF7 to regulate differential growth responses of hypocotyl and formation of lateral roots in Arabidopsis thaliana. Plant Cell **16:** 379-393

Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, Kruger P, Selbig J, Muller LA, Rhee SY, Stitt M (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

Torti S, Fornara F, Vincent C, Andres F, Nordstrom K, Gobel U, Knoll D, Schoof H, Coupland G (2012) Analysis of the Arabidopsis shoot meristem transcriptome during floral transition identifies distinct regulatory patterns and a leucine-rich repeat protein that promotes flowering. Plant Cell **24:** 444-462

van Verk MC, Bol JF, Linthorst HJ (2011) WRKY transcription factors involved in activation of SA biosynthesis genes. BMC Plant Biol **11:** 89

Wen CK, Chang C (2002) Arabidopsis RGL1 encodes a negative regulator of gibberellin responses. Plant Cell **14:** 87-100

Weston DJ, Gunter LE, Rogers A, Wullschleger SD (2008) Connecting genes, coexpression modules, and molecular signatures to environmental stress phenotypes in plants. BMC Syst Biol **2:** 16

Williams EJ, Bowles DJ (2004) Coexpression of neighboring genes in the genome of Arabidopsis thaliana. Genome Res **14:** 1060-1067

Wu HJ, Zhang Z, Wang JY, Oh DH, Dassanayake M, Liu B, Huang Q, Sun HX, Xia R, Wu Y, Wang YN, Yang Z, Liu Y, Zhang W, Zhang H, Chu J, Yan C, Fang S, Zhang J, Wang Y, Zhang F, Wang G, Lee SY, Cheeseman JM, Yang B, Li B, Min J, Yang L, Wang J, Chu C, Chen SY, Bohnert HJ, Zhu JK, Wang XJ, Xie Q (2012) Insights into salt tolerance from the genome of Thellungiella salsuginea. Proc Natl Acad Sci U S A **109:** 12219-12224

Yang SW, Jang IC, Henriques R, Chua NH (2009) FAR-RED ELONGATED HYPOCOTYL1 and FHY1-LIKE associate with the Arabidopsis transcription factors LAF1 and HFR1 to transmit phytochrome A signals for inhibition of hypocotyl elongation. Plant Cell **21:** 1341-1359

Yip AM, Horvath S (2007) Gene network interconnectedness and the generalized topological overlap measure. BMC Bioinformatics **8:** 22

Yuan JS, Galbraith DW, Dai SY, Griffin P, Stewart CN, Jr. (2008) Plant systems biology comes of age. Trends Plant Sci **13:** 165-171

Zhang B, Horvath S (2005) A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol **4:** Article17

Zhang X, Chen Y, Wang ZY, Chen Z, Gu H, Qu LJ (2007) Constitutive expression of CIR1 (RVE2) affects several circadian-regulated processes and seed germination in Arabidopsis. Plant J **51:** 512-525

Zheng Z, Mosher SL, Fan B, Klessig DF, Chen Z (2007) Functional analysis of Arabidopsis WRKY25 transcription factor in plant defense against Pseudomonas syringae. BMC Plant Biol **7:** 2

Zheng Z, Qamar SA, Chen Z, Mengiste T (2006) Arabidopsis WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. Plant J **48:** 592-605

Zhou MQ, Shen C, Wu LH, Tang KX, Lin J (2011) CBF-dependent signaling pathway: a key responder to low temperature stress in plants. Crit Rev Biotechnol **31:** 186-192

Figure legends

Figure 1. Clustering of transcripts to pre-defined expression profiles generating the transcriptional response modes. For each stress combination, transcript sets were created by the union of the 500 most significant transcripts for each single stress and for the combination. These transcripts were clustered to 20 pre-defined expression profiles, each categorizing a potential expression change that may occur when multiple stresses are applied. Each transcript was assigned to the profile with the highest Pearson Correlation Coefficient. Left-side boxes: The 20 pre-defined expression profiles described by the transcript expression pattern for stress 1 (S1), stress 2 (S2) and the combination of stress 1 and stress 2 (S1+S2). The dotted line represents transcript expression with no change compared to the control. Right-side boxes: Each column represents the union of the two single and the combined stress for a given double stress experiment. In each box the value and color represents the percentage of transcripts that correlate with the particular pre-defined expression profile (green is higher). The transcriptional response modes are composed of a given set of the pre-defined expression profiles as indicated on the right: C*ombinatorial*, similar levels in the two individual stresses, but different response to combined stresses; *Cancelled*, transcript response to either or both individual stresses returned to control levels; *Prioritized*, opposing responses to the individual stresses and one stress response prioritized in response to combined stresses; *Independent*, response to only one single stress and similar response to combined stresses. *Similar*, similar responses to both individual stresses and to combined stresses. HL = High light, FLG = Flagellin.

Figure 2. Overview of mode of responses for the combined stress experiments, showing percentage of the transcript responses that cluster in each response mode (A) and per stress combination (B). *Combinatorial*, similar levels in the two individual stresses, but different response to combined stresses; *Cancelled*, transcript response to either or both individual stresses returned to control levels; *Prioritized*, opposing responses to the individual stresses and one stress response prioritized in response to combined stresses; *Independent*, response to only one single stress and similar response to combined stresses. *Similar*, similar responses to both individual stresses and to combined stresses. HL = High light, FLG = Flagellin.

Figure 3. Cumulative log-fold changes of the 500 most significantly responding transcripts in the single stress experiments when the particular stress is combined with another stress in a double stress experiment. The extent of response of significant transcripts upon combination with the other stress in a double stress experiment (eg. Cold and FLG) is given by the length of the bars from the center, where longer bars represent greater response of the transcripts. For example, when the plants are exposed to both heat and HL there is a higher response of the HL transcripts compared to the heat transcripts.

Figure 4. Relationships between four modules and the 11 stress treatments. The heat maps show transcript levels across treatments. Magenta is positive expression, black is neutral, and green is negative expression in comparison to the control treatment. Treatments are shown on the bottom as horizontal axis labels. Bar plots are eigengene values (i.e., the first principle component), calculated from singular value composition for each module.

Tables

Module	No. of transcripts	Functional enrichment analysis
1	66	No significant category detected
2	403	Response to organic substance, response to hormone stimulus, regulation of anion channel activity by blue light, response to abiotic stimulus, maltose metabolic process, response to chemical stimulus
3	328	Response to stress, nucleotide binding, transporter activity, hydrolase activity, electron transport or energy pathways
4	71	Response to abiotic or biotic stimulus, signal transduction, developmental processes, protein metabolism
5	60	Water transport, fluid transport
6	62	Response to abiotic stimulus, cellular response to red or far red light, circadian rhythm, response to radiation, shade avoidance, response to cold, response to hormone stimulus,
7	69	Response to cold, response to blue light, cold acclimation, auxin homeostasis, response to far red light, cellular response to carbohydrate stimulus, sugar mediated signaling pathway, response to non-ionic osmotic stress, response to abscisic acid stimulus, hyperosmotic salinity response, detection of gravity
8	907	Response to biotic stimulus, response to abiotic stimulus, multi-organism process, response to bacterium, response to heat, response to wounding, response to fungus, response to oxidative stress, response to light intensity, innate immune response, response to jasmonic acid stimulus, response to cold, indole glucosinolate metabolic process, flavonol metabolic process, host programmed cell death, response to hormone stimulus, salicylic acid metabolic process, response to ethylene stimulus, ost-translational protein modification, response to ozone, lignin metabolic process
9	166	Response to stress, electron transport or energy pathways, cell organization and biogenesis.
10	104	These transcripts were not placed in any of the modules.

Table 1. Co-expression module gene ontology associations. Number of transcripts and gene ontology functional enrichment of the 10 co-expression modules.

Figure 1. Clustering of transcripts to pre-defined expression profiles generating the transcriptional response modes. For each stress combination, transcript sets were created by the union of the 500 most significant transcripts for each single stress and for the combination. These transcripts were clustered to 20 pre-defined expression profiles, each categorizing a potential expression change that may occur when multiple stresses are applied. Each transcript was assigned to the profile with the highest Pearson Correlation Coefficient. Left-side boxes: The 20 pre-defined expression profiles described by the transcript expression pattern for stress 1 (S1), stress 2 (S2) and the combination of stress 1 and stress 2 (S1+S2). The dotted line represents transcript expression with no change compared to the control. Right-side boxes: Each column represents the union of the two single and the combined stress for a given double stress experiment. In each box the value and color represents the

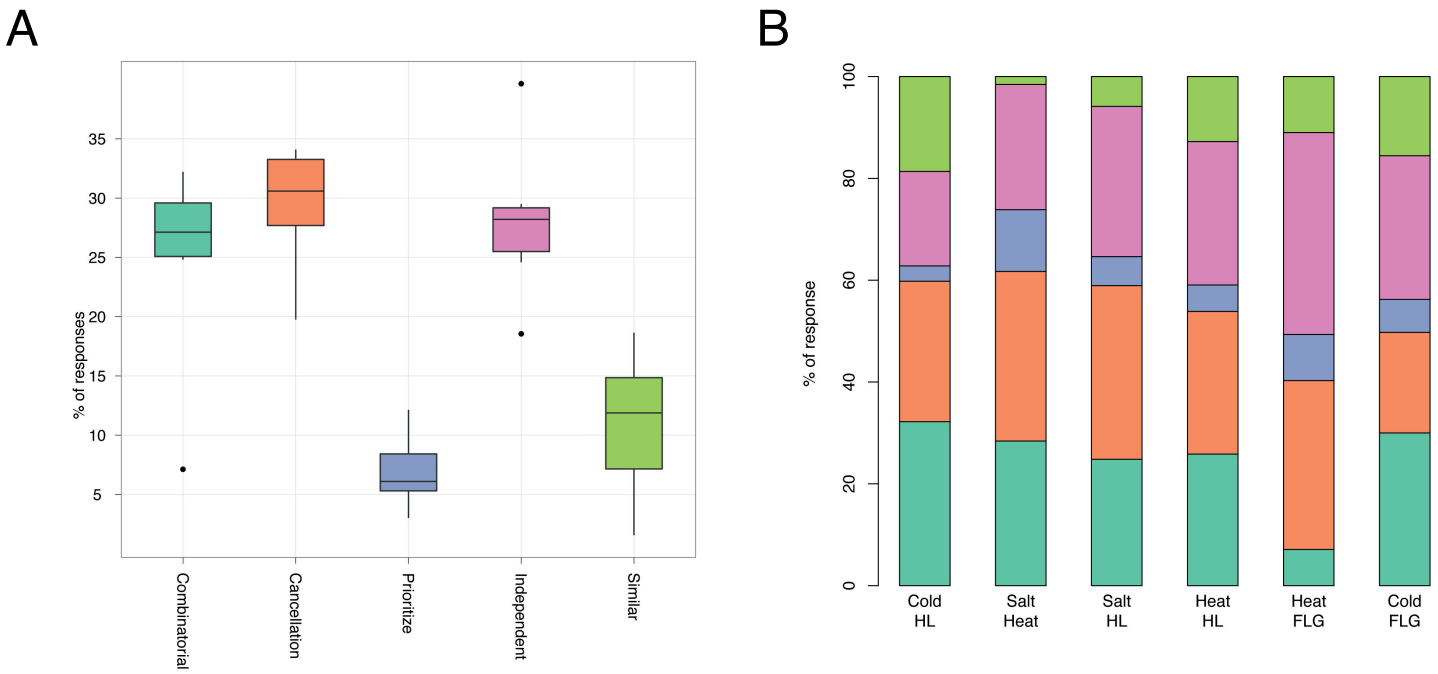


Figure 2. Overview of mode of responses for the combined stress experiments, showing percentage of the transcript responses that cluster in each response mode (A) and per stress combination (B). *Combinatorial*, similar levels in the two individual stresses, but different response to combined stresses; *Cancelled*, transcript response to either or both individual stresses returned to control levels; *Prioritized*, opposing responses to the individual stresses and one stress response prioritized in response to combined stresses; *Independent*, response to only one single stress and similar response to combined stresses. *Similar*, similar responses to both individual stresses and to combined stresses. HL = High light, FLG = Flagellin.

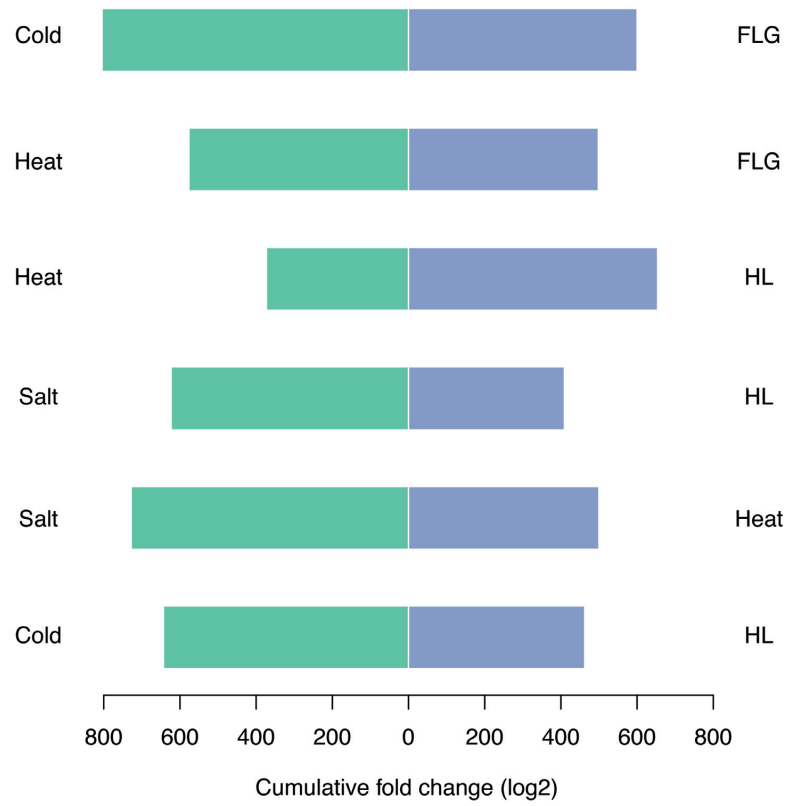


Figure 3. Cumulative log-fold changes of the 500 most significantly responding transcripts in the single stress experiments when the particular stress is combined with another stress in a double stress experiment. The extent of response of significant transcripts upon combination with the other stress in a double stress experiment (eg. Cold and FLG) is given by the length of the bars from the center, where longer bars represent greater response of the transcripts. For example, when the plants are exposed to both heat and HL there is a higher response of the HL transcripts compared to the heat transcripts.

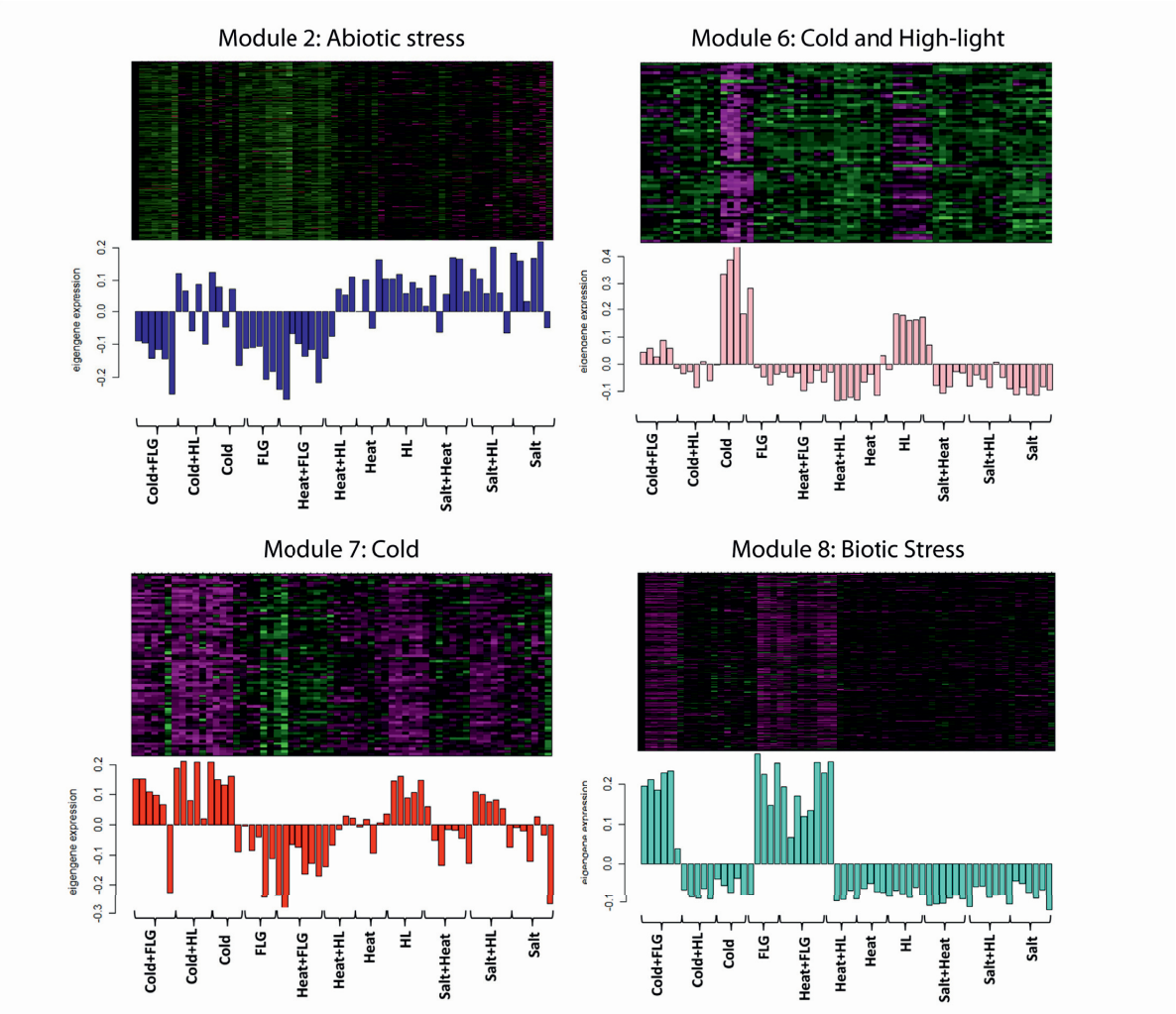


Figure 4. Relationships between four modules and the 11 stress treatments. The heat maps show transcript levels across treatments. Magenta is positive expression, black is neutral, and green is negative expression in comparison to the control treatment. Treatments are shown on the bottom as horizontal axis labels. Bar plots are eigengene values (i.e., the first principle component), calculated from singular value composition for each module.

Single abiotic stress experiments							
Ecotype	**Abbreviation**	**Ctrl**	**Cold**	**FLG**	**Heat**	**High_light**	**Salt**
Antwerpen	An-1	3	3	0	3	2	2
C24	C24	3	3	0	3	2	2
Columbia	Col	3	2	3	3	3	3
Cape Verde Islands	Cvi	2	3	0	2	2	3
Eriengsboda	Eri	3	2	0	3	3	3
Kashmir	Kas-1	3	3	0	3	3	3
Kondara	Kond	2	2	0	3	3	3
Kyoto-2	Kyo-2	3	3	0	3	3	3
Landsberg *erecta*	L*er*	2	3	3	2	3	3
Shakdara / Shahdara	Sha	3	3	0	3	3	3

Double abiotic stress experiments					
Ecotype	**Abbreviation**	**Cold+High_light**	**Heat+High_light**	**Salt+Heat**	**Salt+High_light**
Antwerpen	An-1	0	0	3	0
C24	C24	0	0	0	0
Columbia	Col	3	3	3	3
Cape Verde Islands	Cvi	0	0	3	0
Eriengsboda	Eri	0	0	0	0
Kashmir	Kas-1	0	0	2	0
Kondara	Kond	0	0	3	0
Kyoto-2	Kyo-2	0	0	0	0
Landsberg *erecta*	L*er*	3	1	3	3
Shakdara / Shahdara	Sha	0	0	3	0

Single and double biotic stress experiments					
Ecotype	**Abbreviation**	**Cold+High_light**	**Heat+High_light**	**Salt+Heat**	**Salt+High_light**
Antwerpen	An-1	0	0	3	0
C24	C24	0	0	0	0
Columbia	Col	3	3	3	3
Cape Verde Islands	Cvi	0	0	3	0
Eriengsboda	Eri	0	0	0	0
Kashmir	Kas-1	0	0	2	0
Kondara	Kond	0	0	3	0
Kyoto-2	Kyo-2	0	0	0	0
Landsberg *erecta*	L*er*	3	1	3	3
Shakdara / Shahdara	Sha	0	0	3	0

Supplemental Table S1. Overview of the experimental setup with regard to ecotypes, stress treatments and biological replicates.

Supplementary figures

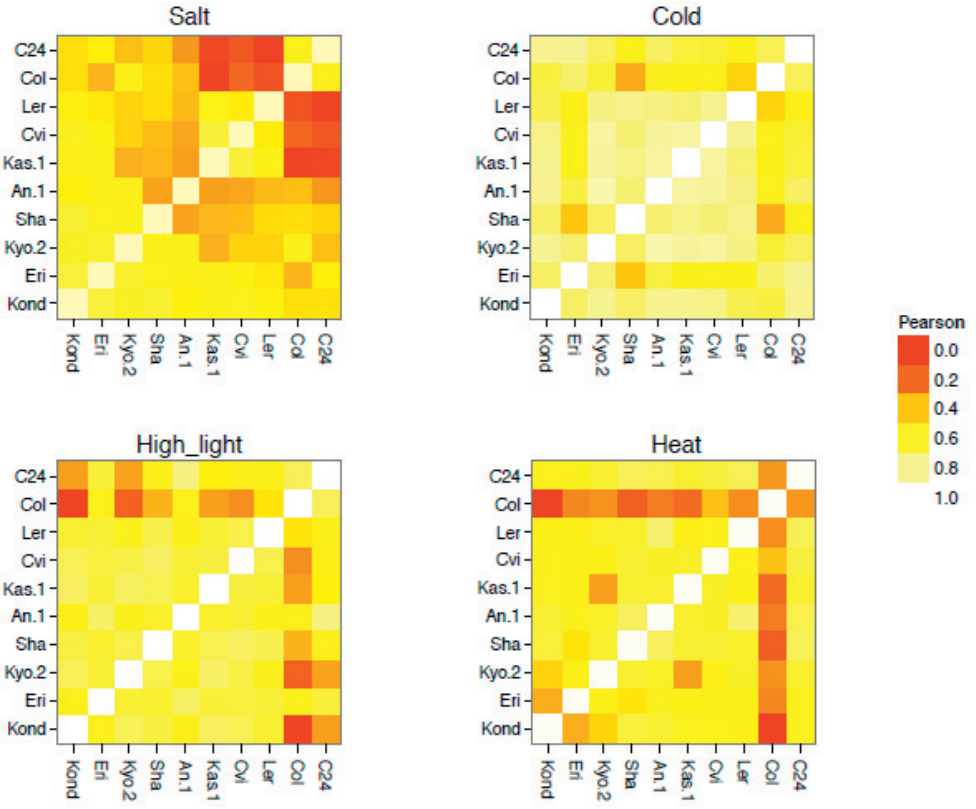


Figure S1. Ecotype response correlations to abiotic stresses. Pearson correlation coefficient of fold changes between the significant transcripts within each single stress experiments and ecotype. For each combination of stresses the correlations were calculated from fold changes of the 500 most significantly responding transcripts from both ecotypes.

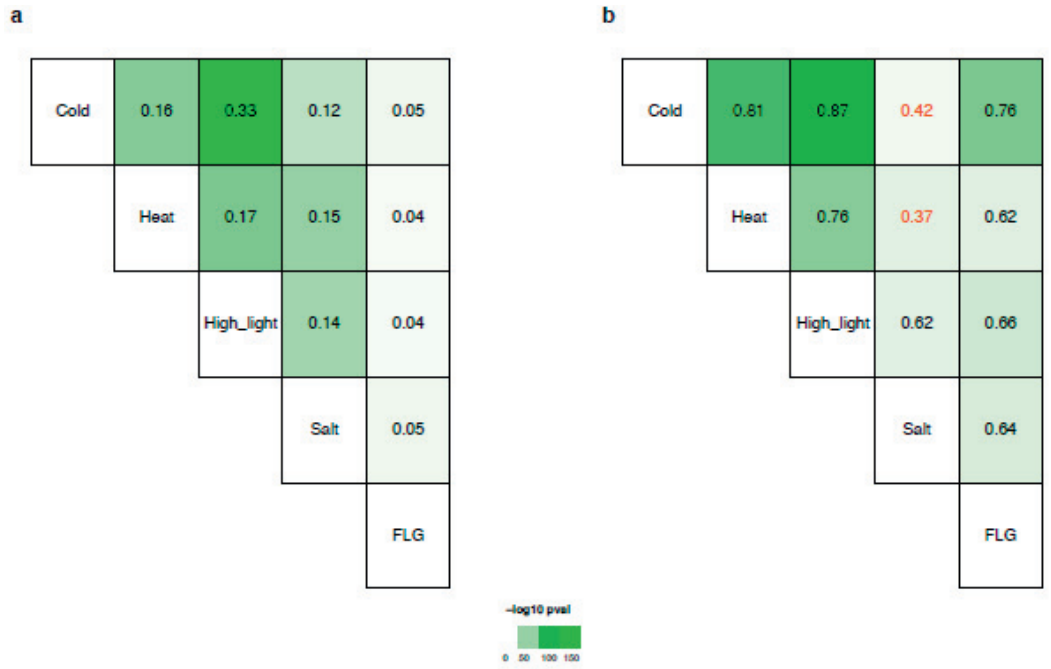


Figure S2. Overlap congruence between responses to individual stress treatments based on comparisons between top 500 transcripts of each single treatment. **a**. Transcript overlap colored by significance of Fisher's exact test. **b**. Congruency (value > 0.5) and dissimilarity (value < 0.5) colored by significance of binomial test, values in red are dissimilar. For both scales, green is more significant.

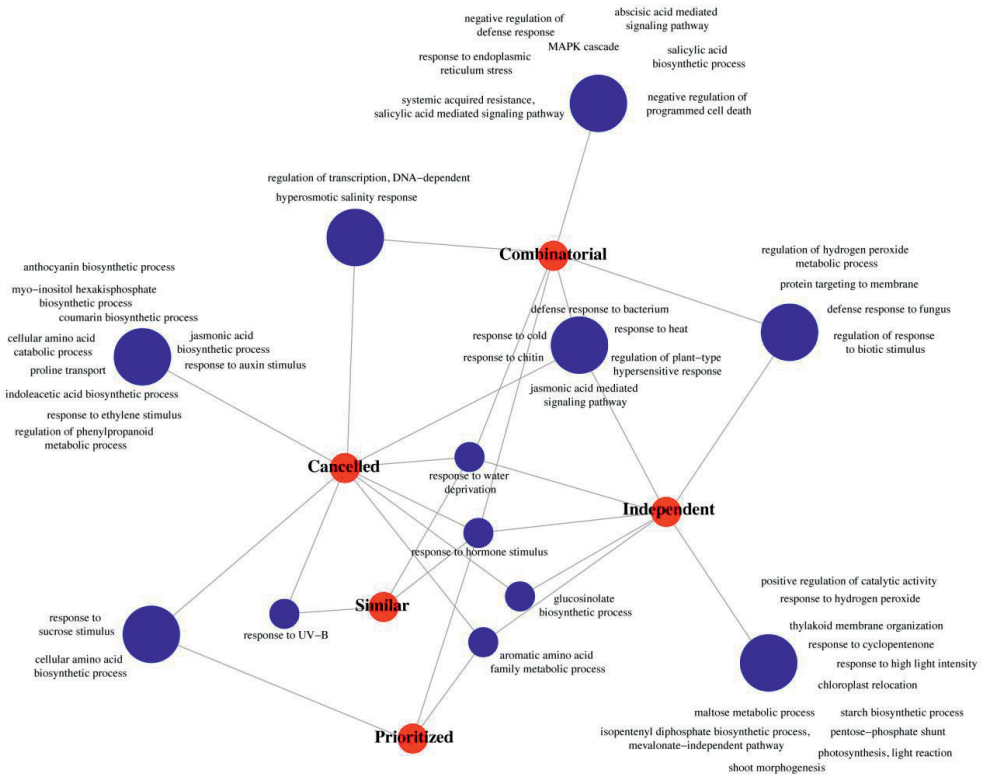


Figure S3. Network of transcript response modes and their associated GO terms. Transcripts from each of the response modes were pooled across experiments and over-representation to a particular response mode was assessed using Fisher's exact test. Red and blue circles represent the transcript response modes and GO terms, respectively. Small blue circles represent single GO terms whereas large blue circles represent several GO terms. Significance was set at p-value < 1e-5.

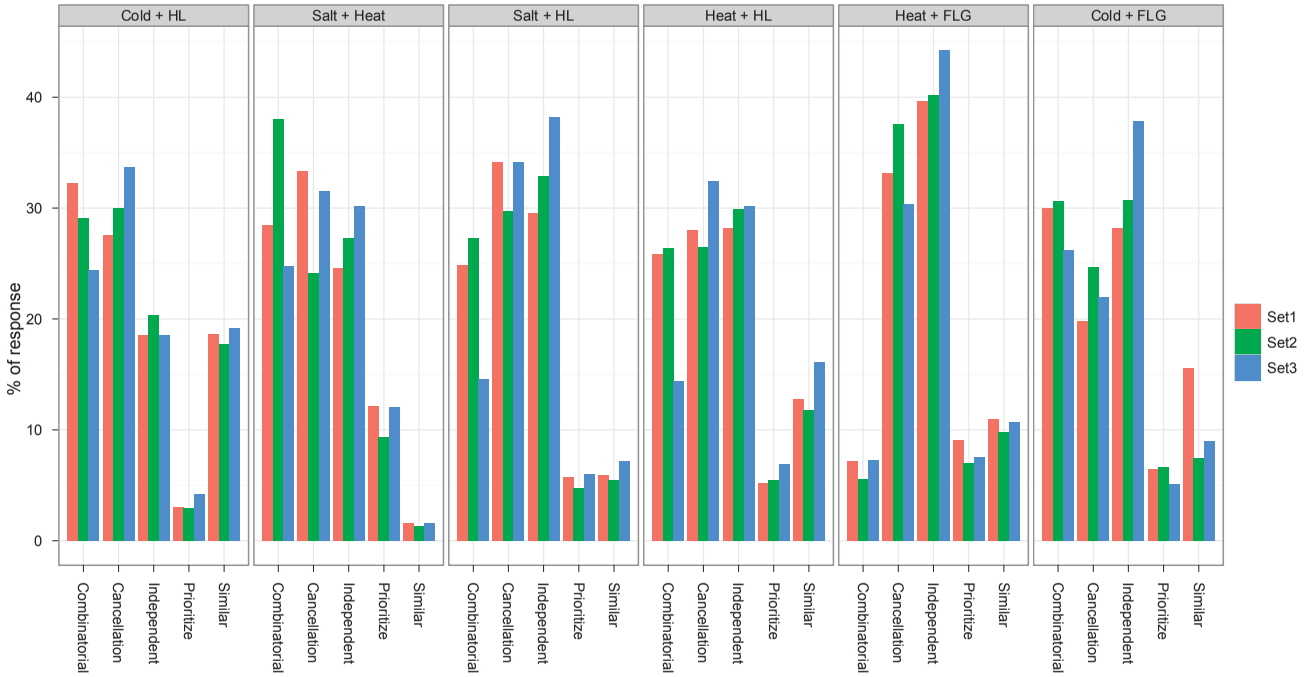


Figure S4. Changing cutoffs. The transcriptional responses investigated using transcript sets derived from different thresholds shows consistent profiles. Set1: 500 most significant transcript, Set2: 1000 most significant transcripts and then filtered for absolute log fold change > 0.5, Set3: all transcripts with p-value < 0.001.

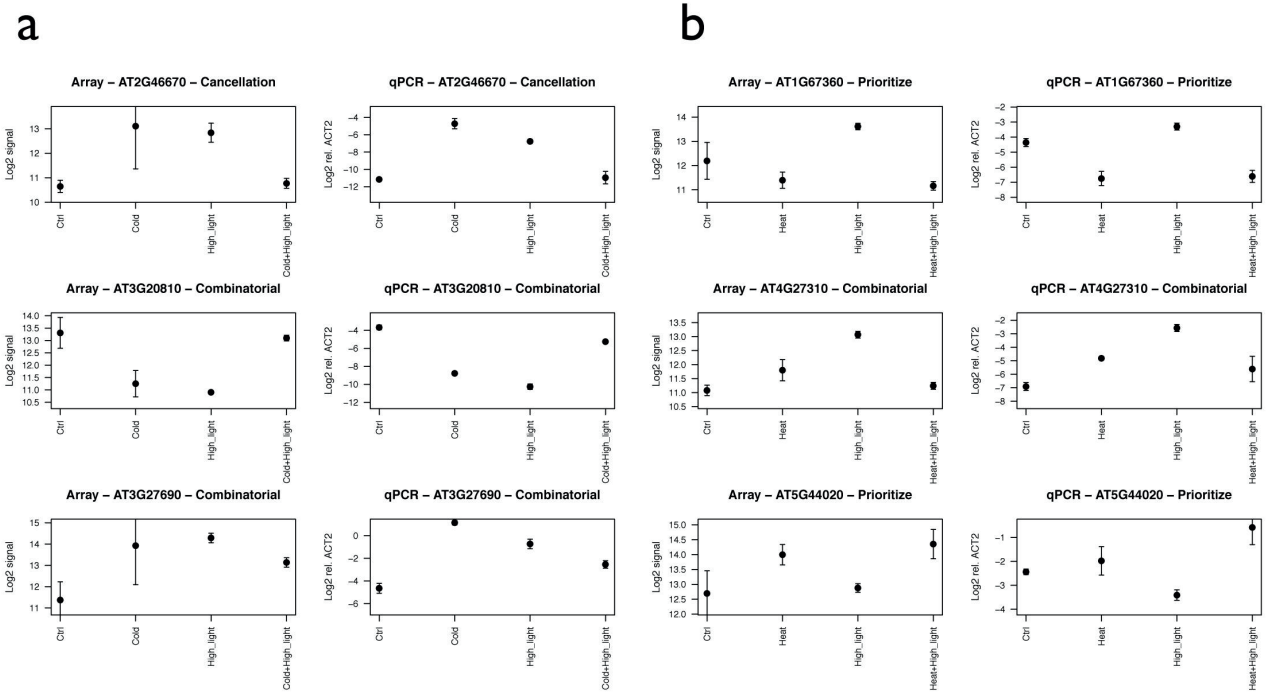


Figure S5. qPCR verification of six transcripts and their transcriptional response profiles using *Col* ecotype, three transcripts from the Cold+High light experiment (a) and three transcripts from the Heat+High light experiment (b). The error bars represents standard deviation based on three biological replicates, except for the Cold experiment with only two biological replicates for *Col*.

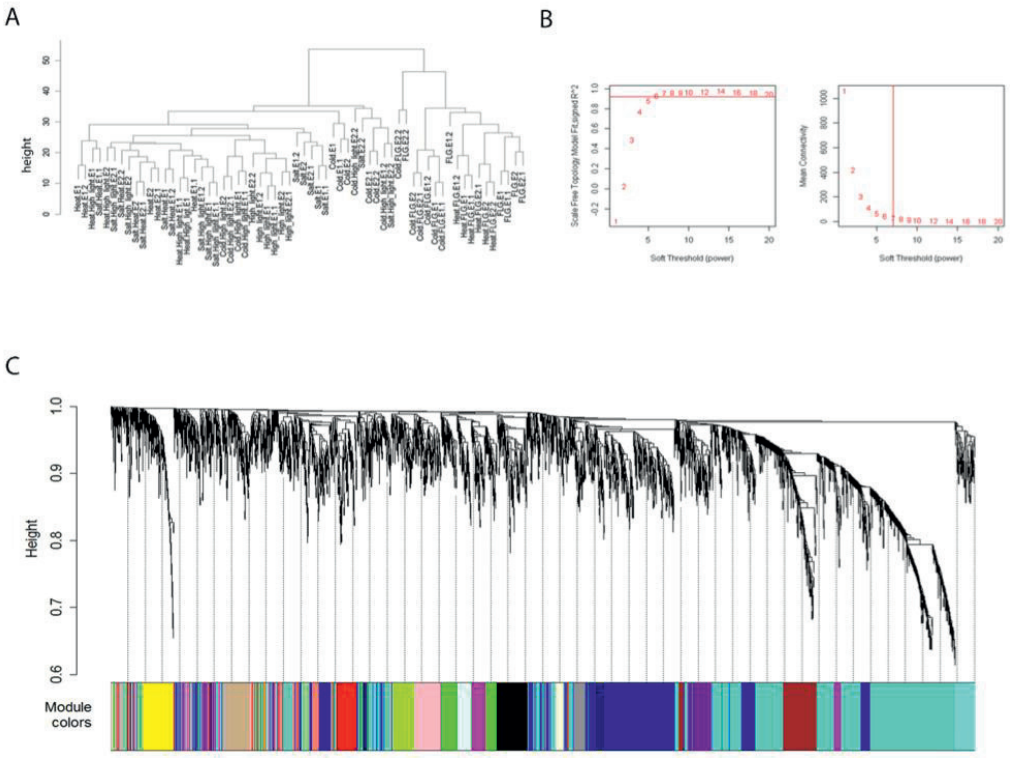


Figure S6. Weighted Gene Co-expression Network Analysis. A. Sample clustering to detect outliers in the data. B. Soft threshold of β=7 satisfied the scale-free topology approximation. C. Clustering dendrogram of transcripts, with dissimilarity based on topological overlap, together with assigned module colors later merged to the final 10 modules. In this dendrogram, each leaf is a short vertical line corresponding to a transcript, and branches of the dendrogram group together in densely interconnected modules of highly co-expressed transcripts.

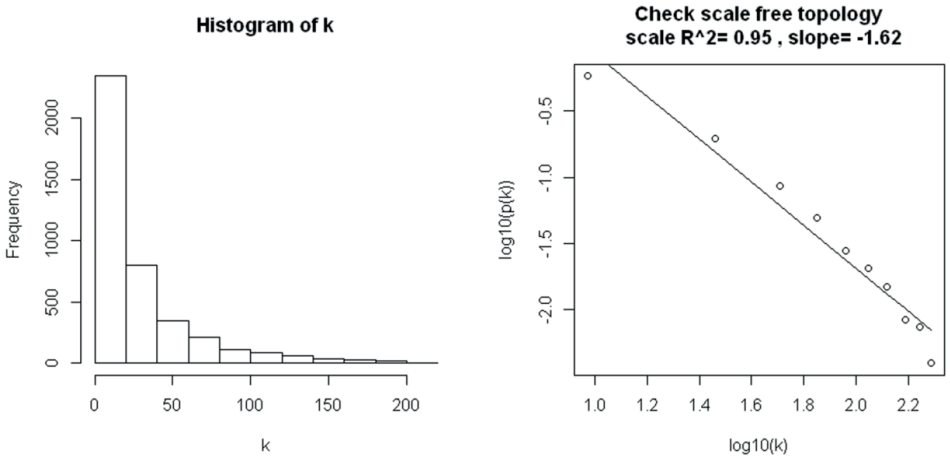


Figure S7. Scale-free behaviour of network node degree distribution. The left panel shows a histogram of network connectivities calculated for the constructed Weighted Gene Co-expression Network. The right panel shows a log-log plot of the same histogram. The approximate straight-line relationship (high R^2 value) shows approximate scale free topology. The constructed WGCN exhibits scale-free topology.

CHAPTER 5

(Paper IV)

Natural Variation in the Cold Response Regulatory Networks of Ten *Arabidopsis thaliana* Ecotypes

Pankaj Barah[1], Naresh Doni Jayavelu[2], Simon Rasmussen[3], Henrik Bjørn Nielsen[3], John Mundy[4] and Atle M Bones[1]*

[1] Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway N-7491

[2] Department of Chemical Engineering, Norwegian University of Science and Technology, Trondheim, Norway N-7491

[3] Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark DK 2800

[4] Department of Biology, University of Copenhagen, Copenhagen, Denmark DK-2200

*Corresponding author:
Atle M. Bones
Department of Biology
Norwegian University of Science and Technology
7491 Trondheim, Norway
atle.bones@bio.ntnu.no

Journal area to which this article is submitted (Original research: Full paper)

1) Environment: global change and environmental stress

2) Evolution: molecular evolution, population genetics

Summary

- Low temperature leads to major crop losses every year. Several studies have been conducted, focusing on diversity of cold tolerance level in multiple phenotypically divergent *Arabidopsis thaliana (A. thaliana)* ecotypes, but genome-scale understanding is still lacking.
- Here we report the genome-scale transcriptional responses to non-freezing cold stress treatments on ten *A. thaliana* ecotypes originated from different geographical locations. The ecotypes exhibited considerable variation in global transcript level responses to the cold stress treatment. A total of 6061 transcripts were identified to be significantly cold regulated ($p<0.01$), which included 498 transcription factors and 315 transposable elements. Ecotype specific transcripts and related gene ontology (GO) categories were identified.
- By using sequence information from the *A. thaliana 1001* genomes project, we further investigated sequence polymorphisms in the core cold stress regulon genes. Significant numbers of non-synonymous amino acid changes were observed in the coding region of the CBF (*C-repeat binding factor*) regulon genes.
- Considering the limited availability of experimental information on regulatory interactions in the model plant *A. thaliana*, we adopted a powerful systems genetics approach, Network Component Analysis (NCA), to re-construct an in-silico transcriptional regulatory network during response to cold stress using the gene expression data from ten ecotypes. Apart from retaining several previously benchmarked regulatory connections, the predicted network model identified new ecotype specific transcription factors and their regulatory interactions. These may be crucial for the geographic adaptation of the ecotypes to cold temperature.

Key words: *Arabidopsis*, ecotypes, Cold stress, natural variation, adaptation, gene expression, regulatory networks, *Arabidopsis 1001* genome, systems biology

Introduction

As sessile organisms, plants have evolved strategies to survive in unfavourable environmental conditions. Intraspecific variation in responses to environmental stress is well documented among plant species (Mitchell-Olds & Schmitt, 2006; Hall *et al.*, 2008; Alonso-Blanco *et al.*, 2009; Horton *et al.*, 2012). Understanding the molecular basis of such adaptations to environmental conditions has proven useful in selecting better traits or target genes for marker assisted breeding (Alonso-Blanco & Koornneef, 2000). Cold is a naturally occurring hazard to world crop production. Cold stress contributes to poor germination, stunted seedlings, chlorosis, reduced leaf expansion and wilting, and may also lead to tissue necrosis (Sanghera *et al.*, 2011). Exposure to cold stress may also drastically reduce the reproductive development of plants. It is thought that plants perceive cold by a receptor at the cell membrane, and that a signal is initiated to activate cold-responsive genes and transcription factors for mediating stress tolerance (Thomashow, 1999; Penfield, 2008). As reported earlier, the CBF pathway plays a major role in cold response, tolerance and acclimation. Nonetheless, there appear to be considerable differences in the sets of cold regulated genes described (Carvallo *et al.*, 2011). Just after few minutes of cold exposure, *CBF* genes are induced which encode a small family of transcription factors known as CBF1, CBF2, and CBF3 (also known as DREB1B, DREB1C and DREB1A). Cold induction of *CBF* genes regulates a set of about 100 downstream genes. Among them, the immediate target genes of CBF1-3 contain *CRT* (*C-repeat*)/*DRE* (*dehydration responsive element*) elements in their promoter regions to which the CBF1-3 proteins bind. The dehydration-responsive element (DRE) is also known as low temperature response element (LTRE) which contributes to cold responsiveness (Yamaguchishinozaki & Shinozaki, 1994). Interestingly, induction of the CBF regulon enhances both cold and drought tolerance (Jaglo-Ottosen *et al.*, 1998). Earlier transcriptome profiling studies have shown that multiple regulatory pathways are activated in

A. thaliana during cold exposure in addition to the CBF cold-response pathway (Fowler & Thomashow, 2002).

Natural variation for cold response and tolerance is an important element in the adaptation and geographic distribution of plant species. More generally, there is clear association between the plasticity of gene expression and the adaptability of an organism (Swindell *et al.*, 2007). Several studies have focused on the diversity of cold tolerance levels in divergent *A. thaliana* ecotypes (Rohde *et al.*, 2004; Alonso-Blanco *et al.*, 2005; McKhann *et al.*, 2008a; Zhen & Ungerer, 2008b; Zhen & Ungerer, 2008a). McKhann et al. reported that *CBF* and *COR* (Cold Regulated) genes responded differently to cold stress in eight accessions, though they could not find clear correlations between gene expression, sequence polymorphisms and freezing tolerance (McKhann *et al.*, 2008b). Thus, little is known about the molecular basis of the natural variation for freezing tolerance.

Transcriptional profiling in the model *Arabidopsis* is a major tool to identify plant genes regulated in response to changing environmental conditions (Somerville & Koornneef, 2002). However, variation in experimental conditions and protocols has made it difficult to extract and compare information from data sets produced in different laboratories (Moreau *et al.*, 2003). To overcome such problems, we subjected 10 ecotypes of *A. thaliana* were subjected to 5 individual stress treatments and 6 combinations of these stress treatments under the same experimental set up and profiling protocols (Rasmussen *et al.*, in press). We have considered all the cold experiments conducted on 10 ecotypes from this large dataset to explore genome-scale transcriptome response signatures of *A. thaliana* during cold stress treatment. By utilising data available from the recently published *A. thaliana* 1001 genome project, we further analysed sequence polymorphisms in CBF regulon genes (Austin *et al.*, 2011).

It is likely that differential expressions, or variation in mRNA stability caused by coding sequence polymorphisms, significantly contribute to natural variation in *A. thaliana* (Shimizu, 2002). Information about differentially regulated genes during different stress conditions is often available as an outcome of microarray experiments. However, in many cases, little is known about the regulation and interaction of these genes (Keurentjes *et al.*, 2007). Being highly dynamic in nature, any biological system continuously responds to environmental and genetic perturbations. Differential dynamic network mapping may facilitate the exploration of previously unknown interactions (Ideker & Krogan, 2012). While the *A. thaliana* genome has ~1922 transcription factors (TFs) (Guo *et al.*, 2005), experimentally confirmed regulatory relations are available for less than 100 TFs only (AGRIS database Sept. 10 2012) (Davuluri *et al.*, 2003). Tirosh & Barkai (2011) have explained how regulatory relationships can be deduced from patterns of evolutionary divergence in molecular properties such as gene expression. To this end, several computational algorithms have been developed to identify regulatory modules and their condition-specific regulators from gene expression data (Alter *et al.*, 2000; Segal *et al.*, 2003; Herrgard *et al.*, 2004). Network Component Analysis (NCA) is such an approach, which has been successfully implemented in several species, including *A. thaliana*, to determine both activities and regulatory influences for a set of transcription factors on target genes in various perspectives (Liao *et al.*, 2003; Kao *et al.*, 2004; Wang, JG *et al.*, 2011). By taking advantage of the NCA method, we have identified ecotype specific regulatory relationships, which provided new information towards understanding the natural variation in cold responses among different ecotypes of the model plant *A. thaliana.*

Materials and methods

Microarray data

We analyzed all the cold stress microarray experiments conducted on 10 ecotypes during the ERA-PG Multi-stress project, to explore genome-scale transcriptome response signatures of *A. thaliana* during cold stress treatment (GSE41935). All the experiments of were set up in environmentally controlled rooms at the plant growth facilities at RISØ DTU National Laboratory for Sustainable Energy (Roskilde, Denmark). A pilot study using wild type plants Col and L*er* was used to find the appropriate conditions at sub-lethal doses. These initial observations indicated that an optimal time before the onset of a phenotypic response (e.g: wilting, dehydration), while avoiding tissue damage, was 3 hours. Ten *A. thaliana* wild ecotypes **(Table1)** were then grown in soil under long day photoperiod and 24°C in a greenhouse setting for one generation to amplify homogeneous seed for all different genotypes. The seeds were then sown into trays and grown in a Conviron growth chamber (Winnipeg, Manitoba, Canada) under a 12hr/12hr photoperiod, 24°C and standard *A. thaliana* growth conditions. Three week-old plants were then placed for three hours in the environmentally controlled growth rooms that were preset to cold stress conditions (10°C). Triplicate trays of independently grown wild type controls were subject to the cold treatment. After the stress treatments, leaf tissue samples were collected and promptly frozen in liquid nitrogen for subsequent microarray experiments.

Statistical analysis of the data

Data from the microarray experiments was pre-processed using the RMA (Irizarry *et al.*, 2003) implementation in the oligo package (Carvalho & Irizarry, 2010) in R programming platform (R Core Team, 2012,). Gene annotation was acquired from TAIR10 (Lamesch *et al.*,

2012) using the biomaRt data mining tool (Guberman *et al.*, 2011). Differentially expressed genes between control and treated plants were identified using t-test ($p<0.01$). Genotype specific responses to stress were identified by the interaction effect from a two-way ANOVA (Kerr *et al.*, 2000; Cui & Churchill, 2003) of the genotype and treatment effect ($p<0.01$). The union of stress responsive genes were further used for network-based analysis. Heat maps were plotted using the TM4 microarray software suite (Saeed *et al.*, 2006).

Gene Set Enrichment Analysis (GSEA)

The Biological Networks Gene Ontology tool (Maere *et al.*, 2005), an open-source Java tool, was used to determine which Gene Ontology (GO) terms (Ashburner *et al.*, 2000) were significantly overrepresented in our differentially regulated gene lists (p-values were Bonferroni corrected).

Sequence analysis

Sequences for CBFs and COR genes were downloaded from the *A. thaliana* 1001 Genome project (http://signal.salk.edu/). Multiple sequence analysis were performed using Clustal w (Chenna *et al.*, 2003). Tajima's D (Tajima, 1989) statistical test to identify sequences which do not fit the neutral theory model at equilibrium between mutation and genetic drift were performed using MEGA5 suit (Tamura *et al.*, 2011).

Network component analysis and network reconstruction

Network component analysis is a computational method for reconstructing hidden regulatory signals or transcription factor activity (TFAs) from gene expression data with known connectivity information in terms of matrix decomposition (Liao *et al.*, 2003; Galbraith *et al.*, 2006). The NCA model assumes the log-linear relationship between target genes expression profiles and TFAs:

$$\frac{E_i(t)}{E_i(0)} = \prod_{j=1}^{L} \left(\frac{TFA_j(t)}{TFA_j(0)}\right)^{CS_{ij}} \quad (1)$$

where $E_i(t)$ and $E_i(0)$ are the expression values of gene i at different measurement conditions and reference point 0, and similarly $TFA_j(t)$ and $TFA_j(0)$ are the activities of TF_j, and CS_{ij} represents the control strength of TF *j* on gene *i*. Taking logarithms, the equation (1) becomes:

$$log[Er] = [CS]log[TFAr] \quad (2)$$

where the matrix *Er* represents the expression values of genes at different measurement conditions, matrix *CS* is the control strength of each TF on each TG, and *TFAr* represents the TFAs of all the TFs. The dimensions of [*Er*] are N ×M (N is the number of genes and M is the number of measurement conditions), [*CS*] is N × L (L is the number of TFs), and for [*TFAr*] is L × M. We can further simplify the above equation (2) as:

$$[E] = [C][T] \quad (3)$$

Here expression matrix [E] corresponds to [Er] in equation (2), connectivity strength matrix [C] is equivalent to [CS] and transcription factor activity matrix [T] corresponds to log[TFAr] in equation (2). Based on the above formulation, the decomposition of [E] into [C] and [T] can be achieved by minimizing the following objective function:

$$\min \| ([E] - [C][T]) \| \quad (4)$$

s.t. C ϵ Z0

Here Z0 is the initial connectivity pattern. The estimation of [C] and [T] is performed by using a two-step least-squares algorithm and normalized through a non-singular matrix [S] according to,

$$[E] = [C][T] = [C][S][S^{-1}][T] \quad (5)$$

In order to guarantee uniqueness of the solution for equation (4) up to a scaling factor, there are certain criteria to be satisfied, termed as NCA criteria: (a) The connectivity matrix [C] must have full-column rank. (b) When a node in the regulatory layer is removed along with

all of the output nodes connected to it, the resulting network must be characterized by a connectivity matrix that still has full-column rank. (c) T matrix must have full row rank.

The algorithm for NCA analysis is implemented in MATLAB by Liao and colleagues and is available online for download (http://www.ee.ucla.edu/~riccardo/NCA/nca.html). With NCA as a reconstruction method, we predicted significant TFs and connectivity strength on target genes and TFAs of TFs.

Results

Transcriptome signatures of 10 *Arabidopsis* ecotypes responding to cold stress

To cover a wide array of phenotypic variation, we selected ten natural accessions of *A. thaliana* representing habitats from 16° to 56.5° north latitudes. These were Cvi (Cape Verde Islands), Kas-1 (Kashmir, India), Kyo-2 (Kyoto, Japan), Sha (Shakdara, Tadjikistan), Col (Columbia, USA), Kond (Kondara, Tadjikistan), C24 (Coimbra, Portugal), L*er* (Landsberg, Poland), An-1 (Antwerpen, Belgium), Eri (Erigsboda, Sweden)**(**details in **Table1).** We have chosen a cut-off of $p \leq 0.01$ to define a gene as differentially expressed. Using the results from these ten ecotypes, we were able to examine the large differences in relative transcript levels that occurred during early hours of cold treatment. The results **(Table 1)** indicated that *A. thaliana* ecotypes have visibly different transcriptome signatures in response to cold treatment. Variable numbers of transcripts were up or down regulated by cold stress. Considering the two extreme ecotypes, Col-0 as a cold tolerant ecotype had significantly less numbers of differentially regulated transcripts, while Cvi, the southernmost ecotype among the 10 used in our experiments, had the highest number of differentially regulated transcripts. Ecotype Cvi (Cape Verde Islands) was associated with a climate temperature comparatively higher than that of ecotype Col-0 (Columbia), which was well reflected in reduced transcript responses to cold treatment. Similar results were also reported by Swindell *et al.* (2007).

A unified list of 6061 cold regulated transcripts ($p<0.01$) were generated from the 10 ecotypes **(Supporting file S1).** Total 498 TFs were differentially regulated in this list. Interestingly, 4553 transcripts (75%) were differentially regulated only in one of the ten ecotypes. The significant list includes most of the previously reported key cold regulated genes. **Fig. 1** displays fold change values (treatment vs. control) calculated from the normalized expression index for the top 1000 genes from the 10 different ecotypes. As a global observation, the heat map indicates differentially regulated transcriptome signatures in response to non-freezing cold treatment in ten different *A. thaliana* ecotypes. Hierarchical clustering (HCL) was performed with Pearson correlation as distance measure and using the average linkage method and 10,000 bootstrapping for the top 1000 cold regulated transcripts based on fold change ratios with respects to their respective controls. *Columbia* (Col-0) is distinctly separated from the other ecotypes, while the southern ecotypes Cvi, Sha and Kyo2 are grouped closely. Zhen et al. (2008) reported a positive correlation between freezing tolerance and latitude of origin based on physiological data collected from 71 *A. thaliana* ecotypes (Zhen & Ungerer, 2008a). Hannah et al. (Hannah *et al.*, 2006) used nine accessions of *A. thaliana* to show that freezing tolerance of natural accessions correlates with habitat winter temperatures. However, our clustering of the gene expression pattern in response to non-freezing cold stress exposure in the 10 ecotypes does not detect a clear latitudinal trend. This may be due to the limited number of ecotypes and a single time point used in our transcriptome experiments, and to the fact that we did not 'freeze' the plants.

Ecotype specific cold regulated transcript lists contain many transcription factors (TFs) and transposable elements (TEs)

In contrast to the relatively small number of transcripts with altered expression shared by all 10 ecotypes, the majority of the transcripts (75%) showed ecotype specific expression pattern (**Supporting file S2**). Thus, each of the ecotypes had unique sets of differentially regulated transcripts in response to cold stress. From the list of differentially regulated transcripts, it was found that 498 encoded for Arabidopsis TFs and 323 TFs were differentially regulated in single ecotypes. The list of differentially regulated transcripts includes many well-known cold regulators like *CBF*s, *DREB1A, DREB1B, DREB2B, RAV1, ERF2,* and *ERF5.* We have surveyed existing available transcription factor - target gene (TF-TG) regulatory interactions available in public databases or literature. In the GO (Gene Ontology) database and TAIR (The Arabidopsis Information Resources), there were 59 TFs reported as associated to cold responses. There were 320 TFs (~ 64% of all the differentially cold regulated TFs) which were regulated only in one of the ten ecotypes. The ecotype specific differentially cold regulated TFs are listed in **Table 2**.

Nimblgen12-plex Arabidopsis microarray chip included 3822 Transposable Element (TEs) probes. In the ecotype specific differentially regulated transcript list, we observed 315 TEs (~10% of the total TE probes printed on the chip). The distribution of the differentially regulated TEs in ten ecotypes were as follows – Col (21), L*er* (81), Cvi (71), Eri (31), Kas-2 (16), Kond (39), Kyo2 (23), C24 (15), Sha (22) and An-1 (8). Somatic events, in particular the activity of transposable elements (TEs), play an important role in plant genome evolution (Ziolkowski *et al.*, 2009). Lee et al. reported that cold-regulated gene expression was not only controlled transcriptionally, but might also be regulated at the posttranscriptional and chromatin levels (Lee *et al.*, 2005). A change in the epigenetic state of TEs by cold stress may

also contribute to the regulatory activities of adjacent genes (Kashkush *et al.*, 2002). Some of the differentially regulated TEs in our cold experiments might therefore be important targets to study diversity of cold stress responses among the ten different ecotypes. Further targeted experiments in this direction can explore the molecular level details of any potential role of these TEs on genomic adaptation of the ecotypes to their local environment.

Gene Set Enrichment Analysis indicates activation of common and unique processes in different ecotypes

To investigate functionally relevant changes, gene ontology based overrepresentation analysis was performed using BinGO software considering the up-regulated gene lists from each of the ten ecotypes (**Supporting file S3**). From this analysis, we created a contingency table by uniting all the statistically significant overrepresented GO categories from each of the ten ecotypes (**Supporting file S4**). Genes showing significant variation in mRNA levels in *A. thaliana* during different stress conditions mainly belong to categories like signal transduction, transcription and stress response (Chen *et al.*, 2005). This reflects the fact that variations among different ecotypes exist in the regulatory mechanisms for these genes. Apart from common cold stress responsive categories such as response to cold stress, response to low temperature, cold acclimation, we observed a few other biological processes to be differentially up-regulated in the various ecotypes **(Table 3).** Some of the interesting and top GO categories are noted below.

Cold response is coupled with light stimulus

Along with the general cold response pathways or processes, there were several overrepresented categories related to 'response to light'. A few genes in these categories were: At1g29920 (*LHCB1*), At5g24470 (*PRR5*), At4g08920 (*OOP2*), At1g02340 (*RSF1*), At1g06040 (*STO*), At3g27690 (*LHCB2.4*), At3g54720 (*PT*), At2g42540 (*COR15A*), At2g26990 (*FUS12*), At5g24120 (*SIGE*), At2g46970 (*PIL1*), At4g18130 (*PHYE*), At5g67030 (*ZEP*), At5g45340 (*CYP707A3*), At1g02400 (*GA2OX6*), At2g46790 (*TL1*), At3g28860 (*PGP19*) At2g46340 (*SPA1*), At4g19230 (*CYP707A1*), and At2g18790 (*PHYB*). Light and cold signals are known to integrate in cold tolerance via a CBF and ABA-independent pathway (Catala *et al.*, 2011). Franklin et al. investigated the modulation of low R/FR signalling by ambient temperature and showed that a low red to far-red ratio (R/FR) light signal increases CBF gene expression in *A. thaliana* in a manner dependent on the circadian clock. Red or far-red light signalling pathway is one of the significantly up-regulated GO categories in some of the ecotype (Franklin & Whitelam, 2007). Such signals stimulate expression of *CBF* genes through ambient temperature–dependent coupling of CBF transcription factors to downstream *COLD REGULATED* (*COR*) genes.

Chlorophyll biosynthetic process

Another overrepresented category from the GO analysis was 'chlorophyll biosynthetic process'. This included several genes including At1g78600 (*STH*3), At5g54190 (*PORA*), At3g59400 (*GUN4*), At3g56940 (*CRD1*), At4g34740 (*CIA1*), At1g78600 (*STH3*), At1g71030 (*MYBL2*), and At5g67030 (*ZEP*). Harvaux et al. reported that *A. thaliana* was able to survive cold stress through light independent xanthophyll cycling by illustrating protective functions of carotenoid and flavonoid pigments against excess visible radiation at cold temperature (Havaux & Kloppstech, 2001). Cold stress also induces synthesis of early light-induced proteins (ELIPs) (Heddad *et al.*, 2006). Low temperature also induces the accumulation of

various antioxidants including carotenoids (except β-carotene), vitamin E (α- and γ-tocopherol) and non-photosynthetic pigments (anthocyanins and other flavonoids) (Rapacz *et al.*, 2008). Genes in the overrepresented category "pigment biosynthetic process" from our analysis support the previous reports.

Cold stress and circadian rhythms

Circadian rhythm is one of the most prominent overrepresented categories in our dataset. It included many well-known genes including At1g22770 (*GI*), At1g68050 (*FKF1*), At1g18330 (*RVE7*), At5g24470 (*PRR5*) (Zuther *et al.*, 2012), At5g17300 (*RVE1*), At2g46790 (*TL1*), and At2g46830 (*CCA1*). Previous studies reported circadian clock regulated induction of *CBF* genes during low-temperature treatment (Harmer *et al.*, 2000; Edwards *et al.*, 2006). The circadian clock gates both gene expression and physiological responses to low R/FR during rapid shade avoidance (Salter *et al.*, 2003; Fowler *et al.*, 2005). Mikkelsen and Thomashow (2009) reported that cold- and clock-regulated gene expression are integrated through regulatory proteins that bind to Evening Element (EE) and Evening Element Like (EEL) elements supported by transcription factors acting at ABA response elements (ABRE) sequences (Mikkelsen & Thomashow, 2009). Our current results are in good agreement with these previous reports.

Co-regulation of cold and biotic stress responsive genes

Few categories in our gene set enrichment analysis (GSEA) were related to biotic stress responses. Among these few categories were response to other organism, response to fungus, response to bacterium, and multi-organism process. Some of the up-regulated genes in these categories include At2g40140 (*ZFAR1*), At5g25110 (*CIPK25*), At5g25910 (*RLP52*), At1g20440 (*RD17/COR47*), At4g37150 (*MES9*), At3g50970 (*XERO2*), At2g42530

(*COR15B*), At2g44490 (*PEN2*), At5g64750 (*ABR1*), At1g51090, At4g12470, At4g36010 (*pathogenesis-related thaumatin family protein*), At3g51660 (*MIF family protein*), At1g20030 (*pathogenesis-related thaumatin family protein*), At3g50260 (*CEJ1*), At3g15210 (*RAP2.5*), At5g58600 (*PMR5*), At3g52400 (*SYP122*), At3g06490 (*MYB108*), At1g19180 (*TIFY10A*), At4g23810 (*WRKY53*). **Supporting file S4** contains all the GO categories from each ecotype including the ecotype specific categories. One important observation was that biotic stress response related categories- response to other organism, response to fungus, response to bacterium, and response nematode were overrepresented mainly in the southern ecotypes such as Cvi, Kas1, Kyo2, and Kond. A possible reason may be that plants in southern latitudes often face such biotic invaders compared to their northern counterparts, and consequently have co-evolved with better and prompt defence responses against them. Based on genetic resources of *A. thaliana*, coupled with 39 years of field data, it has been reported that natural enemies drive geographic variation in plant defences (Zust *et al.*, 2012).

Unlike cold tolerance, molecular mechanism of pathogen resistance obtained through cold treatment is not understood well. Plazek et al. reported that cold treatment of spring barley (*Hordeum vulgare.*), meadow fescue *(Festuca pratensis*) and oilseed winter rape (*Brassica napus var. oleifera*) induced resistance to their specific pathogens (Plazek & Zur, 2003). Zhu et al. identified a temperature sensitive component in disease resistance, and provided a potential means to generate plants adapted to broader temperature ranges (Zhu *et al.*, 2010). Besides the available reports about enhanced disease resistance acquired through cold treatment, it is not yet known if these two traits are regulated by the same signal transduction pathways (Kuwabara & Imai, 2009). We have observed functional overrepresentation of GO categories like steroid hormone mediated signalling pathway, brassinosteroid mediated signalling pathway, and jasmonic acid stimulus. Phytohormones are involved in induction of disease resistance upon pathogen infection. Plant hormones like

salicylic acid (SA) ethylene (ET), jasmonic acid (JA) are known to play important functions in signal transduction during biotic stresses (Fujita *et al.*, 2006). The occurrence of simultaneous biotic and abiotic stresses increases the complexity, as the responses to these are largely controlled by different hormone signalling pathways that may interact and inhibit one another (Atkinson & Urwin, 2012). Interaction of cold temperature and pathogen attack may potentially negatively impact plant growth (Mittler, 2006). Plants grow in heavy snowfall areas need to enhance disease resistance to survive from the attack of pathogens such as snow molds (Hoshino *et al.*, 2009). Hence, as a nascent observation, the co-evolution of regulatory mechanism for co-occurring stress related genes and processes are probable. Further targeted screening of more ecotypes may explore interactions of biotic and abiotic stress on adaptive evolution of plant defence response.

CBF regulon genes exhibit differential expression pattern in *Arabidopsis ecotypes* during cold treatment

The *A. thaliana* CBF cold response pathway has a major role in cold responses. *CBF* genes appear to be present across plant species and are almost always present as a gene family. In *A. thaliana* there are four characterized *CBF* genes: *CBF1, 2* and *3*, located on chromosome 4, are cold induced while *CBF4*, located on chromosome 5, is reported to be involved in drought tolerance (Haake *et al.*, 2002; Medina *et al.*, 2011). All the *CBF* genes as well as the selected *COR* genes were cold regulated in the 10 accessions. However, we observed different levels of expression of *CBF* and *COR* genes in them. All the *CBF* genes were induced, but *COR* genes had preferential expressions in different ecotypes (**Fig. 2**). *DREBA1* expression was consistent in all the accessions. In a previously published study, *CBF1* and *CBF2* were reported to have quite comparable expression levels in 9 accessions except low expression of

both in Cvi (McKhann *et al.*, 2008). Low expression of *CBF1* and *2* in the Cvi ecotype is clearly visible in our data (**Fig. 2**). It was reported that expression of the CBF1, 2 and 3 genes was not correlated with cold tolerance level among ecotypes (Zuther *et al.*, 2012).

Variation in gene expression reflects the interplay between 'robustness' and 'evolvability' and is generally achieved by regulatory divergence. An organism has to preserve a consistent function under different conditions and, at the same time, it needs to sustain the ability to evolve in order to adapt to new environments. The plasticity of gene expression may be achieved by selective accumulation of mutations in the promoter. As about 100 downstream genes and processes are regulated by the CBF and COR proteins, differences could be seen in the expression pattern of down-stream genes, which was visible in the heat-map of 1000 genes and other ecotype specific, differentially regulated genes (**Fig. 1**). We chose to investigate the polymorphism present in the *CBF1, 2* and 3 and a few *COR* genes using recently released data from *A. thaliana* 1001 genome project (Cao *et al.*, 2011).

Sequence Polymorphisms in the *CBF* genes

Sequence variation of *CBF* and *COR* genes could exert effects at two different levels: (i) in the expression of the *CBF* genes themselves, via polymorphism in the respective promoters, or (ii) in the expression of their downstream genes. It has been reported that all three *CBF* genes were highly polymorphic, particularly in their promoters, with *CBF1* the most and *CBF2* the least polymorphic (McKhann *et al.*, 2008b; Zhen & Ungerer, 2008b). Hence, we downloaded the sequence data (including 1kbp upstream of the coding region) available from 1001 genome database, and calculated Tajima's D statistic to evaluate the allele frequency spectrum and to quantify the frequency of rare alleles. We observed significant numbers of non-synonymous amino acid changes in the coding region of the *CBF* genes **(Supporting file**

S5). Initially, sequences from all available ecotypes in the 1001 genome database (706) were downloaded, but incomplete sequences were discarded before the analysis. Apart from the coding regions, we considered 1,000 bp upstream sequences for alignment. We have considered codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. The three *CBF*s have shown significantly negative Tajima's D values (*CBF1*= -1.291, *CBF2*= -2.223, *DREBA1*= -2.165). More negative and significantly lower values of Tajima's D indicate an increased frequency of rare and recent alleles. However, it is known from earlier studies that the average distribution of Tajima's D in the *A. thaliana* genome is biased towards negative values. We could not conclude any direct correlation between sequence polymorphisms on the expression patterns of the *CBF* and *COR* genes. The lack of a clear correlation might have several reasons. There are other CBF independent pathways, and their complex interactions between different components contributes to cold tolerance (Fowler & Thomashow, 2002). So, how these complex interactions of other pathways affect *CBF* and *COR* gene expression would be difficult to predict. Again, *COR* genes are often up-regulated much later, and this is also true for protein expression. So, they might not be clearly visible while looking at single time point transcriptomic response during non-freezing cold treatment. But, apart from genotype variation, the length of cold exposure and treatment temperature also affect the gene expression level that leads to freezing tolerance (Fowler & Limin, 2004). While studying natural variation of transcriptional auxin response networks in *A. thaliana*, Delker et al. reported that differentially regulated signalling networks had a greater role to play than sequence polymorphism (Delker *et al.*, 2010). Considering such facts, we wanted to explore the pattern of regulatory divergence of cold stress response network among the ten *A. thaliana* ecotypes.

Re-construction of a transcriptional regulatory network during cold stress responses

Due to the lack of experimentally validated transcriptional regulatory information in *A. thaliana*, we decided to re-construct an *in silico* transcript regulatory network model during cellular responses to cold stress using gene expression data from ten natural ecotypes. For this purpose, we selected the top 1,509 differentially cold regulated transcripts from the union of the entire cold regulated transcripts list, given the criterion that a transcript had to be significantly differentially regulated in at least 2 of the ecotypes to be included in the network re-construction. This final list contained 178 TFs and 1,331 target genes (TGs). Using the NCA method (materials and methods) we re-constructed the network by putting correlation-coefficient threshold ≥ 0.8. Activation and repression interactions were calculated using the positive and negative correlations. The final network contained 1,275 nodes and 7,720 connections and, of them, 6,731 connections were activations (positive) and 989 were repressions (negative) (**Supporting file S6** and **Fig. 3a**). Some of the highly connected positive regulators (TFs) include *ATTLP7, POSF21, AS1, RTV1, APRR9, BT1, ANAC102, ANAC035, GLK2, ZFN1, WRKY11, HAC5, MYB73, DA1, LBD41, SR1,* and *WRKY70*. Further details of the network including calculated Pearson's correlation coefficient, *p-values* etc. are given in **Supporting file S6**. In the network visualization, transcription factors were marked as green triangles and target genes as red circles. General network topology analysis revealed that the network exhibited power-law degree distribution (Clauset *et al.*, 2009) (**Fig. 3b)**. We also calculated several other graph-theoretic parameters (Barabasi & Oltvai, 2004). Some of the parameters were: clustering coefficient = 0.3, connected components = 3, network diameter = 11, characteristic path length = 3.67, and average number of neighbours = 11.385. The generated network (.cys file) isprovided in Supporting file **S7**. The file can be opened locally and explored interactively using Cytoscape software (Shannon *et al.*, 2003). Note that

the presented view of the annotated network in this manuscript has been simplified manually for easier representation.

Transcription factor activity under cold stress in different *Arabidopsis* ecotypes

The primary assumption in predicting interactions between TFs and their target genes using gene expression data is that mRNA expression levels reflect the activities of the corresponding proteins. However, expression levels between mRNAs and their corresponding proteins in different cell types show diverse correlations for different genes (Pascal *et al.*, 2008). Nonetheless, several studies have reported an overall positive correlation between mRNA and protein expression levels (Tian *et al.*, 2004; Guo *et al.*, 2008). Based on this assumption, Wang et al. (2011) successfully implemented the NCA methodology to construct a dynamic transcript network during *A. thaliana* pollen development (Wang, JG *et al.*, 2011). We adopted the same strategy to explore the differential activity of the previously reported cold regulated transcription factors during non-freezing cold stress treatment among ten ecotypes.

The NCA requires two inputs: a series of gene expression profiles and a pre-defined regulatory network. The *A. thaliana* transcription factor list was collected from the Database of Arabidopsis Transcription Factors (DATF) (Guo *et al.*, 2005), The Arabidopsis Gene Regulatory Information Server (AGRIS) (Yilmaz *et al.*, 2011), and the Plant Transcription Factor Database (PlantTFDB)(Riano-Pachon *et al.*, 2007). A list of 59 cold regulated transcription factors was collected from Gene Ontology database under the annotation category 'response to cold' (Ashburner *et al.*, 2000). The re-constructed network accounted for 30 (~50%) of these previously reported cold regulated TFs.

Correlation between activity and expression values of TFs

We compared the predicted transcription factors activities in ten different ecotypes with their gene expression data for the 30 previously reported cold responsive transcription factors. About 57% of the TFs showed moderate correlation (Pearson correlation coefficient, $|r| > 0.5$) between their activities and expression levels (**Fig. 4**). Thirteen TFs ((*ZFAR1* (At2G40140), At1G28050, *ERF2* (At5G47220), *ZAT10* (At1G27730), At5G46710, *CRF4* (At4G27950), At1G78700, At5G17300, At4G28140, At5G48250, At4G29190, *RAP2* (At1G46768), and *WRKY7* (At4G24240)) exhibited positive correlations ($r > 0.5$). For example, CZF-1 (At2G40140) had a correlation of r=0.9206 suggesting an autoregulation in agreement with previous reporting (AbuQamar *et al.*, 2006) and information available in the Arabidopsis Gene Regulatory Information Server (AGRIS). Four transcription factors ((*WRKY25* (At2G30250), *ERF6* (At4G17490), *DREB2B* (At3G11020) and *TIFY10A* (At1G19180)) showed strong negative correlation ($r < -0.5$) and the remaining TFs displayed low or no correlation at all ($|r| < 0.5$). Three of these predictions have been confirmed from the AGRIS database by downloading all the available interactions and comparing them against our predicted network.

Transcription factor activities in ten different ecotypes

The expression activity profiles of thirty cold regulated TFs have clearly shown the ecotype specific variations in ten of the *A. thaliana* ecotypes (**Fig. 4**). For example, At5G17300 (*RVE1*) is highly active in the Eri, C24, Col ecotypes compared to others. Most of the transcription factors are multiply responsive (active in more than two ecotypes) to cold stress treatment (**Table 4**). We could also identify ecotype specific transcription factors (highly active in only one ecotype). For example, At1G04240 (*SHY2*), At2G46830 (*CCA1*), At3G11020 (*DREB2B*) are active in the Sha ecotype, and At4G25490 (*DREB1B/CBF1*) is

most active in the Eri ecotype. Spatio-temporal regulatory dynamics of *SHY2* (Scacchi *et al.*, 2010), *CCA1* (Li *et al.*, 2011; Wang, W *et al.*, 2011) have been reported earlier. The transcription factor At5G17490 (*RGL3*) (Feng *et al.*, 2008) is active in the Col ecotype. We also found a group of transcription factors which were highly active in a particular set of ecotypes. For example, At1G27730 (*ZAT10*), At1G28050, At3G17609 (*HYH*), AtAt4G27950 (*CRF4*), At5G17300, and At5G48250 were active in the Eri, C24, Col ecotypes, and tAt1G9180, At4G25480 were highly responsive in the Eri and Col ecotypes during cold treatment. All of the ecotypes had at least 7 (out of 30 core cold responsive TFs) active TFs, except the Kond ecotype. This ecotype had only two significantly active TFs (At1G76590 and At4G04450).

Discussion

The analysis indicated that 10 *A. thaliana* ecotypes had different transcriptome level signatures in response to non-freezing cold treatment. Some of the cold-responding transcripts were differentially regulated among the ecotypes, irrespective of their geographical origin. Col-0, a cold tolerant ecotype, had significantly fewer differentially regulated transcripts, while Cvi, the most southern ecotype, had the highest number of differentially regulated transcripts. Among the differentially cold regulated transcripts, 75% showed ecotype specific expression patterns. In the ecotype specific differentially regulated gene list, we observed 315 transposable elements (TEs). These TEs may play an important role in plant genome evolution during adaption to local climatic temperatures. CBF genes were cold induced in all ecotypes, irrespective of their geographic origin. However, we observed different levels of expression among different ecotypes. Expression of the selected COR genes were not consistent in all ecotypes. Sequence data from the 1001-genome project indicated that

polymorphisms in their sequences may contribute to the dramatic differences in expression patterns. We observed significant numbers of non-synonymous amino acid changes in the coding region of the CBF genes. All of the CBFs exhibited significantly negative Tajima's D values, indicating an excess of rare and recent alleles. Gene ontology analysis showed that, apart from common cold stress regulated processes, several other categories of biological processes were differentially regulated in various ecotypes. Some of the important categories included pigment biosynthesis, circadian rhythm, response to light, response to water deprivation, and response to ABA. By looking at the differentially regulated genes related to pathogen responses induced by cold stress, we speculate that there might be co-evolution of concomitant stress related genes and processes.

We furthermore constructed an *in silico* transcriptl regulatory network model during cellular responses to non-freezing cold stress using gene expression data from ten ecotypes. The final network contained 1,275 nodes and 7,720 connections, which included 178 TFs and 1,331 target genes. Apart from retaining several previously known interactions (cross validated using AGRIS), many potentially novel interactions between TFs regulated during the cold stress response were detected. Differential regulatory activities were observed among the cold regulated TFs, which may contribute toward cold adaptation of the ten ecotypes. In addition, since the model is general, it could in principle be used to study networks regulating any biological process in any biological systems. As far as cold stress is concerned, it can be implemented to identify useful molecular markers to develop cold tolerant crop varieties.

Acknowledgments

We acknowledge funding from ERA-NET Plant Genomics (http://www.erapg.org/) for Multi-Stress project. The array experiments were run at the DTU Multi-Assay Core (DMAC) supported by an infrastructure grant from the Technical University of Denmark. PB was supported by the Biotechnology and Functional Genomics (FUGE) program of the Norwegian Research Council through grants NFR 184147/S10, 185173/V40, 184146/S10 and NDJ was supported by the Norwegian Research Council, through grant number 7017430. We thank Drs. Jens Rohloff, Christophe Pelabon and professor Nadav Skjøndal-Bar for critical reading of the manuscript.

References

AbuQamar S, Chen X, Dhawan R, Bluhm B, Salmeron J, Lam S, Dietrich RA, Mengiste T. 2006. Expression profiling and mutant analysis reveals complex regulatory networks involved in Arabidopsis response to Botrytis infection. *Plant J* **48**(1): 28-44.

Alonso-Blanco C, Aarts MGM, Bentsink L, Keurentjes JJB, Reymond M, Vreugdenhil D, Koornneef M. 2009. What Has Natural Variation Taught Us about Plant Development, Physiology, and Adaptation? *Plant Cell* **21**(7): 1877-1896.

Alonso-Blanco C, Gomez-Mena C, Llorente F, Koornneef M, Salinas J, Martinez-Zapater JM. 2005. Genetic and molecular analyses of natural variation indicate CBF2 as a candidate gene for underlying a freezing tolerance quantitative trait locus in Arabidopsis. *Plant Physiology* **139**(3): 1304-1312.

Alonso-Blanco C, Koornneef M. 2000. Naturally occurring variation in Arabidopsis: an underexploited resource for plant genetics. *Trends Plant Sci* **5**(1): 22-29.

Alter O, Brown PO, Botstein D. 2000. Singular value decomposition for genome-wide expression data processing and modeling. *Proceedings of the National Academy of Sciences of the United States of America* **97**(18): 10101-10106.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**(1): 25-29.

Atkinson NJ, Urwin PE. 2012. The interaction of plant biotic and abiotic stresses: from genes to the field. *J Exp Bot*.

Austin RS, Vidaurre D, Stamatiou G, Breit R, Provart NJ, Bonetta D, Zhang J, Fung P, Gong Y, Wang PW, McCourt P, Guttman DS. 2011. Next-generation mapping of Arabidopsis genes. *Plant J* **67**(4): 715-725.

Barabasi AL, Oltvai ZN. 2004. Network biology: understanding the cell's functional organization. *Nat Rev Genet* **5**(2): 101-113.

Cao J, Schneeberger K, Ossowski S, Gunther T, Bender S, Fitz J, Koenig D, Lanz C, Stegle O, Lippert C, Wang X, Ott F, Muller J, Alonso-Blanco C, Borgwardt K, Schmid KJ, Weigel D. 2011. Whole-genome sequencing of multiple Arabidopsis thaliana populations. *Nat Genet* **43**(10): 956-963.

Carvalho BS, Irizarry RA. 2010. A framework for oligonucleotide microarray preprocessing. *Bioinformatics* **26**(19): 2363-2367.

Carvallo MA, Pino MT, Jeknic Z, Zou C, Doherty CJ, Shiu SH, Chen TH, Thomashow MF. 2011. A comparison of the low temperature transcriptomes and CBF regulons of three plant species that differ in freezing tolerance: Solanum commersonii, Solanum tuberosum, and Arabidopsis thaliana. *J Exp Bot* **62**(11): 3807-3819.

Catala R, Medina J, Salinas J. 2011. Integration of low temperature and light signaling during cold acclimation response in Arabidopsis. *Proc Natl Acad Sci U S A* **108**(39): 16475-16480.

Chen WJ, Chang SH, Hudson ME, Kwan WK, Li J, Estes B, Knoll D, Shi L, Zhu T. 2005. Contribution of transcriptional regulation to natural variations in Arabidopsis. *Genome Biol* **6**(4): R32.

Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD. 2003. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Research* **31**(13): 3497-3500.

Clauset A, Shalizi CR, Newman MEJ. 2009. Power-Law Distributions in Empirical Data. *Siam Review* **51**(4): 661-703.

Cui X, Churchill GA. 2003. Statistical tests for differential expression in cDNA microarray experiments. *Genome Biol* **4**(4): 210.

Davuluri RV, Sun H, Palaniswamy SK, Matthews N, Molina C, Kurtz M, Grotewold E. 2003. AGRIS: Arabidopsis gene regulatory information server, an information resource of Arabidopsis cis-regulatory elements and transcription factors. *BMC Bioinformatics* **4**: 25.

Delker C, Poschl Y, Raschke A, Ullrich K, Ettingshausen S, Hauptmann V, Grosse I, Quint M. 2010. Natural variation of transcriptional auxin response networks in Arabidopsis thaliana. *Plant Cell* **22**(7): 2184-2200.

Edwards KD, Anderson PE, Hall A, Salathia NS, Locke JCW, Lynn JR, Straume M, Smith JQ, Millar AJ. 2006. FLOWERING LOCUS C mediates natural variation in the high-temperature response of the Arabidopsis circadian clock. *Plant Cell* **18**(3): 639-650.

Feng S, Martinez C, Gusmaroli G, Wang Y, Zhou J, Wang F, Chen L, Yu L, Iglesias-Pedraz JM, Kircher S, Schafer E, Fu X, Fan LM, Deng XW. 2008. Coordinated regulation of Arabidopsis thaliana development by light and gibberellins. *Nature* **451**(7177): 475-479.

Fowler DB, Limin AE. 2004. Interactions among factors regulating phenological development and acclimation rate determine low-temperature tolerance in wheat. *Annals of Botany* **94**(5): 717-724.

Fowler S, Thomashow MF. 2002. Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. *Plant Cell* **14**(8): 1675-1690.

Fowler SG, Cook D, Thomashow ME. 2005. Low temperature induction of Arabidopsis CBF1, 2, and 3 is gated by the circadian clock. *Plant Physiology* **137**(3): 961-968.

Franklin KA, Whitelam GC. 2007. Light-quality regulation of freezing tolerance in Arabidopsis thaliana. *Nature Genetics* **39**(11): 1410-1413.

Fujita M, Fujita Y, Noutoshi Y, Takahashi F, Narusaka Y, Yamaguchi-Shinozaki K, Shinozaki K. 2006. Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. *Current Opinion in Plant Biology* **9**(4): 436-442.

Galbraith SJ, Tran LM, Liao JC. 2006. Transcriptome network component analysis with limited microarray data. *Bioinformatics* **22**(15): 1886-1894.

Guberman JM, Ai J, Arnaiz O, Baran J, Blake A, Baldock R, Chelala C, Croft D, Cros A, Cutts RJ, Di Genova A, Forbes S, Fujisawa T, Gadaleta E, Goodstein DM, Gundem G, Haggarty B, Haider S, Hall M, Harris T, Haw R, Hu S, Hubbard S, Hsu J, Iyer V, Jones P, Katayama T, Kinsella R, Kong L, Lawson D, Liang Y, Lopez-Bigas N, Luo J, Lush M, Mason J, Moreews F, Ndegwa N, Oakley D, Perez-Llamas C, Primig M, Rivkin E, Rosanoff S, Shepherd R, Simon R, Skarnes B, Smedley D, Sperling L, Spooner W, Stevenson P, Stone K, Teague J, Wang J, Whitty B, Wong DT, Wong-Erasmus M, Yao L, Youens-Clark K, Yung C, Zhang J, Kasprzyk A. 2011. BioMart Central Portal: an open database network for the biological community. *Database (Oxford)* **2011**: bar041.

Guo A, He K, Liu D, Bai S, Gu X, Wei L, Luo J. 2005. DATF: a database of Arabidopsis transcription factors. *Bioinformatics* **21**(10): 2568-2569.

Guo YF, Xiao P, Lei SF, Deng FY, Xiao GG, Liu YZ, Chen XD, Li LM, Wu S, Chen Y, Jiang H, Tan LJ, Xie JY, Zhu XZ, Liang SP, Deng HW. 2008. How is mRNA expression predictive for protein expression? A correlation study on human circulating monocytes. *Acta Biochimica Et Biophysica Sinica* **40**(5): 426-436.

Haake V, Cook D, Riechmann JL, Pineda O, Thomashow MF, Zhang JZ. 2002. Transcription factor CBF4 is a regulator of drought adaptation in Arabidopsis. *Plant Physiology* **130**(2): 639-648.

Hall BK, Hallgrímsson B, Strickberger MW. 2008. *Strickberger's evolution : the integration of genes, organisms and populations*. Sudbury, Mass.: Jones and Bartlett.

Hannah MA, Wiese D, Freund S, Fiehn O, Heyer AG, Hincha DK. 2006. Natural genetic variation of freezing tolerance in Arabidopsis. *Plant Physiology* **142**(1): 98-112.

Harmer SL, Hogenesch LB, Straume M, Chang HS, Han B, Zhu T, Wang X, Kreps JA, Kay SA. 2000. Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. *Science* **290**(5499): 2110-2113.

Havaux M, Kloppstech K. 2001. The protective functions of carotenoid and flavonoid pigments against excess visible radiation at chilling temperature investigated in Arabidopsis npq and tt mutants. *Planta* **213**(6): 953-966.

Heddad M, Noren H, Reiser V, Dunaeva M, Andersson B, Adamska I. 2006. Differential expression and localization of early light-induced proteins in Arabidopsis. *Plant Physiology* **142**(1): 75-87.

Herrgard MJ, Covert MW, Palsson BO. 2004. Reconstruction of microbial transcriptional regulatory networks. *Curr Opin Biotechnol* **15**(1): 70-77.

Horton MW, Hancock AM, Huang YS, Toomajian C, Atwell S, Auton A, Muliyati NW, Platt A, Sperone FG, Vilhjalmsson BJ, Nordborg M, Borevitz JO, Bergelson J. 2012. Genome-wide patterns of genetic variation in worldwide Arabidopsis thaliana accessions from the RegMap panel. *Nat Genet* **44**(2): 212-216.

Hoshino T, Xiao N, Tkachenko OB. 2009. Cold adaptation in the phytopathogenic fungi causing snow molds. *Mycoscience* **50**(1): 26-38.

Ideker T, Krogan NJ. 2012. Differential network biology. *Mol Syst Biol* **8**: 565.

Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. 2003. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Research* **31**(4): e15.

Jaglo-Ottosen KR, Gilmour SJ, Zarka DG, Schabenberger O, Thomashow MF. 1998. Arabidopsis CBF1 overexpression induces COR genes and enhances freezing tolerance. *Science* **280**(5360): 104-106.

Kao KC, Yang YL, Boscolo R, Sabatti C, Roychowdhury V, Liao JC. 2004. Transcriptome-based determination of multiple transcription regulator activities in Escherichia coli by using network component analysis. *Proc Natl Acad Sci U S A* **101**(2): 641-646.

Kashkush K, Feldman M, Levy AA. 2002. Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. *Genetics* **160**(4): 1651-1659.

Kerr MK, Martin M, Churchill GA. 2000. Analysis of variance for gene expression microarray data. *J Comput Biol* **7**(6): 819-837.

Keurentjes JJB, Fu JY, Terpstra IR, Garcia JM, van den Ackerveken G, Snoek LB, Peeters AJM, Vreugdenhil D, Koornneef M, Jansen RC. 2007. Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci. *Proceedings of the National Academy of Sciences of the United States of America* **104**(5): 1708-1713.

Kuwabara C, Imai R. 2009. Molecular Basis of Disease Resistance Acquired through Cold Acclimation in Overwintering Plants. *Journal of Plant Biology* **52**(1): 19-26.

Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander DL, Garcia-Hernandez M, Karthikeyan AS, Lee CH, Nelson WD, Ploetz L, Singh S, Wensel A, Huala E. 2012. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Research* **40**(Database issue): D1202-1210.

Lee BH, Henderson DA, Zhu JK. 2005. The Arabidopsis cold-responsive transcriptome and its regulation by ICE1. *Plant Cell* **17**(11): 3155-3175.

Li G, Siddiqui H, Teng Y, Lin R, Wan XY, Li J, Lau OS, Ouyang X, Dai M, Wan J, Devlin PF, Deng XW, Wang H. 2011. Coordinated transcriptional regulation underlying the circadian clock in Arabidopsis. *Nat Cell Biol* **13**(5): 616-622.

Liao JC, Boscolo R, Yang YL, Tran LM, Sabatti C, Roychowdhury VP. 2003. Network component analysis: reconstruction of regulatory signals in biological systems. *Proc Natl Acad Sci U S A* **100**(26): 15522-15527.

Maere S, Heymans K, Kuiper M. 2005. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* **21**(16): 3448-3449.

McKhann HI, Gery C, Berard A, Leveque S, Zuther E, Hincha DK, De Mita S, Brunel D, Teoule E. 2008a. Natural variation in CBF gene sequence, gene expression and freezing tolerance in the Versailles core collection of Arabidopsis thaliana. *Bmc Plant Biology* **8**.

McKhann HI, Gery C, Berard A, Leveque S, Zuther E, Hincha DK, De Mita S, Brunel D, Teoule E. 2008b. Natural variation in CBF gene sequence, gene expression and freezing tolerance in the Versailles core collection of Arabidopsis thaliana. *BMC Plant Biol* **8**: 105.

Medina J, Catala R, Salinas J. 2011. The CBFs: Three arabidopsis transcription factors to cold acclimate. *Plant Science* **180**(1): 3-11.

Mikkelsen MD, Thomashow MF. 2009. A role for circadian evening elements in cold-regulated gene expression in Arabidopsis. *Plant Journal* **60**(2): 328-339.
Mitchell-Olds T, Schmitt J. 2006. Genetic mechanisms and evolutionary significance of natural variation in Arabidopsis. *Nature* **441**(7096): 947-952.
Mittler R. 2006. Abiotic stress, the field environment and stress combination. *Trends Plant Sci* **11**(1): 15-19.
Moreau Y, Aerts S, De Moor B, De Strooper B, Dabrowski M. 2003. Comparison and meta-analysis of microarray data: from the bench to the computer desk. *Trends Genet* **19**(10): 570-577.
Pascal LE, True LD, Campbell DS, Deutsch EW, Risk M, Coleman IM, Eichner LJ, Nelson PS, Liu AY. 2008. Correlation of mRNA and protein levels: cell type-specific gene expression of cluster designation antigens in the prostate. *BMC Genomics* **9**: 246.
Penfield S. 2008. Temperature perception and signal transduction in plants. *New Phytol* **179**(3): 615-628.
Plazek A, Zur I. 2003. Cold-induced plant resistance to necrotrophic pathogens and antioxidant enzyme activities and cell membrane permeability. *Plant Science* **164**(6): 1019-1028.
R Core Team 2012,. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
Rapacz M, Wolanin B, Hura K, Tyrka M. 2008. The effects of cold acclimation on photosynthetic apparatus and the expression of COR14b in four genotypes of barley (Hordeum vulgare) contrasting in their tolerance to freezing and high-light treatment in cold conditions. *Annals of Botany* **101**(5): 689-699.
Rasmussen S, Barah P, Suarez-Rodriguez MC, Friis P, Costantino P, Bones AM, Nielsen HB, Mundy J. in press. Transcriptome response to combinations of stress in Arabidopsis thaliana. *Plant Physiology*: (In press).
Riano-Pachon DM, Ruzicic S, Dreyer I, Mueller-Roeber B. 2007. PlnTFDB: an integrative plant transcription factor database. *BMC Bioinformatics* **8**: 42.
Rohde P, Hincha DK, Heyer AG. 2004. Heterosis in the freezing tolerance of crosses between two Arabidopsis thaliana accessions (Columbia-0 and C24) that show differences in non-acclimated and acclimated freezing tolerance. *Plant Journal* **38**(5): 790-799.
Saeed AI, Hagabati NK, Braisted JC, Liang W, Sharov V, Howe EA, Li JW, Thiagarajan M, White JA, Quackenbush J. 2006. TM4 microarray software suite. *DNA Microarrays, Part B: Databases and Statistics* **411**: 134-+.
Salter MG, Franklin KA, Whitelam GC. 2003. Gating of the rapid shade-avoidance response by the circadian clock in plants. *Nature* **426**(6967): 680-683.
Sanghera GS, Wani SH, Hussain W, Singh NB. 2011. Engineering cold stress tolerance in crop plants. *Curr Genomics* **12**(1): 30-43.
Scacchi E, Salinas P, Gujas B, Santuari L, Krogan N, Ragni L, Berleth T, Hardtke CS. 2010. Spatio-temporal sequence of cross-regulatory events in root meristem growth. *Proc Natl Acad Sci U S A* **107**(52): 22734-22739.
Segal E, Shapira M, Regev A, Pe'er D, Botstein D, Koller D, Friedman N. 2003. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nature Genetics* **34**(2): 166-176.
Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**(11): 2498-2504.
Shimizu KK. 2002. Ecology meets molecular genetics in Arabidopsis. *Population Ecology* **44**(3): 221-233.
Somerville C, Koornneef M. 2002. A fortunate choice: the history of Arabidopsis as a model plant. *Nat Rev Genet* **3**(11): 883-889.
Swindell WR, Huebner M, Weber AP. 2007. Plastic and adaptive gene expression patterns associated with temperature stress in Arabidopsis thaliana. *Heredity* **99**(2): 143-150.

Tajima F. 1989. Statistical-Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics* **123**(3): 585-595.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* **28**(10): 2731-2739.

Thomashow MF. 1999. PLANT COLD ACCLIMATION: Freezing Tolerance Genes and Regulatory Mechanisms. *Annu Rev Plant Physiol Plant Mol Biol* **50**: 571-599.

Tian Q, Stepaniants SB, Mao M, Weng L, Feetham MC, Doyle MJ, Yi EC, Dai HY, Thorsson V, Eng J, Goodlett D, Berger JP, Gunter B, Linseley PS, Stoughton RB, Aebersold R, Collins SJ, Hanlon WA, Hood LE. 2004. Integrated genomic and proteomic analyses of gene expression in mammalian cells. *Molecular & Cellular Proteomics* **3**(10): 960-969.

Tirosh I, Barkai N. 2011. Inferring regulatory mechanisms from patterns of evolutionary divergence. *Mol Syst Biol* **7**: 530.

Wang JG, Qiu XJ, Li YH, Deng YP, Shi TL. 2011. A transcriptional dynamic network during Arabidopsis thaliana pollen development. *Bmc Systems Biology* **5**.

Wang W, Barnaby JY, Tada Y, Li H, Tor M, Caldelari D, Lee DU, Fu XD, Dong X. 2011. Timing of plant immune responses by a central circadian regulator. *Nature* **470**(7332): 110-114.

Yamaguchishinozaki K, Shinozaki K. 1994. A Novel Cis-Acting Element in an Arabidopsis Gene Is Involved in Responsiveness to Drought, Low-Temperature, or High-Salt Stress. *Plant Cell* **6**(2): 251-264.

Yilmaz A, Mejia-Guerra MK, Kurz K, Liang X, Welch L, Grotewold E. 2011. AGRIS: the Arabidopsis Gene Regulatory Information Server, an update. *Nucleic Acids Research* **39**(Database issue): D1118-1122.

Zhen Y, Ungerer MC. 2008a. Clinal variation in freezing tolerance among natural accessions of Arabidopsis thaliana. *New Phytologist* **177**(2): 419-427.

Zhen Y, Ungerer MC. 2008b. Relaxed Selection on the CBF/DREB1 Regulatory Genes and Reduced Freezing Tolerance in the Southern Range of Arabidopsis thaliana. *Molecular Biology and Evolution* **25**(12): 2547-2555.

Zhu Y, Qian WQ, Hua J. 2010. Temperature Modulates Plant Defense Responses through NB-LRR Proteins. *Plos Pathogens* **6**(4).

Ziolkowski PA, Koczyk G, Galganski L, Sadowski J. 2009. Genome sequence comparison of Col and Ler lines reveals the dynamic nature of Arabidopsis chromosomes. *Nucleic Acids Research* **37**(10): 3189-3201.

Zust T, Heichinger C, Grossniklaus U, Harrington R, Kliebenstein DJ, Turnbull LA. 2012. Natural enemies drive geographic variation in plant defenses. *Science* **338**(6103): 116-119.

Zuther E, Schulz E, Childs LH, Hincha DK. 2012. Clinal variation in the non-acclimated and cold-acclimated freezing tolerance of Arabidopsis thaliana accessions. *Plant Cell Environ*.

Tables

Table1: Geographic distribution of the 10 *A. thaliana* ecotypes and numbers of cold regulated genes in each ecotype ($p \leq 0.01$). Up and down regulation was calculated based on fold change ratios compared to respective untreated controls in individual ecotypes. Ecotype Col, a cold tolerant ecotype, has fewer cold regulated transcripts compared to the others. In contrast, Cvi, the southernmost accession from a warm climate, exhibits the greatest number of cold responsive transcripts.

*Unique= Unique to the respective ecotype

	Geographic location of the ecotypes		Differentially regulated transcripts					
Eotype	**Origin**	**Latitude (°North)**	**Total**	**Total Up**	**Total Down**	**Unique (Total)**	**Unique (up)**	**Unique (Down)**
Cvi	Cape Verdia Islands	16	2004	603	1401	1230	276	954
Kas-1	Kashmir, India	34	1097	487	610	442	153	289
Kyo-2	Kyoto city, western part of Hoshu Island, Japan	35.5	877	458	419	305	104	201
Sha	Pamiro-Alay, Tadjikistan	39	620	215	405	268	70	198
Col	Columbia, United States	38.5	185	120	65	89	48	41
Kond	Khurmatov, Tadjikistan	38.8	814	428	386	384	167	217
C24	Coimbra, Portugal	40	1427	931	496	758	460	208
Ler	Landsberg, Poland	48	619	195	424	348	91	257
An-1	Antwerpern, Belgium	51.5	632	308	324	188	69	119
Eri	Erigsboda, Sweden	56	967	804	163	541	427	114

Table 2: List of cold regulated (up and down) TFs unique for each of the 10 ecotypes (significance threshold $p \leq 0.01$).

Ecotype	Unique TFs (Up regulated)	Unique TFs (Down regulated)
Cvi	ANAC014, ANAC042, ANAC058, AtHB24, AtHB32, AtMYB103, CUC1, HEC1, HSFB2A, LBD27, LBD35, LD, PIL6, SRL2, At3g20880, At4g00150, At1g09060, At3g11450, At5g45270	AIL6, ANAC041, ANAC074, ANAC103, ARR2, ASML2, AtHB23, AtMYB11, AtMYB17, AtMYB86, AtNAC3, AtY13, bt5, EMB3008, ETC1, GNC, HAt22, HDG12, IAA18, LBD23, MAF4, MYB113, MYB3, MYB33, MYB92, MYC6.2, NTL9, PAN, PCF1, PUX2, SDG40, SGR1, SPL5, SUVR4, tcp17, TCP3, TGA6, VND1, WOX13, WRKY50, At1g16640, At3g06160, At4g34400, At4g00940, At5g49300, At3g57480, At5g10970, At2g05160, At5g40880, At3g16940, t5g38140, At2g20110, At1g07520, At1g63100, At1g44810, At4g00232, At4g26170, At1g09710, At1g33420, At2g01810, At3g53370, At5g51910, At1g76870, At1g26260, At1g62975, At4g00870, At4g14410, At4g29930, At5g46830, At5g65320
Kas.1	AtGRF3, BPC6, HAt3, HSFA8, IAA29, SSL2, SWN, WRKY3, WRKY32, WRKY66, At3g45260, At1g67310, At2g45460	AL1, BT4, DUO1, GBF6, HSFB1, HSFB4, MNP, TED5, TIFY3B, U2AF35B, ZBF1, ZFN3, ZFP4, At5G52020, At5G06770, At5G41920, At4G22140, At5G50670, At1G03040, At3G23690
Kyo.2	AtIDD2, CAL1, HAt14, LCL1, PHE2, RAP2.9, SNZ, SPL3, TOC1, At4g18870, At5g51790	HDT3, PMG1, TRFL6, UNE12, WER1, At5g61190, At4g23800, At2g45800, At1g69170, At1g68920, At2g46510,
Col	At3G50750, At2G27630, At5G22990, At1G48195, At4G37850	ACD6, ACS3, CYP71A28, CYP81K1, MEA, MLP28, MYB24, PSRP5, RCK, STR16, XIJ, At3g21570, At5g33260, At2g21930, At5g26930, At5g15620, At3g18840, At1g31370, At2g01031, At2g09850, At2g24930, At1g79120, At4g09070, At3g56600, At5g39150, At2g35300, At1g23570, At5g02140, At1g23060, At3g14750, At1g27300, At3g16840, At3g03920, At2g07671, At1g53060, At5g66230, At5g58570, At5g26690, At1g27330, At1g18720, At5g18850
Kond	AGL79, ANAC077, AtGRF6, LBD14, SCL11, TIFY9 (JAZ10), WRKY10, At3g06410, At3g51950,	ADOF2, ANAC097, AtHB27, BME3-ZF, bZIP61, GIF2, HAP3A, HAP5B,ING1, RAP2.11, TGA1, At3g51080, At2g40970, At5g47390
Sha	AGL58, ANAC009, SET1, STY2, At4g33880	AGL24, AGL43, CRF1, ETT, IAA7, LBD38, PHB-1D, PTAC1, SRS8, ZFHD2, At5g02460, At5g41030,
C24	ADOF1, ANAC045, ANAC061, ANAC069, AtIDD16, AtIDD5, CDF1, CDF2 , COL2 , DREB2A, HSF A4A, IAA1, LBD32, MP, MYB51, MYB77, RVE2, SPL1R2, SSRP1,SUVH3, WRKY26, WRKY33, WRKY40, WRKY55, At5g51190, At2g17410, At4g17570, At1g26610, At4g15420, At5g26610, At5g12440, At3g52250, At5g06110, At1g20640, At1g64530, At2g18090, At2g37520, At1g01260 , At5g57150	ANAC065, AtHB16, AtMYB63, BPC5,HSFB2B, LBD22, RGA2, SCL3, ZFHD3, At1g49475, At1g68520, At4g24060, At3g24050, At4g14540, At2g44730, At2g21235
Ler	ddf2, HSFC1	AGL26, ARF21, DAR7, emb2746, LBD1, MYB105, NAI1, ZFP6, At4g31680, At5g12850, At1g75530, At5g41765, At2g17150
An.1	AGL56, RD26, SOC1, SUVH1, tify5a, At1g79700, At4g15250, At2g42150,	AtGRF6, LBD14 , At3g51950
Eri	ALC, ANAC011, ANAC019, ANAC044, ANAC046, AtAF2, AtNAC3, AZF2, BPC4, COL9, DAG2, ERF5, ERF8, IAA17, IAA5, MYB59, MYBR1, RAP2.10, SHN3, SHY1, TIFY10B, WRKY22, WRKY27, WRKY28, ZFHD4, At4g01580, At2g40340, At2g40350, At4g32800, At2g45050, At3g49930, At3g60580, At3g08505, At3g14020, At1g25550, At3g12730, At5g01200, At5g05790, At3g21210, At3g53680, At2g18850, At3g21330, At3g23210 , At1g19490	AtbZIP,BZR2, DAG2 , WOX12, ZFP8, At3g23140, At1g19040

Table 3: GO category attribute matrix from the significantly regulated gene-list for each ecotype generated using BiNGO (Hypergeometric test, Benjamini & Hochberg False Discovery Rate FDR correction, significance level 0.05). Only some of the overlapping GO categories are included in this table. Detailed categories, including ecotype specific categories and corresponding gene lists, are included in **Supporting file S3.**

GO-terms	An.1	Col 0	Cvi	Eri	Kas.1	Kond	Kyo.2	Ler	C24	Sha
Response to abiotic stimulus	√	√	√	√	√	√	√	√	√	√
Response to chemical stimulus	√	√	√	√	√	√	√	√	√	√
Response to cold	√	√	√	√	√	√	√	√	√	√
Response to organic substance	√	√	√	√	√	√	√	√	√	√
Response to endogenous stimulus	√	√	√	√	√	√	√	√	√	√
Response to hormone stimulus	√	√	√	√	√	.	√	√	√	√
Circadian rhythm	√	√	√	√	√	.	√	√	√	√
Response to light stimulus	√	.	√	√	√	.	√	√	√	√
Response to water	√	√	.	√	√	.	√	√	√	√
Response to jasmonic acid stimulus	√	.	√	√	.	√	√	√	√	√
Response to water deprivation	√	√	.	√	√	.	√	√	√	√
Response to red or far red light	√	.	√	√	√	.	√	√	√	√
Cold acclimation	√	.	.	√	√	√	√	√	√	√
Rresponse to other organism	√	.	√	.	√	√	√	√	√	√
Response to blue light	√	.	√	√	√	.	√	√	√	√
Response to abscisic acid stimulus	√	.	√	.	√	√	√	.	√	√
Response to far red light	√	.	√	√	√	.	√	√	√	.
Multi-organism process	√	.	√	.	√	√	√	.	√	√
Response to red light	√	.	√	√	√	.	√	√	√	.
Response to fungus	√	.	√	.	√	.	√	√	√	√
Response to carbohydrate stimulus	√	.	.	√	√	√	√	.	√	√
Regulation of transcription	√	.	.	√	√	.	√	.	√	√
Regulation of macromolecule biosynthetic process	√	.	.	√	√	.	√	.	√	√
Regulation of nucleobase, acid metabolic process	√	.	.	√	√	.	√	.	√	√
Pigment biosynthetic process	√	.	√	.	.	√	√	√	√	.
Regulation of biosynthetic process	√	.	.	√	√	.	√	.	√	√
Regulation of nitrogen compound metabolic process	√	.	.	√	√	.	√	.	√	√
Regulation of gene expression	√	.	.	√	√	.	√	.	√	√
Regulation of cellular metabolic process	√	.	.	√	√	.	√	.	√	√
Regulation of primary metabolic process	√	.	.	√	√	.	√	.	√	√
Regulation of macromolecule metabolic process	√	.	.	√	√	.	√	.	√	√
Regulation of metabolic process	√	.	.	√	√	.	√	.	√	√
Response to chitin	√	√	.	√	.	.	√	.	√	√
Regulation of cellular process	√	.	.	√	√	.	√	.	√	√
Response to osmotic stress	√	.	.	√	√	.	√	.	√	.
Response to ethylene stimulus	√	.	.	√	√	.	√	.	√	.
Regulation of transcription, DNA-dependent	√	.	.	√	√	.	√	.	√	.
Regulation of RNA metabolic process	√	.	.	√	√	.	√	.	√	.
Chlorophyll biosynthetic process	√	.	√	.	.	.	√	√	√	.
Porphyrin biosynthetic process	√	.	√	.	.	.	√	√	√	.

Tetrapyrrole biosynthetic process	√	.	√	.	.	.	√	√	√	.
Regulation of biological process	√	.	.	√	.	.	√	.	√	√
Red or far-red light signalling pathway	.	.	√	.	√	.	.	√	√	√
Cellular response to radiation	.	.	√	.	√	.	.	√	√	√

Table 4: Ecotype specific transcript level profiles of 30 cold responsive TFs in the ten eco-types

TF ID	Alias	Ecotypes
At1G01060	LHY1	Eri, Col, Cvi, Kyo-2
At1G04240	SHY2	Sha
At1G13260	RAV1	Eri, C24
At1G19180	TIFY10A	Eri, Col
At1G27730	ZAT10	Eri, C24, Col
At1G28050	At1G28050	Eri, C24, Col
At1G46768	RAP2.1	An-1, Sha, L*er*, Cvi
At1G76590	At1G76590	Eri, Kond, C24, An-1, Col, Sha, L*er*, Cvi, Kyo-2, Kas-1
At1G78700	At1G78700	Sha, L*er*, Cvi, Kas-1
At2G30250	WRKY25	Sha, L*er*, Cvi
At2G40140	ZFAR1/CZF1	Eri, C24
At2G46830	CCA1	Sha
At3G11020	DREB2B	Sha
At3G17609	HYH	Eri, C24, Col
At4G04450	WRKY42	Eri, Kond, C24, An-1, Col, Cvi, Kyo-2, Kas-1
At4G17490	ERF-6-6	An-1, Sha, L*er*, Cvi, Kyo-2, Kas-1
At4G24240	WRKY7	Eri, Col
At4G25470	CBF2	C24, Col, Kas-1
At4G25480	DREB1A	Eri, Col
At4G25490	DREB1B/CBF1	Eri
At4G27950	CRF4	Eri, C24, Col
At4G28140	At4G28140	An-1, Sha, L*er*, Cvi, Kas-1
At4G29190	At4G29190	Eri, An-1, Cvi, Kyo-, Kas-1
At5G17300	At5G17300	Eri, C24, Col
At5G17490	RGL3	Col
At5G24470	PRR5	An-1, L*er*, Cvi, Kyo-2, Kas-1
At5G46710	At5G46710	An-1, Sha, L*er*, Cvi, Kas-1
At5G47220	ERF2	Sha, L*er*, Kas-1
At5G48250	At5G48250	Eri, C24, Col
At5G61270	PIF7	Col, Kas-1

Figure legends:

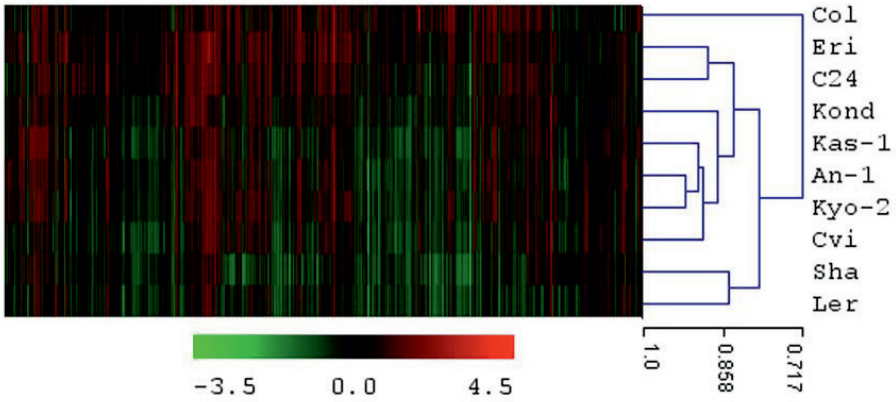


Fig. 1

Heat map visualization and hierarchical clustering (with Pearson's correlation coefficient using average linkage method) of top 1000 cold regulated transcripts based on fold change ratios compared to their respective controls from 10 different ecotypes. Genes are shown as columns and ecotypes as rows. As a global observation, this heat map indicates differential regulation signatures in response to non-freezing cold treatment in different *A. thaliana* ecotypes. Cold tolerant ecotype Col-0 ecotype is distinctly separated out from others.

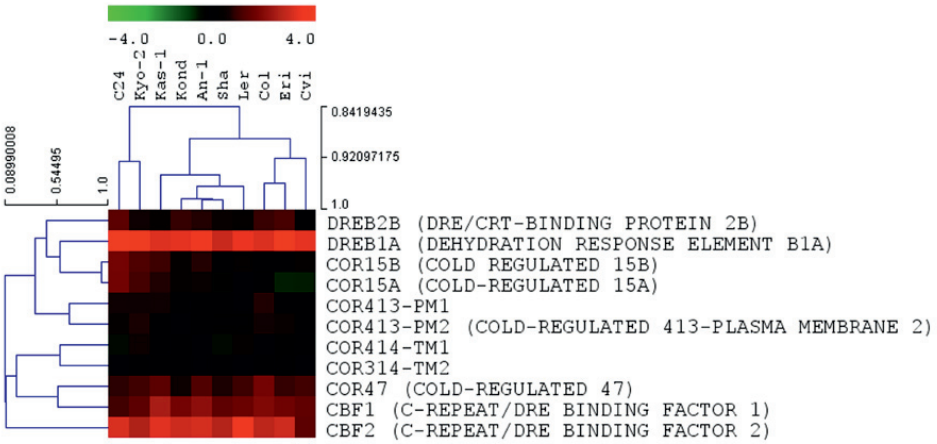


Fig. 2

CBF and selected *COR* genes were cold regulated in all accessions. However, there were noticeable differences in the levels of expression among the ten ecotypes.

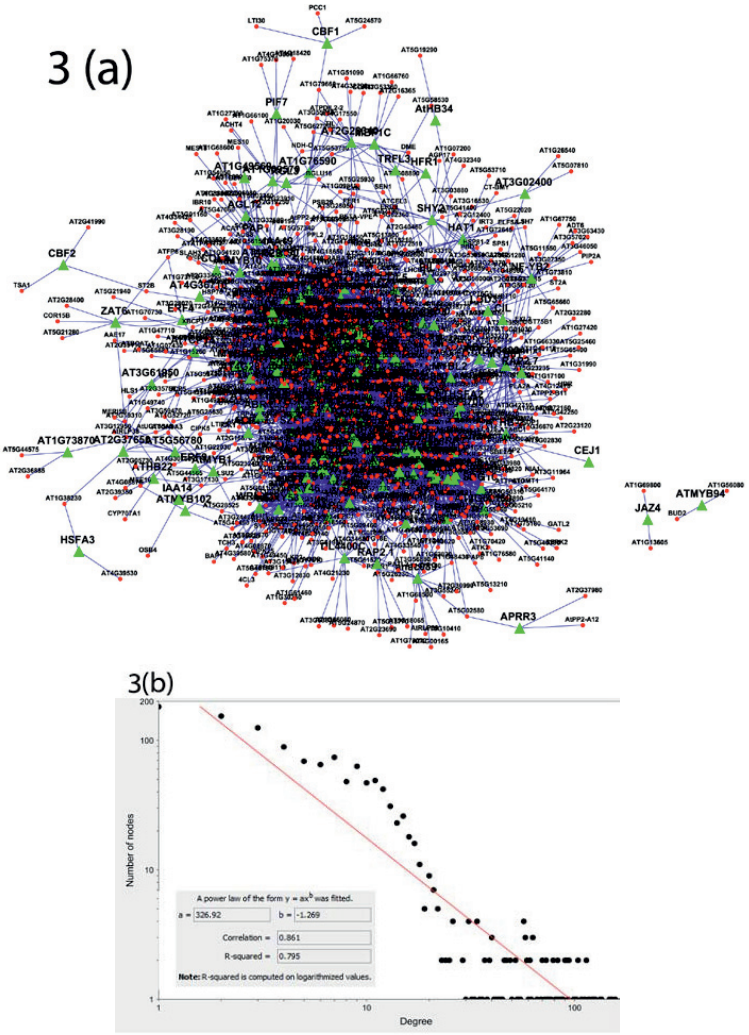


Fig. 3

(a) Topological overview of the transcript regulatory network constructed using cold stress microarray data from 10 ecotypes. This network contains 1,275 nodes and 7,720 connections. Transcription factors are marked as green triangles and target genes as red circles. Predicted regulatory interactions are shown as arrow (→) for activation and down-horizontal bar (⊣) as repression. Network visualized in Cytoscape software using 'Force-Directed Layout'.

(b) Power-law degree distribution of the network *P(k)* at correlation thresholds (≥ 0.8) . k indicates connectivity, and *P(k)* the connectivity distribution.

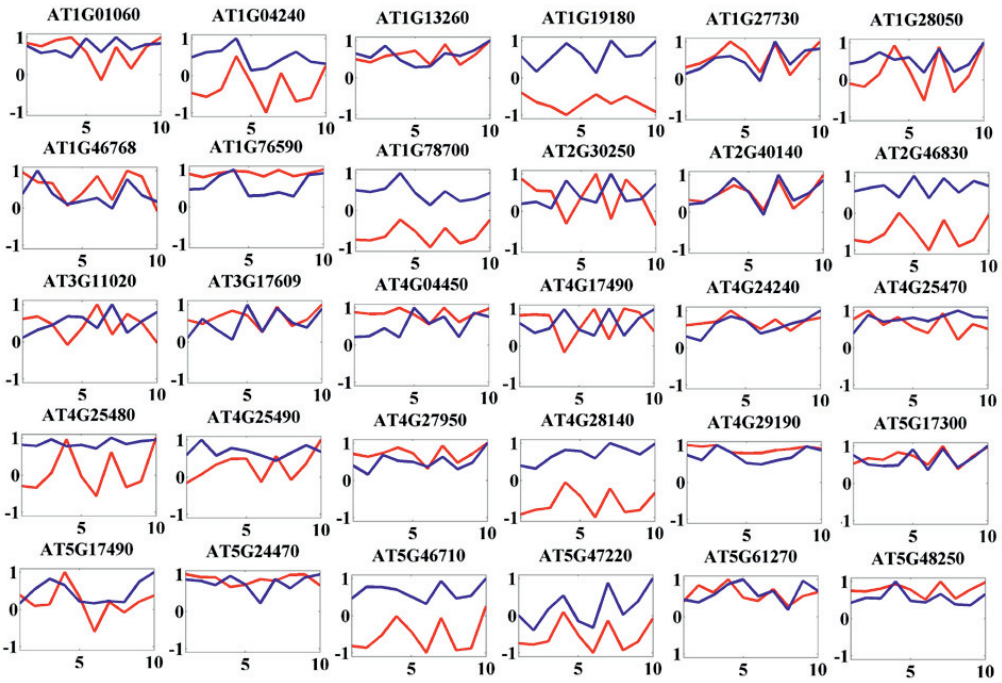


Fig. 4

Differential activities of 30 known cold regulated transcription factors in the ten ecotypes predicted using NCA. Rows represent the TFs and columns different eco-type responses to cold treatment. Transcription factor activities were shown in blue their expression values were represented as red colour. Here, values are scaled for direct comparison purposes. X-axis represents the different eco-types (1=Cvi, 2=Kas1, 3=Kyo.2, 4=Col, 5=Kond, 6=Sha, 7=C24, 8=Ler, 9=An.1, 10=Eri).

Supporting Information

S1: List of all transcripts from 10 ecotypes with annotations, p-values and fold change values during cold treatment. List of cold regulated transcripts and TFs are in separate sheet of the same excel file.

S2: List of ecotype specific gene expression during cold treatment from 10 ecotypes.

S3: Results of Gene Set Enrichment Analysis using BinGO.

S4: GO category contingency table from the significantly regulated gene-list (extended version of **Table 3**).

S5: Analysis of sequence polymorphism in CBF and COR genes.

S6: Predicted TF-TG regulatory connections including patterns of regulation and connectivity strengths.

S7: Predicted TF-TG regulatory networks as a .cys file, which can be opened locally for interactive exploration using cytoscape software from http://www.cytoscape.org. The presented view of the annotated network in this manuscript has been simplified manually for easier representation.

(S5)

Results from Tajima's Neutrality Test [1]

Genes	m	S	ps	T	p	D
CBF1	501	680	.0687560	.0101220	.005514	-1.291374
CBF2	496	590	.0595360	.0087770	.001834	-2.222324
CBF3/DREBA1	476	700	.0704930	.0104570	.002521	-2.164592

*Abbreviations: m = number of sequences, S = Number of segregating sites, ps = S/m, T = ps/a1, p = nucleotide diversity, and D is the Tajima test statistic (see chapter 12 in ref. [3] for details).

The analysis involved nucleotide sequences downloaded from Salk Arabidopsis thaliana 1001 Genomes (http://signal.salk.edu/atg1001/accessions.php). Initially sequences from all available ecotypes (724) were downloaded, but incomplete sequences were discarded before the analysis. Apart from the coding regions we have considered 1000 bp upstream sequences for alignment. We have considered Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA5 [2,3]. A negative Tajima's D signifies an excess of low frequency polymorphisms relative to expectation, indicating population size expansion (e.g., after a bottleneck or a selective sweep) and/or purifying selection. A very rough rule of thumb to significance is that values greater than +2 or less than -2 are likely to be significant. This rule is based on an appeal to asymptotic properties of some statistics, and thus +/- 2 does not actually represent a critical value for a significance test.

REFERENCES:

1. Tajima F. (1989). Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. Genetics 123:585-595.

2. Tamura K., Peterson D., Peterson N., Stecher G., Nei M., and Kumar S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution (In Press).

3. Nei M. and Kumar S. (2000). Molecular Evolution and Phylogenetics. Oxford University Press, New York.

Sequence Polymorphisms seen in the CBF genes (coding regions)

We obs erved s ignificant num ber of non -synonymous amino a cid c hanges i n t he c oding region of the CBF genes.

Nucleotide

A SNPs -- Red line
C SNPs -- Blue line
G SNPs -- Green line
T SNPs -- Yellow line
1 bp deletions -- Black line
Unsequenced regions -- .. (dot) or grey area

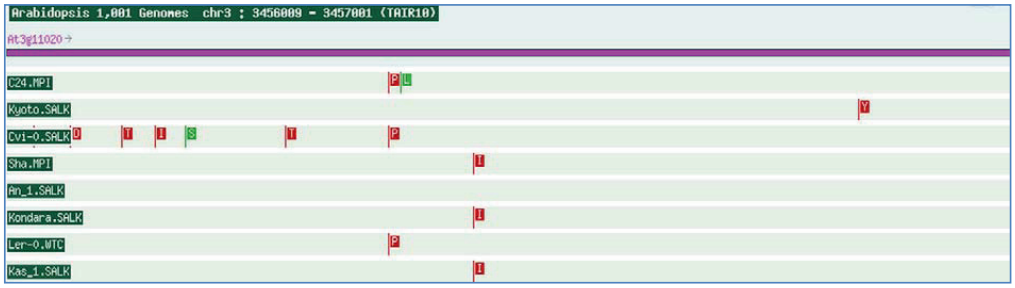
Amino Acide

Synonymous amino acid -- Green line
Non-synonymous amino acid -- Red line

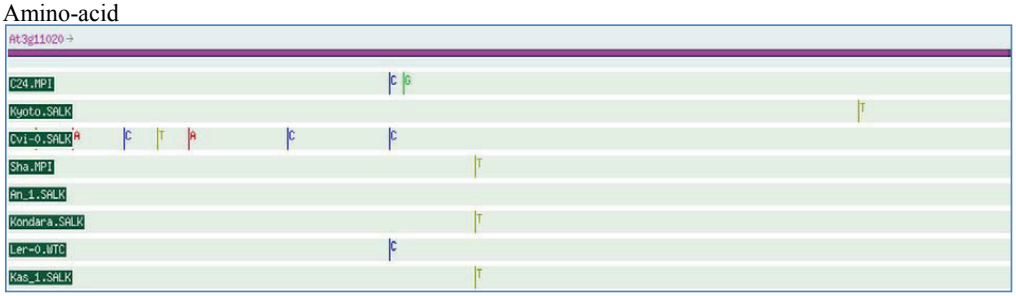
All standard amino acide codes are used. In addition, a Z - indicates an unknown amino acid, possibly a deletion and an X - indicates an amino acid has become a stop codon.

1) DREB2B (DRE/CRT-BINDING PROTEIN 2B)

Nucleotide

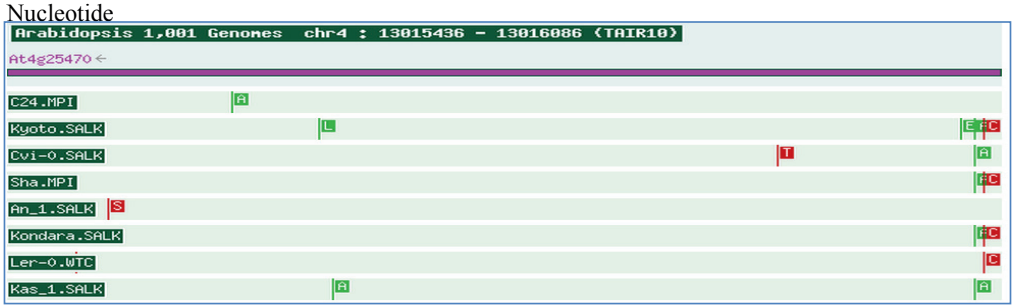


Amino-acid

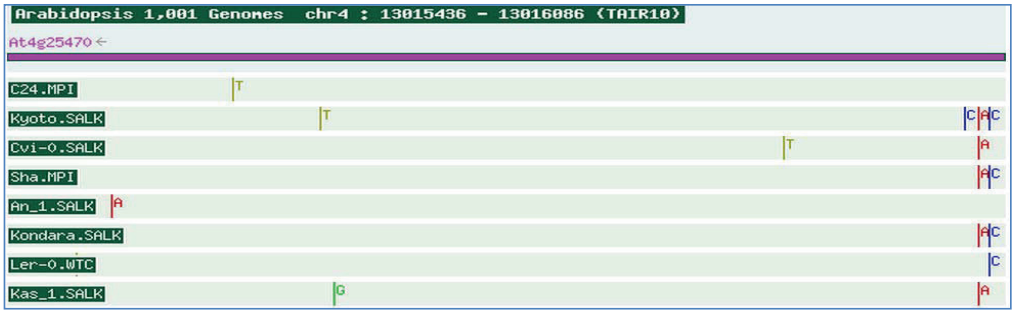


2) CBF2 (C-REPEAT/DRE BINDING FACTOR 2)

Nucleotide

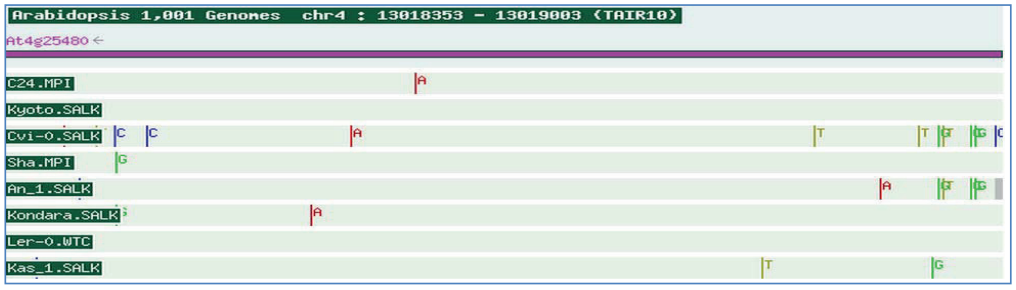


Amino-acid

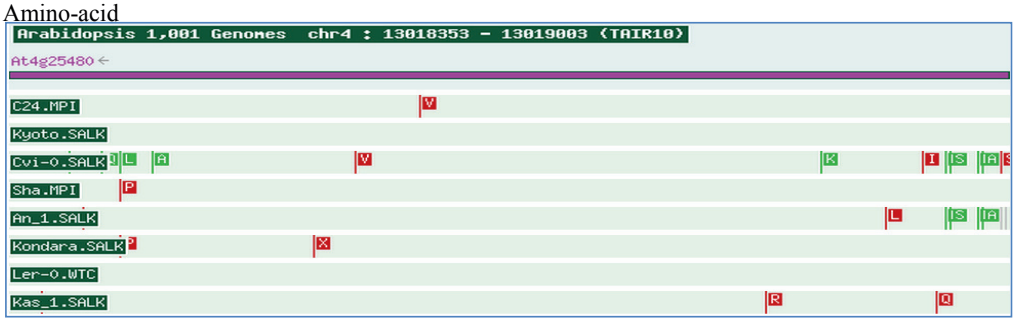


3) DREB1A (DEHYDRATION RESPONSE ELEMENT B1A)

Nucleotide

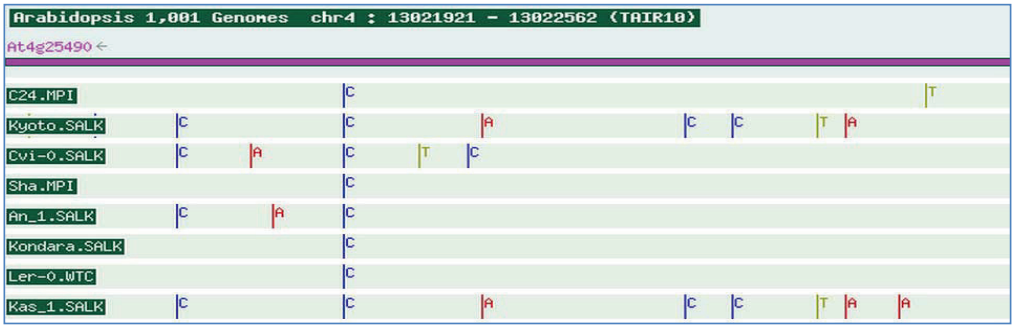


Amino-acid

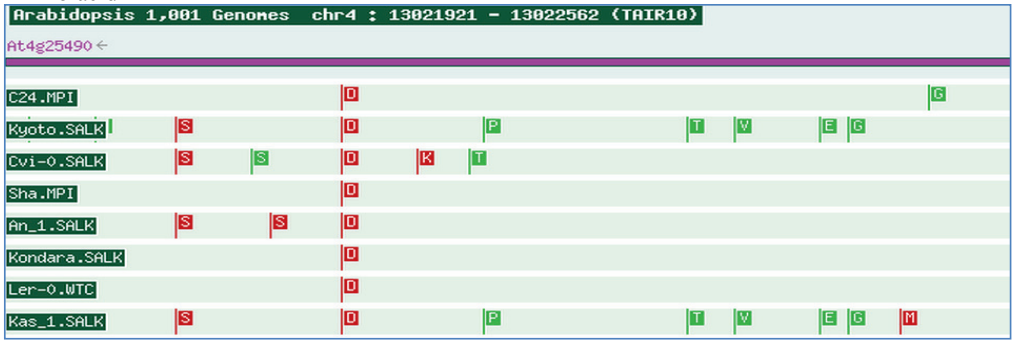


4) CBF1 (C-REPEAT/DRE BINDING FACTOR 1)

Nucleotide



Amino-acid



CHAPTER 6

(Paper V)

Is not included due to copyright

CHAPTER 7.a

(Paper VI)

Is not included due to copyright

CHAPTER 7.b

(Paper VII)

Is not included due to copyright

Doctoral theses in Biology
Norwegian University of Science and Technology
Department of Biology

Year	Name	Degree	Title
1974	Tor-Henning Iversen	Dr. philos Botany	The roles of statholiths, auxin transport, and auxin metabolism in root gravitropism
1978	Tore Slagsvold	Dr. philos Zoology	Breeding events of birds in relation to spring temperature and environmental phenology
1978	Egil Sakshaug	Dr.philos Botany	"The influence of environmental factors on the chemical composition of cultivated and natural populations of marine phytoplankton"
1980	Arnfinn Langeland	Dr. philos Zoology	Interaction between fish and zooplankton populations and their effects on the material utilization in a freshwater lake
1980	Helge Reinertsen	Dr. philos Botany	The effect of lake fertilization on the dynamics and stability of a limnetic ecosystem with special reference to the phytoplankton
1982	Gunn Mari Olsen	Dr. scient Botany	Gravitropism in roots of *Pisum sativum* and *Arabidopsis thaliana*
1982	Dag Dolmen	Dr. philos Zoology	Life aspects of two sympartic species of newts (*Triturus, Amphibia*) in Norway, with special emphasis on their ecological niche segregation
1984	Eivin Røskaft	Dr. philos Zoology	Sociobiological studies of the rook *Corvus frugilegus*
1984	Anne Margrethe Cameron	Dr. scient Botany	Effects of alcohol inhalation on levels of circulating testosterone, follicle stimulating hormone and luteinzing hormone in male mature rats
1984	Asbjørn Magne Nilsen	Dr. scient Botany	Alveolar macrophages from expectorates – Biological monitoring of workers exosed to occupational air pollution. An evaluation of the AM-test
1985	Jarle Mork	Dr. philos Zoology	Biochemical genetic studies in fish
1985	John Solem	Dr. philos Zoology	Taxonomy, distribution and ecology of caddisflies (*Trichoptera*) in the Dovrefjell mountains
1985	Randi E. Reinertsen	Dr. philos Zoology	Energy strategies in the cold: Metabolic and thermoregulatory adaptations in small northern birds
1986	Bernt-Erik Sæther	Dr. philos Zoology	Ecological and evolutionary basis for variation in reproductive traits of some vertebrates: A comparative approach
1986	Torleif Holthe	Dr. philos Zoology	Evolution, systematics, nomenclature, and zoogeography in the polychaete orders *Oweniimorpha* and *Terebellomorpha*, with special reference to the Arctic and Scandinavian fauna
1987	Helene Lampe	Dr. scient Zoology	The function of bird song in mate attraction and territorial defence, and the importance of song

			repertoires
1987	Olav Hogstad	Dr. philos Zoology	Winter survival strategies of the Willow tit *Parus montanus*
1987	Jarle Inge Holten	Dr. philos Botany	Autecological investigations along a coust-inland transect at Nord-Møre, Central Norway
1987	Rita Kumar	Dr. scient Botany	Somaclonal variation in plants regenerated from cell cultures of *Nicotiana sanderae* and *Chrysanthemum morifolium*
1987	Bjørn Åge Tømmerås	Dr. scient. Zoolog	Olfaction in bark beetle communities: Interspecific interactions in regulation of colonization density, predator - prey relationship and host attraction
1988	Hans Christian Pedersen	Dr. philos Zoology	Reproductive behaviour in willow ptarmigan with special emphasis on territoriality and parental care
1988	Tor G. Heggberget	Dr. philos Zoology	Reproduction in Atlantic Salmon (*Salmo salar*): Aspects of spawning, incubation, early life history and population structure
1988	Marianne V. Nielsen	Dr. scient Zoology	The effects of selected environmental factors on carbon allocation/growth of larval and juvenile mussels (*Mytilus edulis*)
1988	Ole Kristian Berg	Dr. scient Zoology	The formation of landlocked Atlantic salmon (*Salmo salar* L.)
1989	John W. Jensen	Dr. philos Zoology	Crustacean plankton and fish during the first decade of the manmade Nesjø reservoir, with special emphasis on the effects of gill nets and salmonid growth
1989	Helga J. Vivås	Dr. scient Zoology	Theoretical models of activity pattern and optimal foraging: Predictions for the Moose *Alces alces*
1989	Reidar Andersen	Dr. scient Zoology	Interactions between a generalist herbivore, the moose *Alces alces*, and its winter food resources: a study of behavioural variation
1989	Kurt Ingar Draget	Dr. scient Botany	Alginate gel media for plant tissue culture
1990	Bengt Finstad	Dr. scient Zoology	Osmotic and ionic regulation in Atlantic salmon, rainbow trout and Arctic charr: Effect of temperature, salinity and season
1990	Hege Johannesen	Dr. scient Zoology	Respiration and temperature regulation in birds with special emphasis on the oxygen extraction by the lung
1990	Åse Krøkje	Dr. scient Botany	The mutagenic load from air pollution at two work-places with PAH-exposure measured with Ames Salmonella/microsome test
1990	Arne Johan Jensen	Dr. philos Zoology	Effects of water temperature on early life history, juvenile growth and prespawning migrations of Atlantic salmion (*Salmo salar*) and brown trout (*Salmo trutta*): A summary of studies in Norwegian streams
1990	Tor Jørgen Almaas	Dr. scient Zoology	Pheromone reception in moths: Response characteristics of olfactory receptor neurons to intra- and interspecific chemical cues
1990	Magne Husby	Dr. scient Zoology	Breeding strategies in birds: Experiments with the Magpie *Pica pica*
1991	Tor Kvam	Dr. scient Zoology	Population biology of the European lynx (*Lynx lynx*) in Norway
1991	Jan Henning L'Abêe	Dr.	Reproductive biology in freshwater fish, brown trout

	Lund	philos Zoology	*Salmo trutta* and roach *Rutilus rutilus* in particular
1991	Asbjørn Moen	Dr. philos Botany	The plant cover of the boreal uplands of Central Norway. I. Vegetation ecology of Sølendet nature reserve; haymaking fens and birch woodlands
1991	Else Marie Løbersli	Dr. scient Botany	Soil acidification and metal uptake in plants
1991	Trond Nordtug	Dr. scient Zoology	Reflctometric studies of photomechanical adaptation in superposition eyes of arthropods
1991	Thyra Solem	Dr. scient Botany	Age, origin and development of blanket mires in Central Norway
1991	Odd Terje Sandlund	Dr. philos Zoology	The dynamics of habitat use in the salmonid genera *Coregonus* and *Salvelinus*: Ontogenic niche shifts and polymorphism
1991	Nina Jonsson	Dr. philos	Aspects of migration and spawning in salmonids
1991	Atle Bones	Dr. scient Botany	Compartmentation and molecular properties of thioglucoside glucohydrolase (myrosinase)
1992	Torgrim Breiehagen	Dr. scient Zoology	Mating behaviour and evolutionary aspects of the breeding system of two bird species: the Temminck's stint and the Pied flycatcher
1992	Anne Kjersti Bakken	Dr. scient Botany	The influence of photoperiod on nitrate assimilation and nitrogen status in timothy (*Phleum pratense* L.)
1992	Tycho Anker-Nilssen	Dr. scient Zoology	Food supply as a determinant of reproduction and population development in Norwegian Puffins *Fratercula arctica*
1992	Bjørn Munro Jenssen	Dr. philos Zoology	Thermoregulation in aquatic birds in air and water: With special emphasis on the effects of crude oil, chemically treated oil and cleaning on the thermal balance of ducks
1992	Arne Vollan Aarset	Dr. philos Zoology	The ecophysiology of under-ice fauna: Osmotic regulation, low temperature tolerance and metabolism in polar crustaceans.
1993	Geir Slupphaug	Dr. scient Botany	Regulation and expression of uracil-DNA glycosylase and O^6-methylguanine-DNA methyltransferase in mammalian cells
1993	Tor Fredrik Næsje	Dr. scient Zoology	Habitat shifts in coregonids.
1993	Yngvar Asbjørn Olsen	Dr. scient Zoology	Cortisol dynamics in Atlantic salmon, *Salmo salar* L.: Basal and stressor-induced variations in plasma levels ans some secondary effects.
1993	Bård Pedersen	Dr. scient Botany	Theoretical studies of life history evolution in modular and clonal organisms
1993	Ole Petter Thangstad	Dr. scient Botany	Molecular studies of myrosinase in Brassicaceae
1993	Thrine L. M. Heggberget	Dr. scient Zoology	Reproductive strategy and feeding ecology of the Eurasian otter *Lutra lutra*.
1993	Kjetil Bevanger	Dr. scient. Zoology	Avian interactions with utility structures, a biological approach.
1993	Kåre Haugan	Dr. scient Bothany	Mutations in the replication control gene trfA of the broad host-range plasmid RK2
1994	Peder Fiske	Dr. scient. Zoology	Sexual selection in the lekking great snipe (*Gallinago media*): Male mating success and female behaviour at the lek
1994	Kjell Inge Reitan	Dr. scient	Nutritional effects of algae in first-feeding of marine

		Botany	fish larvae
1994	Nils Røv	Dr. scient Zoology	Breeding distribution, population status and regulation of breeding numbers in the northeast-Atlantic Great Cormorant *Phalacrocorax carbo carbo*
1994	Annette-Susanne Hoepfner	Dr. scient Botany	Tissue culture techniques in propagation and breeding of Red Raspberry (*Rubus idaeus* L.)
1994	Inga Elise Bruteig	Dr. scient Bothany	Distribution, ecology and biomonitoring studies of epiphytic lichens on conifers
1994	Geir Johnsen	Dr. scient Botany	Light harvesting and utilization in marine phytoplankton: Species-specific and photoadaptive responses
1994	Morten Bakken	Dr. scient Zoology	Infanticidal behaviour and reproductive performance in relation to competition capacity among farmed silver fox vixens, *Vulpes vulpes*
1994	Arne Moksnes	Dr. philos Zoology	Host adaptations towards brood parasitism by the Cockoo
1994	Solveig Bakken	Dr. scient Bothany	Growth and nitrogen status in the moss *Dicranum majus* Sm. as influenced by nitrogen supply
1994	Torbjørn Forseth	Dr. scient Zoology	Bioenergetics in ecological and life history studies of fishes.
1995	Olav Vadstein	Dr. philos Botany	The role of heterotrophic planktonic bacteria in the cycling of phosphorus in lakes: Phosphorus requirement, competitive ability and food web interactions
1995	Hanne Christensen	Dr. scient Zoology	Determinants of Otter *Lutra lutra* distribution in Norway: Effects of harvest, polychlorinated biphenyls (PCBs), human population density and competition with mink *Mustela vision*
1995	Svein Håkon Lorentsen	Dr. scient Zoology	Reproductive effort in the Antarctic Petrel *Thalassoica antarctica*; the effect of parental body size and condition
1995	Chris Jørgen Jensen	Dr. scient Zoology	The surface electromyographic (EMG) amplitude as an estimate of upper trapezius muscle activity
1995	Martha Kold Bakkevig	Dr. scient Zoology	The impact of clothing textiles and construction in a clothing system on thermoregulatory responses, sweat accumulation and heat transport
1995	Vidar Moen	Dr. scient Zoology	Distribution patterns and adaptations to light in newly introduced populations of *Mysis relicta* and constraints on Cladoceran and Char populations
1995	Hans Haavardsholm Blom	Dr. philos Bothany	A revision of the *Schistidium apocarpum* complex in Norway and Sweden
1996	Jorun Skjærmo	Dr. scient Botany	Microbial ecology of early stages of cultivated marine fish; inpact fish-bacterial interactions on growth and survival of larvae
1996	Ola Ugedal	Dr. scient Zoology	Radiocesium turnover in freshwater fishes
1996	Ingibjørg Einarsdottir	Dr. scient Zoology	Production of Atlantic salmon (*Salmo salar*) and Arctic charr (*Salvelinus alpinus*): A study of some physiological and immunological responses to rearing routines
1996	Christina M. S. Pereira	Dr. scient Zoology	Glucose metabolism in salmonids: Dietary effects and hormonal regulation
1996	Jan Fredrik Børseth	Dr. scient Zoology	The sodium energy gradients in muscle cells of *Mytilus edulis* and the effects of organic xenobiotics

1996	Gunnar Henriksen	Dr. scient Zoology	Status of Grey seal *Halichoerus grypus* and Harbour seal *Phoca vitulina* in the Barents sea region
1997	Gunvor Øie	Dr. scient Bothany	Eevalution of rotifer *Brachionus plicatilis* quality in early first feeding of turbot *Scophtalmus maximus* L. larvae
1997	Håkon Holien	Dr. scient Botany	Studies of lichens in spurce forest of Central Norway. Diversity, old growth species and the relationship to site and stand parameters
1997	Ole Reitan	Dr. scient. Zoology	Responses of birds to habitat disturbance due to damming
1997	Jon Arne Grøttum	Dr. scient. Zoology	Physiological effects of reduced water quality on fish in aquaculture
1997	Per Gustav Thingstad	Dr. scient. Zoology	Birds as indicators for studying natural and human-induced variations in the environment, with special emphasis on the suitability of the Pied Flycatcher
1997	Torgeir Nygård	Dr. scient Zoology	Temporal and spatial trends of pollutants in birds in Norway: Birds of prey and Willow Grouse used as Biomonitors
1997	Signe Nybø	Dr. scient. Zoology	Impacts of long-range transported air pollution on birds with particular reference to the dipper *Cinclus cinclus* in southern Norway
1997	Atle Wibe	Dr. scient. Zoology	Identification of conifer volatiles detected by receptor neurons in the pine weevil (*Hylobius abietis*), analysed by gas chromatography linked to electrophysiology and to mass spectrometry
1997	Rolv Lundheim	Dr. scient Zoology	Adaptive and incidental biological ice nucleators
1997	Arild Magne Landa	Dr. scient Zoology	Wolverines in Scandinavia: ecology, sheep depredation and conservation
1997	Kåre Magne Nielsen	Dr. scient Botany	An evolution of possible horizontal gene transfer from plants to sail bacteria by studies of natural transformation in *Acinetobacter calcoacetius*
1997	Jarle Tufto	Dr. scient Zoology	Gene flow and genetic drift in geographically structured populations: Ecological, population genetic, and statistical models
1997	Trygve Hesthagen	Dr. philos Zoology	Population responces of Arctic charr (*Salvelinus alpinus* (L.)) and brown trout (*Salmo trutta* L.) to acidification in Norwegian inland waters
1997	Trygve Sigholt	Dr. philos Zoology	Control of Parr-smolt transformation and seawater tolerance in farmed Atlantic Salmon (*Salmo salar*) Effects of photoperiod, temperature, gradual seawater acclimation, NaCl and betaine in the diet
1997	Jan Østnes	Dr. scient Zoology	Cold sensation in adult and neonate birds
1998	Seethaledsumy Visvalingam	Dr. scient Botany	Influence of environmental factors on myrosinases and myrosinase-binding proteins
1998	Thor Harald Ringsby	Dr. scient Zoology	Variation in space and time: The biology of a House sparrow metapopulation
1998	Erling Johan Solberg	Dr. scient. Zoology	Variation in population dynamics and life history in a Norwegian moose (*Alces alces*) population: consequences of harvesting in a variable environment
1998	Sigurd Mjøen Saastad	Dr. scient Botany	Species delimitation and phylogenetic relationships between the Sphagnum recurvum complex (Bryophyta): genetic variation and phenotypic

			plasticity
1998	Bjarte Mortensen	Dr. scient Botany	Metabolism of volatile organic chemicals (VOCs) in a head liver S9 vial equilibration system in vitro
1998	Gunnar Austrheim	Dr. scient Botany	Plant biodiversity and land use in subalpine grasslands. – A conservtaion biological approach
1998	Bente Gunnveig Berg	Dr. scient Zoology	Encoding of pheromone information in two related moth species
1999	Kristian Overskaug	Dr. scient Zoology	Behavioural and morphological characteristics in Northern Tawny Owls *Strix aluco*: An intra- and interspecific comparative approach
1999	Hans Kristen Stenøien	Dr. scient Bothany	Genetic studies of evolutionary processes in various populations of nonvascular plants (mosses, liverworts and hornworts)
1999	Trond Arnesen	Dr. scient Botany	Vegetation dynamics following trampling and burning in the outlying haylands at Sølendet, Central Norway
1999	Ingvar Stenberg	Dr. scient Zoology	Habitat selection, reproduction and survival in the White-backed Woodpecker *Dendrocopos leucotos*
1999	Stein Olle Johansen	Dr. scient Botany	A study of driftwood dispersal to the Nordic Seas by dendrochronology and wood anatomical analysis
1999	Trina Falck Galloway	Dr. scient Zoology	Muscle development and growth in early life stages of the Atlantic cod (*Gadus morhua* L.) and Halibut (*Hippoglossus hippoglossus* L.)
1999	Marianne Giæver	Dr. scient Zoology	Population genetic studies in three gadoid species: blue whiting (*Micromisistius poutassou*), haddock (*Melanogrammus aeglefinus*) and cod (*Gradus morhua*) in the North-East Atlantic
1999	Hans Martin Hanslin	Dr. scient Botany	The impact of environmental conditions of density dependent performance in the boreal forest bryophytes *Dicranum majus*, *Hylocomium splendens*, *Plagiochila asplenigides*, *Ptilium crista-castrensis* and *Rhytidiadelphus lokeus*
1999	Ingrid Bysveen Mjølnerød	Dr. scient Zoology	Aspects of population genetics, behaviour and performance of wild and farmed Atlantic salmon (*Salmo salar*) revealed by molecular genetic techniques
1999	Else Berit Skagen	Dr. scient Botany	The early regeneration process in protoplasts from *Brassica napus* hypocotyls cultivated under various g-forces
1999	Stein-Are Sæther	Dr. philos Zoology	Mate choice, competition for mates, and conflicts of interest in the Lekking Great Snipe
1999	Katrine Wangen Rustad	Dr. scient Zoology	Modulation of glutamatergic neurotransmission related to cognitive dysfunctions and Alzheimer's disease
1999	Per Terje Smiseth	Dr. scient Zoology	Social evolution in monogamous families: mate choice and conflicts over parental care in the Bluethroat (*Luscinia s. svecica*)
1999	Gunnbjørn Bremset	Dr. scient Zoology	Young Atlantic salmon (*Salmo salar* L.) and Brown trout (*Salmo trutta* L.) inhabiting the deep pool habitat, with special reference to their habitat use, habitat preferences and competitive interactions
1999	Frode Ødegaard	Dr. scient Zoology	Host spesificity as parameter in estimates of arhrophod species richness
1999	Sonja Andersen	Dr. scient Bothany	Expressional and functional analyses of human, secretory phospholipase A2
2000	Ingrid Salvesen	Dr. scient	Microbial ecology in early stages of marine fish:

		Botany	Development and evaluation of methods for microbial management in intensive larviculture
2000	Ingar Jostein Øien	Dr. scient Zoology	The Cuckoo (*Cuculus canorus*) and its host: adaptions and counteradaptions in a coevolutionary arms race
2000	Pavlos Makridis	Dr. scient Botany	Methods for the microbial econtrol of live food used for the rearing of marine fish larvae
2000	Sigbjørn Stokke	Dr. scient Zoology	Sexual segregation in the African elephant (*Loxodonta africana*)
2000	Odd A. Gulseth	Dr. philos Zoology	Seawater tolerance, migratory behaviour and growth of Charr, (*Salvelinus alpinus*), with emphasis on the high Arctic Dieset charr on Spitsbergen, Svalbard
2000	Pål A. Olsvik	Dr. scient Zoology	Biochemical impacts of Cd, Cu and Zn on brown trout (*Salmo trutta*) in two mining-contaminated rivers in Central Norway
2000	Sigurd Einum	Dr. scient Zoology	Maternal effects in fish: Implications for the evolution of breeding time and egg size
2001	Jan Ove Evjemo	Dr. scient Zoology	Production and nutritional adaptation of the brine shrimp *Artemia* sp. as live food organism for larvae of marine cold water fish species
2001	Olga Hilmo	Dr. scient Botany	Lichen response to environmental changes in the managed boreal forset systems
2001	Ingebrigt Uglem	Dr. scient Zoology	Male dimorphism and reproductive biology in corkwing wrasse (*Symphodus melops* L.)
2001	Bård Gunnar Stokke	Dr. scient Zoology	Coevolutionary adaptations in avian brood parasites and their hosts
2002	Ronny Aanes	Dr. scient	Spatio-temporal dynamics in Svalbard reindeer (*Rangifer tarandus platyrhynchus*)
2002	Mariann Sandsund	Dr. scient Zoology	Exercise- and cold-induced asthma. Respiratory and thermoregulatory responses
2002	Dag-Inge Øien	Dr. scient Botany	Dynamics of plant communities and populations in boreal vegetation influenced by scything at Sølendet, Central Norway
2002	Frank Rosell	Dr. scient Zoology	The function of scent marking in beaver (*Castor fiber*)
2002	Janne Østvang	Dr. scient Botany	The Role and Regulation of Phospholipase A_2 in Monocytes During Atherosclerosis Development
2002	Terje Thun	Dr.philos Biology	Dendrochronological constructions of Norwegian conifer chronologies providing dating of historical material
2002	Birgit Hafjeld Borgen	Dr. scient Biology	Functional analysis of plant idioblasts (Myrosin cells) and their role in defense, development and growth
2002	Bård Øyvind Solberg	Dr. scient Biology	Effects of climatic change on the growth of dominating tree species along major environmental gradients
2002	Per Winge	Dr. scient Biology	The evolution of small GTP binding proteins in cellular organisms. Studies of RAC GTPases in *Arabidopsis thaliana* and the Ral GTPase from *Drosophila melanogaster*
2002	Henrik Jensen	Dr. scient Biology	Causes and consequenses of individual variation in fitness-related traits in house sparrows
2003	Jens Rohloff	Dr. philos Biology	Cultivation of herbs and medicinal plants in Norway – Essential oil production and quality control
2003	Åsa Maria O. Espmark Wibe	Dr. scient Biology	Behavioural effects of environmental pollution in threespine stickleback *Gasterosteus aculeatur* L.
2003	Dagmar Hagen	Dr. scient	Assisted recovery of disturbed arctic and alpine

		Biology	vegetation – an integrated approach
2003	Bjørn Dahle	Dr. scient Biology	Reproductive strategies in Scandinavian brown bears
2003	Cyril Lebogang Taolo	Dr. scient Biology	Population ecology, seasonal movement and habitat use of the African buffalo (*Syncerus caffer*) in Chobe National Park, Botswana
2003	Marit Stranden	Dr.scient Biology	Olfactory receptor neurones specified for the same odorants in three related Heliothine species (*Helicoverpa armigera, Helicoverpa assulta* and *Heliothis virescens*)
2003	Kristian Hassel	Dr.scient Biology	Life history characteristics and genetic variation in an expanding species, *Pogonatum dentatum*
2003	David Alexander Rae	Dr.scient Biology	Plant- and invertebrate-community responses to species interaction and microclimatic gradients in alpine and Artic environments
2003	Åsa A Borg	Dr.scient Biology	Sex roles and reproductive behaviour in gobies and guppies: a female perspective
2003	Eldar Åsgard Bendiksen	Dr.scient Biology	Environmental effects on lipid nutrition of farmed Atlantic salmon (*Salmo Salar* L.) parr and smolt
2004	Torkild Bakken	Dr.scient Biology	A revision of Nereidinae (Polychaeta, Nereididae)
2004	Ingar Pareliussen	Dr.scient Biology	Natural and Experimental Tree Establishment in a Fragmented Forest, Ambohitantely Forest Reserve, Madagascar
2004	Tore Brembu	Dr.scient Biology	Genetic, molecular and functional studies of RAC GTPases and the WAVE-like regulatory protein complex in *Arabidopsis thaliana*
2004	Liv S. Nilsen	Dr.scient Biology	Coastal heath vegetation on central Norway; recent past, present state and future possibilities
2004	Hanne T. Skiri	Dr.scient Biology	Olfactory coding and olfactory learning of plant odours in heliothine moths. An anatomical, physiological and behavioural study of three related species (*Heliothis virescens, Helicoverpa armigera* and *Helicoverpa assulta*)
2004	Lene Østby	Dr.scient Biology	Cytochrome P4501A (CYP1A) induction and DNA adducts as biomarkers for organic pollution in the natural environment
2004	Emmanuel J. Gerreta	Dr. philos Biology	The Importance of Water Quality and Quantity in the Tropical Ecosystems, Tanzania
2004	Linda Dalen	Dr.scient Biology	Dynamics of Mountain Birch Treelines in the Scandes Mountain Chain, and Effects of Climate Warming
2004	Lisbeth Mehli	Dr.scient Biology	Polygalacturonase-inhibiting protein (PGIP) in cultivated strawberry (*Fragaria* x *ananassa*): characterisation and induction of the gene following fruit infection by *Botrytis cinerea*
2004	Børge Moe	Dr.scient Biology	Energy-Allocation in Avian Nestlings Facing Short-Term Food Shortage
2005	Matilde Skogen Chauton	Dr.scient Biology	Metabolic profiling and species discrimination from High-Resolution Magic Angle Spinning NMR analysis of whole-cell samples
2005	Sten Karlsson	Dr.scient Biology	Dynamics of Genetic Polymorphisms
2005	Terje Bongard	Dr.scient Biology	Life History strategies, mate choice, and parental investment among Norwegians over a 300-year period
2005	Tonette Røstelien	ph.d	Functional characterisation of olfactory receptor

		Biology	neurone types in heliothine moths
2005	Erlend Kristiansen	Dr.scient Biology	Studies on antifreeze proteins
2005	Eugen G. Sørmo	Dr.scient Biology	Organochlorine pollutants in grey seal (*Halichoerus grypus*) pups and their impact on plasma thyrid hormone and vitamin A concentrations
2005	Christian Westad	Dr.scient Biology	Motor control of the upper trapezius
2005	Lasse Mork Olsen	ph.d Biology	Interactions between marine osmo- and phagotrophs in different physicochemical environments
2005	Åslaug Viken	ph.d Biology	Implications of mate choice for the management of small populations
2005	Ariaya Hymete Sahle Dingle	ph.d Biology	Investigation of the biological activities and chemical constituents of selected *Echinops* spp. growing in Ethiopia
2005	Anders Gravbrøt Finstad	ph.d Biology	Salmonid fishes in a changing climate: The winter challenge
2005	Shimane Washington Makabu	ph.d Biology	Interactions between woody plants, elephants and other browsers in the Chobe Riverfront, Botswana
2005	Kjartan Østbye	Dr.scient Biology	The European whitefish *Coregonus lavaretus* (L.) species complex: historical contingency and adaptive radiation
2006	Kari Mette Murvoll	ph.d Biology	Levels and effects of persistent organic pollutans (POPs) in seabirds Retinoids and α-tocopherol – potential biomakers of POPs in birds?
2006	Ivar Herfindal	Dr.scient Biology	Life history consequences of environmental variation along ecological gradients in northern ungulates
2006	Nils Egil Tokle	ph.d Biology	Are the ubiquitous marine copepods limited by food or predation? Experimental and field-based studies with main focus on *Calanus finmarchicus*
2006	Jan Ove Gjershaug	Dr.philos Biology	Taxonomy and conservation status of some booted eagles in south-east Asia
2006	Jon Kristian Skei	Dr.scient Biology	Conservation biology and acidification problems in the breeding habitat of amphibians in Norway
2006	Johanna Järnegren	ph.d Biology	Acesta Oophaga and Acesta Excavata – a study of hidden biodiversity
2006	Bjørn Henrik Hansen	ph.d Biology	Metal-mediated oxidative stress responses in brown trout (*Salmo trutta*) from mining contaminated rivers in Central Norway
2006	Vidar Grøtan	ph.d Biology	Temporal and spatial effects of climate fluctuations on population dynamics of vertebrates
2006	Jafari R Kideghesho	ph.d Biology	Wildlife conservation and local land use conflicts in western Serengeti, Corridor Tanzania
2006	Anna Maria Billing	ph.d Biology	Reproductive decisions in the sex role reversed pipefish *Syngnathus typhle*: when and how to invest in reproduction
2006	Henrik Pärn	ph.d Biology	Female ornaments and reproductive biology in the bluethroat
2006	Anders J. Fjellheim	ph.d Biology	Selection and administration of probiotic bacteria to marine fish larvae
2006	P. Andreas Svensson	ph.d Biology	Female coloration, egg carotenoids and reproductive success: gobies as a model system
2007	Sindre A. Pedersen	ph.d Biology	Metal binding proteins and antifreeze proteins in the beetle *Tenebrio molitor* - a study on possible competition for the semi-

			essential amino acid cysteine
2007	Kasper Hancke	ph.d Biology	Photosynthetic responses as a function of light and temperature: Field and laboratory studies on marine microalgae
2007	Tomas Holmern	ph.d Biology	Bushmeat hunting in the western Serengeti: Implications for community-based conservation
2007	Kari Jørgensen	ph.d Biology	Functional tracing of gustatory receptor neurons in the CNS and chemosensory learning in the moth *Heliothis virescens*
2007	Stig Ulland	ph.d Biology	Functional Characterisation of Olfactory Receptor Neurons in the Cabbage Moth, (*Mamestra brassicae* L.) (Lepidoptera, Noctuidae). Gas Chromatography Linked to Single Cell Recordings and Mass Spectrometry
2007	Snorre Henriksen	ph.d Biology	Spatial and temporal variation in herbivore resources at northern latitudes
2007	Roelof Frans May	ph.d Biology	Spatial Ecology of Wolverines in Scandinavia
2007	Vedasto Gabriel Ndibalema	ph.d Biology	Demographic variation, distribution and habitat use between wildebeest sub-populations in the Serengeti National Park, Tanzania
2007	Julius William Nyahongo	ph.d Biology	Depredation of Livestock by wild Carnivores and Illegal Utilization of Natural Resources by Humans in the Western Serengeti, Tanzania
2007	Shombe Ntaraluka Hassan	ph.d Biology	Effects of fire on large herbivores and their forage resources in Serengeti, Tanzania
2007	Per-Arvid Wold	ph.d Biology	Functional development and response to dietary treatment in larval Atlantic cod *(Gadus morhua* L.) Focus on formulated diets and early weaning
2007	Anne Skjetne Mortensen	ph.d Biology	Toxicogenomics of Aryl Hydrocarbon- and Estrogen Receptor Interactions in Fish: Mechanisms and Profiling of Gene Expression Patterns in Chemical Mixture Exposure Scenarios
2008	Brage Bremset Hansen	ph.d Biology	The Svalbard reindeer (*Rangifer tarandus platyrhynchus*) and its food base: plant-herbivore interactions in a high-arctic ecosystem
2008	Jiska van Dijk	ph.d Biology	Wolverine foraging strategies in a multiple-use landscape
2008	Flora John Magige	ph.d Biology	The ecology and behaviour of the Masai Ostrich (Struthio camelus massaicus) in the Serengeti Ecosystem, Tanzania
2008	Bernt Rønning	ph.d Biology	Sources of inter- and intra-individual variation in basal metabolic rate in the zebra finch, (*Taeniopygia guttata*)
2008	Sølvi Wehn	ph.d Biology	Biodiversity dynamics in semi-natural mountain landscapes. - A study of consequences of changed agricultural practices in Eastern Jotunheimen
2008	Trond Moxness Kortner	ph.d Biology	"The Role of Androgens on previtellogenic oocyte growth in Atlantic cod (*Gadus morhua*): Identification and patterns of differentially expressed genes in relation to Stereological Evaluations"
2008	Katarina Mariann Jørgensen	Dr.Scient Biology	The role of platelet activating factor in activation of growth arrested keratinocytes and re-epithelialisation

2008	Tommy Jørstad	ph.d Biology	Statistical Modelling of Gene Expression Data
2008	Anna Kusnierczyk	ph.d Bilogy	*Arabidopsis thaliana* Responses to Aphid Infestation
2008	Jussi Evertsen	ph.d Biology	Herbivore sacoglossans with photosynthetic chloroplasts
2008	John Eilif Hermansen	ph.d Biology	Mediating ecological interests between locals and globals by means of indicators. A study attributed to the asymmetry between stakeholders of tropical forest at Mt. Kilimanjaro, Tanzania
2008	Ragnhild Lyngved	ph.d Biology	Somatic embryogenesis in *Cyclamen persicum.* Biological investigations and educational aspects of cloning
2008	Line Elisabeth Sundt-Hansen	ph.d Biology	Cost of rapid growth in salmonid fishes
2008	Line Johansen	ph.d Biology	Exploring factors underlying fluctuations in white clover populations – clonal growth, population structure and spatial distribution
2009	Astrid Jullumstrø Feuerherm	ph.d Biology	Elucidation of molecular mechanisms for pro-inflammatory phospholipase A2 in chronic disease
2009	Pål Kvello	ph.d Biology	Neurons forming the network involved in gustatory coding and learning in the moth *Heliothis virescens:* Physiological and morphological characterisation, and integration into a standard brain atlas
2009	Trygve Devold Kjellsen	ph.d Biology	Extreme Frost Tolerance in Boreal Conifers
2009	Johan Reinert Vikan	ph.d Biology	Coevolutionary interactions between common cuckoos *Cuculus canorus* and *Fringilla* finches
2009	Zsolt Volent	ph.d Biology	Remote sensing of marine environment: Applied surveillance with focus on optical properties of phytoplankton, coloured organic matter and suspended matter
2009	Lester Rocha	ph.d Biology	Functional responses of perennial grasses to simulated grazing and resource availability
2009	Dennis Ikanda	ph.d Biology	Dimensions of a Human-lion conflict: Ecology of human predation and persecution of African lions (*Panthera leo*) in Tanzania
2010	Huy Quang Nguyen	ph.d Biology	Egg characteristics and development of larval digestive function of cobia (*Rachycentron canadum*) in response to dietary treatments -Focus on formulated diets
2010	Eli Kvingedal	ph.d Biology	Intraspecific competition in stream salmonids: the impact of environment and phenotype
2010	Sverre Lundemo	ph.d Biology	Molecular studies of genetic structuring and demography in *Arabidopsis* from Northern Europe
2010	Iddi Mihijai Mfunda	ph.d Biology	Wildlife Conservation and People's livelihoods: Lessons Learnt and Considerations for Improvements. Tha Case of Serengeti Ecosystem, Tanzania
2010	Anton Tinchov Antonov	ph.d Biology	Why do cuckoos lay strong-shelled eggs? Tests of the puncture resistance hypothesis
2010	Anders Lyngstad	ph.d Biology	Population Ecology of *Eriophorum latifolium*, a Clonal Species in Rich Fen Vegetation
2010	Hilde Færevik	ph.d Biology	Impact of protective clothing on thermal and cognitive responses
2010	Ingerid Brænne Arbo	ph.d	Nutritional lifestyle changes – effects of dietary

		Medical technology	carbohydrate restriction in healthy obese and overweight humans
2010	Yngvild Vindenes	ph.d Biology	Stochastic modeling of finite populations with individual heterogeneity in vital parameters
2010	Hans-Richard Brattbakk	ph.d Medical technology	The effect of macronutrient composition, insulin stimulation, and genetic variation on leukocyte gene expression and possible health benefits
2011	Geir Hysing Bolstad	ph.d Biology	Evolution of Signals: Genetic Architecture, Natural Selection and Adaptive Accuracy
2011	Karen de Jong	ph.d Biology	Operational sex ratio and reproductive behaviour in the two-spotted goby (*Gobiusculus flavescens*)
2011	Ann-Iren Kittang	ph.d Biology	*Arabidopsis thaliana* L. adaptation mechanisms to microgravity through the EMCS MULTIGEN-2 experiment on the ISS:– The science of space experiment integration and adaptation to simulated microgravity
2011	Aline Magdalena Lee	ph.d Biology	Stochastic modeling of mating systems and their effect on population dynamics and genetics
2011	Christopher Gravningen Sørmo	ph.d Biology	Rho GTPases in Plants: Structural analysis of ROP GTPases; genetic and functional studies of MIRO GTPases in *Arabidopsis thaliana*
2011	Grethe Robertsen	ph.d Biology	Relative performance of salmonid phenotypes across environments and competitive intensities
2011	Line-Kristin Larsen	ph.d Biology	Life-history trait dynamics in experimental populations of guppy (*Poecilia reticulata*): the role of breeding regime and captive environment
2011	Maxim A. K. Teichert	ph.d Biology	Regulation in Atlantic salmon (*Salmo salar*): The interaction between habitat and density
2011	Torunn Beate Hancke	ph.d Biology	Use of Pulse Amplitude Modulated (PAM) Fluorescence and Bio-optics for Assessing Microalgal Photosynthesis and Physiology
2011	Sajeda Begum	ph.d Biology	Brood Parasitism in Asian Cuckoos: Different Aspects of Interactions between Cuckoos and their Hosts in Bangladesh
2011	Kari J. K. Attramadal	ph.d Biology	Water treatment as an approach to increase microbial control in the culture of cold water marine larvae
2011	Camilla Kalvatn Egset	ph.d Biology	The Evolvability of Static Allometry: A Case Study
2011	AHM Raihan Sarker	ph.d Biology	Conflict over the conservation of the Asian elephant (*Elephas maximus*) in Bangladesh
2011	Gro Dehli Villanger	ph.d Biology	Effects of complex organohalogen contaminant mixtures on thyroid hormone homeostasis in selected arctic marine mammals
2011	Kari Bjørneraas	ph.d Biology	Spatiotemporal variation in resource utilisation by a large herbivore, the moose
2011	John Odden	ph.d Biology	The ecology of a conflict: Eurasian lynx depredation on domestic sheep
2011	Simen Pedersen	ph.d Biology	Effects of native and introduced cervids on small mammals and birds
2011	Mohsen Falahati-Anbaran	ph.d Biology	Evolutionary consequences of seed banks and seed dispersal in *Arabidopsis*
2012	Jakob Hønborg Hansen	ph.d Biology	Shift work in the offshore vessel fleet: circadian rhythms and cognitive performance

2012	Elin Noreen	ph.d Biology	Consequences of diet quality and age on life-history traits in a small passerine bird
2012	Irja Ida Ratikainen	ph.d Biology	Theoretical and empirical approaches to studying foraging decisions: the past and future of behavioural ecology
2012	Aleksander Handå	ph.d Biology	Cultivation of mussels (*Mytilus edulis):*Feed requirements, storage and integration with salmon (*Salmo salar*) farming
2012	Morten Kraabøl	ph.d Biology	Reproductive and migratory challenges inflicted on migrant brown trour (*Salmo trutta* L) in a heavily modified river
2012	Jisca Huisman	ph.d Biology	Gene flow and natural selection in Atlantic salmon
2012	Maria Bergvik	ph.d Biology	Lipid and astaxanthin contents and biochemical post-harvest stability in *Calanus finmarchicus*
2012	Bjarte Bye Løfaldli	ph.d Biology	Functional and morphological characterization of central olfactory neurons in the model insect *Heliothis virescens*.
2012	Karen Marie Hammer	ph.d Biology.	Acid-base regulation and metabolite responses in shallow- and deep-living marine invertebrates during environmental hypercapnia
2012	Øystein Nordrum Wiggen	ph.d Biology	Optimal performance in the cold
2012	Robert Dominikus Fyumagwa	Dr. Philos.	Anthropogenic and natural influence on disease prevalence at the human –livestock-wildlife interface in the Serengeti ecosystem, Tanzania
2012	Jenny Bytingsvik	ph.d Biology	Organohalogenated contaminants (OHCs) in polar bear mother-cub pairs from Svalbard, Norway Maternal transfer, exposure assessment and thyroid hormone disruptive effects in polar bear cubs
2012	Christer Moe Rolandsen	ph.d Biology	The ecological significance of space use and movement patterns of moose in a variable environment
2012	Erlend Kjeldsberg Hovland	ph.d Biology	Bio-optics and Ecology in *Emiliania huxleyi* Blooms: Field and Remote Sensing Studies in Norwegian Waters
2012	Lise Cats Myhre	ph.d Biology	Effects of the social and physical environment on mating behaviour in a marine fish
2012	Tonje Aronsen	ph.d Biology	Demographic, environmental and evolutionary aspects of sexual selection
2012	Bin Liu	ph.d Biology	Molecular genetic investigation of cell separation and cell death regulation in *Arabidopsis thaliana*
2013	Jørgen Rosvold	ph.d Biology	Ungulates in a dynamic and increasingly human dominated landscape – A millennia-scale perspective
2013	Pankaj Barah	ph.d Biology (Systems Biology)	Integrated Systems Approaches to Study Plant Stress Responses