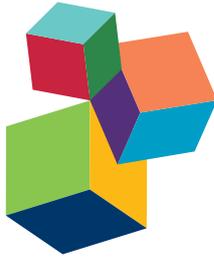


PLANT SINGLE CELL TYPE SYSTEMS BIOLOGY

EDITED BY : Marc Libault and Sixue Chen
PUBLISHED IN : Frontiers in Plant Science



frontiers

Frontiers Copyright Statement

© Copyright 2007-2016 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88919-948-8

DOI 10.3389/978-2-88919-948-8

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

PLANT SINGLE CELL TYPE SYSTEMS BIOLOGY

Topic Editors:

Marc Libault, University of Oklahoma, USA

Sixue Chen, University of Florida, USA

Citation: Libault, M., Chen, S., eds. (2016). Plant Single Cell Type Systems Biology. Lausanne: Frontiers Media. doi: 10.3389/978-2-88919-948-8

Table of Contents

05 Editorial: Plant Single Cell Type Systems Biology

Marc Libault and Sixue Chen

Chapter 1. Plant Single Cell Type Models for –omics and systems biology approaches

07 Molecular phenotyping of plant single cell-types enhances forward genetic analyses

John Schiefelbein

11 System approaches to study root hairs as a single cell plant model: current status and future perspectives

Md Shakhawat Hossain, Trupti Joshi and Gary Stacey

18 The female gametophyte: an emerging model for cell type-specific systems biology in plant development

Marc W. Schmid, Anja Schmidt and Ueli Grossniklaus

36 Methods to isolate a large amount of generative cells, sperm cells and vegetative nuclei from tomato pollen for “omics” analysis

Yunlong Lu, Liqin Wei and Tai Wang

Chapter 2. Dissection of the plant single cell type transcriptional activity

45 Re-analysis of RNA-seq transcriptome data reveals new aspects of gene activity in Arabidopsis root hairs

Wenfeng Li and Ping Lan

59 Spatial dissection of the Arabidopsis thaliana transcriptional response to downy mildew using Fluorescence Activated Cell Sorting

Timothy L. R. Coker, Volkan Cevik, Jim L. Beynon and Miriam L. Gifford

72 Identification of a core set of rhizobial infection genes using data from single cell-types

Da-Song Chen, Cheng-Wu Liu, Sonali Roy, Donna Cousins, Nicola Stacey and Jeremy D. Murray

Chapter 3. Biochemical analysis of plant single cell types

83 Spatially resolved in vivo plant metabolomics by laser ablation-based mass spectrometry imaging (MSI) techniques: LDI-MSI and LAESI

Benjamin Bartels and Aleš Svatoš

90 Cytological and proteomic analyses of horsetail (Equisetum arvense L.) spore germination

Qi Zhao, Jing Gao, Jinwei Suo, Sixue Chen, Tai Wang and Shaojun Dai

- 110** *Proteasome targeting of proteins in Arabidopsis leaf mesophyll, epidermal and vascular tissues*
Julia Svozil, Wilhelm Gruissem and Katja Baerenfaller
- 127** *The guard cell metabolome: functions in stomatal movement and global food security*
Biswapriya B. Misra, Biswa R. Acharya, David Granot, Sarah M. Assmann and Sixue Chen
- 140** *Single cell-type comparative metabolomics of epidermal bladder cells from the halophyte Mesembryanthemum crystallinum*
Bronwyn J. Barkla and Rosario Vera-Estrella



Editorial: Plant Single Cell Type Systems Biology

Marc Libault^{1*} and Sixue Chen²

¹ Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK, USA, ² Department of Biology, Interdisciplinary Center for Biotechnology Research, Genetics Institute, Plant Molecular and Cellular Biology Program, University of Florida, Gainesville, FL, USA

Keywords: systems biology, molecular phenotype, omic analyses, single cell types, root hair, trichome

Editorial on the Research Topic

Plant Single Cell Type Systems Biology

The molecular responses of a plant to a stress and the molecular profiles of plant organs during their development and differentiation are the reflection of the contribution from different cell types composing the plant and the organs. Hence, a major limitation to understanding plant cellular and molecular responses in different cells is the multicellular complexity of the plant or organs used to decipher them. For instance, as mentioned by Coker et al. changes in the expression levels of plant cells infected by pathogenic microbial organisms are diluted by the relative abundance of uninfected cells. This constraint has led plant biologists to select model single plant cell types such as pollen, trichomes, cotton fiber, guard cells of stomata, and various root cell types including the root hair cells, and to develop new technologies (e.g., microscopic, biochemical, and omics) to decipher their biology (Dai and Chen, 2012; Misra et al.). However, as mentioned by Schmid et al. (2015) profiling single cell types is dependent on the quantity and purity of the samples isolated as well as the use of sensitive and accurate profiling methods. The 12 articles published in this Research Topic highlight interesting methodology and biological systems applied by plant scientists to advance our knowledge in plant biology using single cell type models.

Working at the level of single cell types is motivated by the need for analyzing specific biological information in the relevant cell types, which would otherwise be missed when using tissues or organs (Dai and Chen, 2012; Misra et al.). This is especially true when working on plant-microbe interactions where only a subset of cells are infected by pathogenic or mutualistic microbes. For instance, to precisely characterize the transcriptional response of *Arabidopsis thaliana* during infection by the oomycete *Hyaloperonospora arabidopsidis* (*Hpa*), Coker et al. applied Fluorescent Activated Cell Sorting (FACS) in separation of haustoriated and non-haustoriated *Arabidopsis* cells for transcriptomic analysis, allowing the discovery of 139 new *Hpa*-responsive genes and characterization of the local and systemic responses of the plant cells. Similarly, working on the infection of the soybean root hair cells by rhizobium, the nitrogen-fixing symbiotic soil bacterium, Hossain et al. described the integration of the transcriptomic, proteomic, phosphoproteomic, and metabolomic datasets to generate a comprehensive network of the early stage of the nodulation process. Another strategy applied by Chen et al. to gain a better understanding of the nodulation process is to compare the transcriptomes of different rhizobium-infected plant cell types. Specifically, they looked for the *Medicago truncatula* genes controlling infection thread formation and elongation by analyzing transcriptomic data obtained from inoculated root hair cells and the infection zone of the *M. truncatula* nodule. Studying plant reproduction, another complex biological process, can also benefit from single cell type analyses. Schmid et al. detailed novel methods to analyze single cell type molecular profiles such as the female gametophyte, which is composed of antipodal, central, egg, and synergid cells. Similarly, working on the male

OPEN ACCESS

Edited by:

Joshua L. Heazlewood,
The University of Melbourne, Australia

Reviewed by:

Berit Ebert,
The University of Melbourne, Australia

*Correspondence:

Marc Libault
libaultm@ou.edu

Specialty section:

This article was submitted to
Plant Systems and Synthetic Biology,
a section of the journal
Frontiers in Plant Science

Received: 20 November 2015

Accepted: 11 January 2016

Published: 09 February 2016

Citation:

Libault M and Chen S (2016) Editorial:
Plant Single Cell Type Systems
Biology. *Front. Plant Sci.* 7:35.
doi: 10.3389/fpls.2016.00035

gametophyte, Lu et al. developed a pollen culture system for isolating generative cells, sperm cells and vegetative nuclei from tomato pollen grains. Plant reproduction studies can also benefit from the utility of unique single cell type models, such as *Equisetum arvense*, an herbaceous plant characterized by its spore reproduction. Zhao et al. analyzed the cellular and proteomic profiles of *E. arvense* during spore germination, revealing the high level activities of the heterotrophic and autotrophic metabolisms.

The generation of unambiguous datasets from single cell types is an asset for generating systems biology models as demonstrated by Kwak et al. (2008), Sun et al. (2014) and Hossain et al. Single cell types are also considered attractive systems to precisely depict molecular phenotypes. As noted by Schiefelbein, access to single cell types now opens a new area to phenotype mutants: the establishment of molecular phenotypes (i.e., distinct molecular profiles between wild-type and mutants and their changes in response to environmental stresses). Such an approach is often limited by efficient methods to generate high quality single plant cell type samples and by the limited amount of material available for analyzing the molecular phenotype. Thus, technological development must continue to meet the needs of addressing questions at the single cell type level. Nucleic acid sequencing technologies associated with the use of performant bioinformatics tools are now enabling an accurate and sensitive quantification of single cell type transcriptomes and epigenomes. As an example, the analysis of previously published Arabidopsis root hair transcriptome data sets allowed the characterization of 5409 genes differentially expressed in root hairs versus non-root hair epidermal cells and the generation of a co-expression network (Li and Lan). Similarly, biochemical methods are quickly developing allowing access to single plant cell type proteome (Svozil et al.) and metabolome (Barkla and Vera-Estrella; Bartels and Svatos; Misra et al.). Specifically, Barkla and Vera-Estrella described the differential metabolome between specialized trichome cells from *Mesembryanthemum crystallinum* named epidermal bladder cells (EBC). This analysis can be expected to provide a systems level of understanding of EBC when integrated with the existing

proteomic and transcriptomic data sets. Similarly, Misra et al. reviewed the most recent advances in our understanding of the guard cell metabolome. This knowledge is essential to advance our understanding of stomatal opening and closing, which have a major impact on plant transpiration, CO₂ uptake and pathogen immunity. At the proteome level, Svozil et al. applied Meselect, an innovative methodology to isolate leaf epidermal, vascular and mesophyll cells. Using these samples, the authors established a proteome map of each cell type and revealed cell type specific processes. These types of studies are going to be revolutionized by the development of new imaging techniques. For instance, applying infrared-laser ablation electrospray ionization (LAESI) and UV-laser desorption/ionization (LDI) methods, less intrusive and spatially-resolved analyses of the metabolomes of single plant cell types are described in this ebook (Bartels and Svatos). These technological developments have greatly enhanced our capabilities in analyzing molecular components in different cells at an unprecedented scope and depth through omics for modeling and hypothesis generation. The integration of hypothesis generation and hypothesis testing in systems biology research will ultimately lead to a holistic view of cellular processes and molecular networks in plants and will create stepping stones toward molecular breeding and biotechnology for enhanced crop stress tolerance, yield and bioenergy.

AUTHOR CONTRIBUTIONS

ML drafted the manuscript. SC edited the manuscript.

FUNDING

Research on plant single cell-type regulatory networks in the Libault laboratory has been supported by NSF grants IOS-1453613, IOS-1339194, by DOE grant DE-SC0012629, and by the Oklahoma Center for the Advancement of Science and Technology PS14-025 to ML. Research on single cell-type proteomics and metabolomics in the Chen laboratory has been supported by NSF grants MCB-0818051, MCB-1158000, and MCB-1412547 to SC.

REFERENCES

- Dai, S., and Chen, S. (2012). Single-cell-type proteomics: toward a holistic understanding of plant function. *Mol. Cell Proteomics* 11, 1622–1630. doi: 10.1074/mcp.R112.021550
- Kwak, J. M., Mäser, P., and Schroeder, J. I. (2008). The clickable guard cell, interactive model of guard cell signal transduction mechanisms and pathways. *Arabidopsis Book* 6:e0114. doi: 10.1199/tab.0114
- Schmid, M. W., Grob, S., and Grossniklaus, U. (2015). HiCdat: a fast and easy-to-use Hi-C data analysis tool. *BMC Bioinformatics* 16:277. doi: 10.1186/s12859-015-0678-x
- Sun, Z., Jin, X., Albert, R., and Assmann, S. M. (2014). Multi-level modeling of light-induced stomatal opening offers new insights into its regulation by drought. *PLoS Comput Biol.* 10:e1003930. doi: 10.1371/journal.pcbi.1003930

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer BE and handling Editor declared their shared affiliation, and the handling editor states that the process nevertheless met the standards of a fair and objective review.

Copyright © 2016 Libault and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Molecular phenotyping of plant single cell-types enhances forward genetic analyses

John Schiefelbein*

Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, MI, USA

Recent advances in the isolation of single cell-types in plants provides an opportunity to conduct detailed analyses of their molecular characteristics at high resolution. This kind of cell-type specific molecular phenotyping is likely to enhance forward genetics studies to dissect the effect of mutations and thereby aid gene function assignment. Recent experimental results support this view, demonstrating that different cell-types exhibit substantial variation in transcript, protein, and metabolite accumulation and these molecular phenotypes are often sensitive to genetic and environmental alterations. The use of single cell-type molecular phenotyping approach to define plant gene function is most amenable to cell-types with well-characterized molecular tools and isolation protocols.

OPEN ACCESS

Edited by:

Marc Libault,
University of Oklahoma, USA

Reviewed by:

Stefan Kempa,
Helmholtz Association, Germany
Jiangxin Wang,
Arizona State University, USA

*Correspondence:

John Schiefelbein,
Department of Molecular, Cellular,
and Developmental Biology, University
of Michigan, 830 North University
Avenue, Ann Arbor, MI 48109, USA
schiefel@umich.edu

Specialty section:

This article was submitted to
Plant Systems and Synthetic Biology,
a section of the journal
Frontiers in Plant Science

Received: 01 May 2015

Accepted: 25 June 2015

Published: 06 July 2015

Citation:

Schiefelbein J (2015) Molecular
phenotyping of plant single cell-types
enhances forward genetic analyses.
Front. Plant Sci. 6:509.
doi: 10.3389/fpls.2015.00509

Keywords: genetic analysis, mutant, molecular phenotype, cell biology, transcriptomics

Introduction

The use of forward genetic analysis has historically been an effective approach for defining gene function. By analyzing the phenotypic effect of a mutated gene, it has been possible to assign hundreds of genes in numerous organisms to particular developmental or metabolic pathways. However, a limitation of this approach is its reliance on the detection and measurement of phenotypic alterations. Indeed, if one cannot observe a change in the phenotype of a genetically altered individual, then little insight is gained concerning the affected gene's role in the process of interest.

The recent development of approaches to isolate and analyze plant single cell-types are likely to help overcome this limitation of forward genetics in plant biology. Specifically, recent studies suggest that individual plant cell-types exhibit distinct molecular characteristics (molecular phenotypes) and that these are differentially affected by genetic or environmental changes. Thus, forward genetic analyses that focus on phenotyping single cell-types are expected to have the capacity to detect and measure molecular alterations and thereby provide greater insight into gene function.

This perspective article summarizes recent studies showing distinct molecular phenotypes in plant single cell-types, their sensitivity to genetic and environmental perturbation, and evidence that forward genetic analyses are aided by single cell-type studies. This article also examines the factors likely to be critical for the successful use of molecular phenotyping to study gene function at the level of plant single cell-types.

Molecular Phenotypes in Single Cell-Types of Plants

The molecular analysis of a bulk cell population yields the composite molecular phenotype of many cell and tissue types, leading to an average assessment of the transcriptome, proteome, or

metabolome of the population. Although such bulk data are useful for some purposes, recent studies have shown that individual cell-types within a plant organ possess widely disparate molecular characteristics (Brandt, 2005). For example, extensive microarray analyses of cell-sorted *Arabidopsis* roots show cell-type specific transcript accumulation that differs substantially from the total root RNA population (Birnbaum et al., 2003; Brady et al., 2007). Similar cell-specific expression has been observed in several different isolated cell-types from rice, yielding a “cell-type transcriptome atlas” (Jiao et al., 2009). Further, it has been shown that distinct protein and metabolite accumulation patterns exist in different root cell populations as compared with the entire root organ (Petricka et al., 2012; Moussaieff et al., 2013). Although the evidence for cell-type specific accumulation of biological molecules is greatest for roots, the analysis of other isolated cell-types of plants, including trichomes and stomata, also reveals transcriptome patterns distinct from the larger organ-level accumulation patterns (Lieckfeldt et al., 2008; Adrian et al., 2015).

Furthermore, it is now clear that the molecular phenotype of a given cell-type is altered in a unique manner following genetic or environmental perturbation. That is, individual cell-types appear to respond in different ways to a particular alteration, such as a mutation or an environmental stress. For example, the analysis of single-gene mutants shows distinct transcriptome responses at the cell-type level in roots (Brady et al., 2011; Bruex et al., 2012), cotton fibers (Wan et al., 2014), and trichomes (Jakoby et al., 2008). Further, distinct cell-type-specific alterations in transcriptional profiles are observed following a specific change in the environmental conditions, including iron deprivation, nitrogen availability, or salt stress, that likely reflect the physiological response appropriate for a particular cell-type (Dinnyen et al., 2008; Gifford et al., 2008; Geng et al., 2013). These results indicate that single cell-types of plants have distinct molecular programs that are not apparent from the analysis of whole organs or whole plants.

An important implication of these studies is that high resolution molecular analyses at the single cell level may be useful to improve forward genetic analyses. That is, although a given mutant individual may not exhibit an observable alteration at the level of the whole plant or from a bulk cell population, it may be possible to detect a molecular phenotype if single cell-type analyses were performed. Indeed, the results of several recent studies support this view. First, the molecular analysis of different wild-type *Arabidopsis* lines (ecotypes) has demonstrated substantial differences in their transcript, protein, and metabolite profiles (Keurentjes et al., 2006; Kliebenstein et al., 2006; Fu et al., 2009; Terpstra et al., 2010), indicating a surprising degree of underlying molecular variation that does not cause observable morphological differences, perhaps due to “phenotypic buffering.” Interestingly, these studies also enable the identification of new kinds of quantitative trait loci (QTLs) that likely mediate the responses of large numbers of genes for ecotype-specific traits. Further, the reverse genetic analysis of a collection of stele-enriched *Arabidopsis* transcription factor genes showed that, among the resulting single-gene mutants, 65% of them exhibited reproducible transcriptome changes whereas only 16% exhibited observable morphological changes in the root (Brady et al.,

2011). Together, these studies suggest that a substantial degree of molecular variation exists that does not impact the plant's phenotype.

Finally, direct demonstration of the value of single cell-type analyses for forward genetic analyses has recently been reported. Several studies have shown that a comparative transcriptome analysis of single cell-types from mutants versus wild-type provides enhanced resolution for transcript changes. Comparing cotton fiber transcriptomes from lines differing in fiber production, specific transcription factors and metabolic pathways were identified as fiber associated (Wan et al., 2014). Further, transcript profiles from trichomes of a wild-type and immature trichome mutant lines uncovered new genes required for normal trichome formation (Marks et al., 2009). In another study, several single-gene mutants associated with *Arabidopsis* root epidermis development, but lacking an observable morphological phenotype due to redundancy, were found to alter root epidermis transcript profiles in a manner that reflects the known biological role of the genes (Simon et al., 2013). In addition, the function of an uncharacterized gene in this root epidermal network (*TTG2*) was deduced by inspection of the cell-specific transcriptome of its corresponding mutant, which lacked a morphological abnormality (Simon et al., 2013). Interestingly, the transcript changes showed that *TTG2* normally promotes root hair cell differentiation, rather than non-hair cell differentiation as previously suspected. This shows that the high resolution analysis of single cell-types can reveal molecular differences associated with gene function.

Issues to Consider for Molecular Phenotyping of Single Cell-Types

The studies described above suggests it is possible to better understand gene function by conducting detailed molecular phenotyping of genetically altered lines (e.g., mutant lines). In particular, it is notable that even mutants lacking an outward (morphological) phenotype may exhibit a molecular phenotype at the cell-type level. This is exciting because the majority of single-gene knockouts in plants lack observable changes (Pickett and Meeks-Wagner, 1995; Bouche and Bouchez, 2001; Hanada et al., 2009; Perez-Perez et al., 2009; Lloyd and Meinke, 2012), and so this approach may enable their associated genes to be assigned a particular function.

For the successful application of this approach, there are several issues that should be considered. First, the biological material to be analyzed should be as specific as possible. Ideally, a single cell-type or cell state (specific stage of a given cell-type) should be analyzed. This is important as it is becoming apparent from many different studies that there is tremendous cell–cell variation within multicellular organisms. For example, the oligodendrocyte of the mammalian brain has historically been considered to be a single cell-type, but it is now known to be composed of six distinct subpopulations in the adult mouse brain (Zeisel et al., 2015). One of the limiting factors in isolating single cells is the accessibility of individual cell-types, with epidermal cells (e.g., root hairs, trichomes, cotton fibers) being among the first to be analyzed due to their extension from the plant surface

(Qiao and Libault, 2013; Becker et al., 2014). The difficulty in isolating other single cell-types is being mitigated by new advances in protoplasting/fluorescence activated cell sorting (FACS), laser capture microdissection (LCM), or nuclear tagging in specific cell-types (INTACT) that enable cells of internal tissues to be purified in plants (Kerk et al., 2003; Deal and Henikoff, 2010). Another limiting factor can be the small amount of biological material collected, which may prevent accurate assessment of molecule accumulation given current transcriptome, proteome, and metabolome methodologies. Indeed, this may prevent the detection of a “phenotype” in some mutants that elicit relatively minor effects. As the sensitivity of these large-scale methods increases, the ultimate goal is to conduct these molecular analyses on individual cells of a single cell-type. For example, it is conceivable that transcriptomes will soon be possible from single cells through the development of single cell RNA-seq methods (Saliba et al., 2014). It is likely that the combined technical advances in single cell isolation and molecular analysis methods will be required for future success.

Second, the likely success of using this approach is improved if the investigator has some knowledge of the likely defect in the mutant line. This knowledge enables the investigator to focus their molecular analysis on one (or a limited number of) cell-type(s) in the mutant line. In the absence of some knowledge or interest in a particular cell-type(s), it would be difficult to employ this approach to deducing the role of a gene, due to the time and

resources required to survey a large number of cell/tissue types and developmental stages of the mutant plant. Having said this, at some point in the future, it may be feasible to consider the ultimate goal of defining the molecular impact of every gene knockout on every cell-type in the plant. It is likely that this information would profoundly change our view of gene activity and function in plants.

Third, this approach is most likely to be successful if there is a known molecular pathway or molecular markers for the cell-type or trait of interest already available to aid in the assignment of gene function. For example, if some of the genes have been identified in a transcriptional regulatory network for the cell-type of interest, then it is a relatively straightforward exercise to determine whether/how the mutant lines alter specific genes or subdomains of the gene network using transcriptomic approaches. This feature has been exploited in the successful applications of this approach to date, using the detailed knowledge available in the trichome and root hair systems (Marks et al., 2009; Simon et al., 2013).

In summary, recent advances in the isolation and characterization of plant single cell-types shows great promise in enhancing the effectiveness of mutant phenotyping. By focusing on single cell-types, rather than entire organs or entire plants, it is possible to detect specific molecular changes in mutant lines. This represents a valuable tool for the future detailed dissection of gene function in plants.

References

- Adrian, J., Chang, J., Ballenger, C. E., Bargmann, B. O., Alassimone, J., Davies, K. A., et al. (2015). Transcriptome dynamics of the stomatal lineage: birth, amplification, and termination of a self-renewing population. *Dev. Cell* 33, 107–118. doi: 10.1016/j.devcel.2015.01.025
- Becker, J. D., Takeda, S., Borges, F., Dolan, L., and Feijo, J. A. (2014). Transcriptional profiling of *Arabidopsis* root hairs and pollen defines an apical cell growth signature. *BMC Plant Biol.* 14:197. doi: 10.1186/s12870-014-0197-3
- Birnbaum, K., Shasha, D. E., Wang, J. Y., Jung, J. W., Lambert, G. M., Galbraith, D. W., et al. (2003). A gene expression map of the *Arabidopsis* root. *Science* 302, 1956–1960. doi: 10.1126/science.1090022
- Bouche, N., and Bouchez, D. (2001). *Arabidopsis* gene knockout: phenotypes wanted. *Curr. Opin. Plant Biol.* 4, 111–117. doi: 10.1016/S1369-5266(00)00145-X
- Brady, S. M., Orlando, D. A., Lee, J. Y., Wang, J. Y., Koch, J., Dinneny, J. R., et al. (2007). A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science* 318, 801–806. doi: 10.1126/science.1146265
- Brady, S. M., Zhang, L., Megraw, M., Martinez, N. J., Jiang, E., Yi, C. S., et al. (2011). A stele-enriched gene regulatory network in the *Arabidopsis* root. *Mol. Syst. Biol.* 7, 459. doi: 10.1038/msb.2010.114
- Brandt, S. P. (2005). Microgenomics: gene expression analysis at the tissue-specific and single-cell levels. *J. Exp. Bot.* 56, 495–505. doi: 10.1093/jxb/eri066
- Bruex, A., Kainkaryam, R. M., Wiecek, Y., Kang, Y. H., Bernhardt, C., Xia, Y., et al. (2012). A gene regulatory network for root epidermis cell differentiation in *Arabidopsis*. *PLoS Genet.* 8:e1002446. doi: 10.1371/journal.pgen.1002446
- Deal, R. B., and Henikoff, S. (2010). A simple method for gene expression and chromatin profiling of individual cell types within a tissue. *Dev. Cell* 18, 1030–1040. doi: 10.1016/j.devcel.2010.05.013
- Dinneny, J. R., Long, T. A., Wang, J. Y., Jung, J. W., Mace, D., Pointer, S., et al. (2008). Cell identity mediates the response of *Arabidopsis* roots to abiotic stress. *Science* 320, 942–945. doi: 10.1126/science.1153795
- Fu, J., Keurentjes, J. J., Bouwmeester, H., America, T., Verstappen, F. W., Ward, J. L., et al. (2009). System-wide molecular evidence for phenotypic buffering in *Arabidopsis*. *Nat. Genet.* 41, 166–167. doi: 10.1038/ng.308
- Geng, Y., Wu, R., Wee, C. W., Xie, F., Wei, X., Chan, P. M., et al. (2013). A spatiotemporal understanding of growth regulation during the salt stress response in *Arabidopsis*. *Plant Cell* 25, 2132–2154. doi: 10.1105/tpc.113.112896
- Gifford, M. L., Dean, A., Gutierrez, R. A., Coruzzi, G. M., and Birnbaum, K. D. (2008). Cell-specific nitrogen responses mediate developmental plasticity. *Proc. Natl. Acad. Sci. U.S.A.* 105, 803–808. doi: 10.1073/pnas.0709559105
- Hanada, K., Kuromori, T., Myouga, F., Toyoda, T., Li, W. H., and Shinozaki, K. (2009). Evolutionary persistence of functional compensation by duplicate genes in *Arabidopsis*. *Genome Biol. Evol.* 1, 409–414. doi: 10.1093/gbe/evp043
- Jakoby, M. J., Falkenhan, D., Mader, M. T., Brininstool, G., Wischnitzki, E., Platz, N., et al. (2008). Transcriptional profiling of mature *Arabidopsis* trichomes reveals that NOECK encodes the MIXTA-like transcriptional regulator MYB106. *Plant Physiol.* 148, 1583–1602. doi: 10.1104/pp.108.126979
- Jiao, Y., Tausta, S. L., Gandotra, N., Sun, N., Liu, T., Clay, N. K., et al. (2009). A transcriptome atlas of rice cell types uncovers cellular, functional and developmental hierarchies. *Nat. Genet.* 41, 258–263. doi: 10.1038/ng.282
- Kerk, N. M., Ceserani, T., Tausta, S. L., Sussex, I. M., and Nelson, T. M. (2003). Laser capture microdissection of cells from plant tissues. *Plant Physiol.* 132, 27–35. doi: 10.1104/pp.102.018127
- Keurentjes, J. J., Fu, J., de Vos, C. H., Lommen, A., Hall, R. D., Bino, R. J., et al. (2006). The genetics of plant metabolism. *Nat. Genet.* 38, 842–849. doi: 10.1038/ng1815
- Kliebenstein, D. J., West, M. A., van Leeuwen, H., Loudet, O., Doerge, R. W., and St Clair, D. A. (2006). Identification of QTLs controlling gene expression networks defined *a priori*. *BMC Bioinformatics* 7:308. doi: 10.1186/1471-2105-7-308
- Lieckfeldt, E., Simon-Rosin, U., Kose, F., Zoeller, D., Schliep, M., and Fisahn, J. (2008). Gene expression profiling of single epidermal, basal and trichome cells of *Arabidopsis thaliana*. *J. Plant Physiol.* 165, 1530–1544. doi: 10.1016/j.jplph.2007.06.017
- Lloyd, J., and Meinke, D. (2012). A comprehensive dataset of genes with a loss-of-function mutant phenotype in *Arabidopsis*. *Plant Physiol.* 158, 1115–1129. doi: 10.1104/pp.111.192393
- Marks, M. D., Wenger, J. P., Gilding, E., Jilk, R., and Dixon, R. A. (2009). Transcriptome analysis of *Arabidopsis* wild-type and gl3-sst sim trichomes identifies four additional genes required for trichome development. *Mol. Plant* 2, 803–822. doi: 10.1093/mp/ssp037

- Moussaieff, A., Rogachev, I., Brodsky, L., Malitsky, S., Toal, T. W., Belcher, H., et al. (2013). High-resolution metabolic mapping of cell types in plant roots. *Proc. Natl. Acad. Sci. U.S.A.* 110, E1232–E1241. doi: 10.1073/pnas.1302019110
- Perez-Perez, J. M., Candela, H., and Micol, J. L. (2009). Understanding synergy in genetic interactions. *Trends Genet.* 25, 368–376. doi: 10.1016/j.tig.2009.06.004
- Petricka, J. J., Schauer, M. A., Megraw, M., Breakfield, N. W., Thompson, J. W., Georgiev, S., et al. (2012). The protein expression landscape of the *Arabidopsis* root. *Proc. Natl. Acad. Sci. U.S.A.* 109, 6811–6818. doi: 10.1073/pnas.1202546109
- Pickett, F. B., and Meeks-Wagner, D. R. (1995). Seeing double: appreciating genetic redundancy. *Plant Cell* 7, 1347–1356. doi: 10.1105/tpc.7.9.1347
- Qiao, Z., and Libault, M. (2013). Unleashing the potential of the root hair cell as a single plant cell type model in root systems biology. *Front. Plant Sci.* 4:484. doi: 10.3389/fpls.2013.00484
- Saliba, A. E., Westermann, A. J., Gorski, S. A., and Vogel, J. (2014). Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* 42, 8845–8860. doi: 10.1093/nar/gku555
- Simon, M., Bruex, A., Kainkaryam, R. M., Zheng, X., Huang, L., Woolf, P., et al. (2013). Tissue-specific profiling reveals transcriptome alterations in *Arabidopsis* mutants lacking morphological phenotypes. *Plant Cell* 25, 3175–3185. doi: 10.1105/tpc.113.115121
- Terpstra, I. R., Snoek, L. B., Keurentjes, J. J., Peeters, A. J., and van den Ackerveken, G. (2010). Regulatory network identification by genetical genomics: signaling downstream of the *Arabidopsis* receptor-like kinase ERECTA. *Plant Physiol.* 154, 1067–1078. doi: 10.1104/pp.110.159996
- Wan, Q., Zhang, H., Ye, W., Wu, H., and Zhang, T. (2014). Genome-wide transcriptome profiling revealed cotton fuzz fiber development having a similar molecular model as *Arabidopsis* trichome. *PLoS ONE* 9:e97313. doi: 10.1371/journal.pone.0097313
- Zeisel, A., Munoz-Manchado, A. B., Codeluppi, S., Lonnerberg, P., La Manno, G., Jureus, G., et al. (2015). Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* 347, 1138–1142. doi: 10.1126/science.aaa1934

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Schiefelbein. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

System approaches to study root hairs as a single cell plant model: current status and future perspectives

Md Shakhawat Hossain¹, Trupti Joshi² and Gary Stacey^{1*}

¹ Division of Plant Sciences and Biochemistry, National Center for Soybean Biotechnology, Christopher S. Bond Life Sciences Center, University of Missouri, Columbia, MO, USA, ² Department of Computer Science, Informatics Institute and Christopher S. Bond Life Sciences Center, University of Missouri, Columbia, MO, USA

OPEN ACCESS

Edited by:

Marc Libault,
University of Oklahoma, USA

Reviewed by:

Chuang Ma,
Northwest Agriculture and Forest
University, China
María E. Zanetti,
CONICET and Universidad Nacional
de LA Plata, Argentina

*Correspondence:

Gary Stacey,
Division of Plant Sciences
and Biochemistry, National Center
for Soybean Biotechnology,
Christopher S. Bond Life Sciences
Center, University of Missouri,
1201 Rollins Street, Columbia, MO
65211, USA
staceyg@missouri.edu

Specialty section:

This article was submitted to
Plant Systems and Synthetic Biology,
a section of the journal
Frontiers in Plant Science

Received: 17 March 2015

Accepted: 06 May 2015

Published: 19 May 2015

Citation:

Hossain MS, Joshi T and Stacey G
(2015) System approaches to study
root hairs as a single cell plant model:
current status and future
perspectives.
Front. Plant Sci. 6:363.
doi: 10.3389/fpls.2015.00363

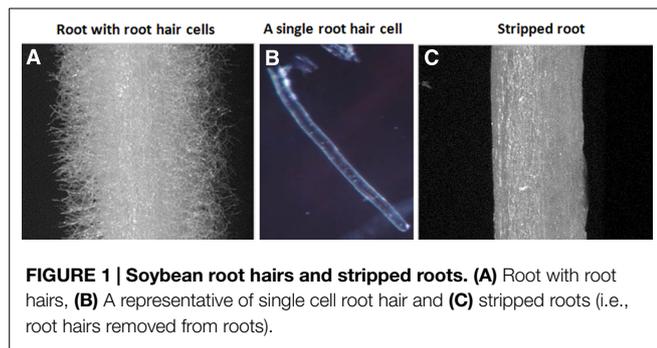
Our current understanding of plant functional genomics derives primarily from measurements of gene, protein and/or metabolite levels averaged over the whole plant or multicellular tissues. These approaches risk diluting the response of specific cells that might respond strongly to the treatment but whose signal is diluted by the larger proportion of non-responding cells. For example, if a gene is expressed at a low level, does this mean that it is indeed lowly expressed or is it highly expressed, but only in a few cells? In order to avoid these issues, we adopted the soybean root hair cell, derived from a single, differentiated root epidermal cell, as a single-cell model for functional genomics. Root hair cells are intrinsically interesting since they are major conduits for root water and nutrient uptake and are also the preferred site of infection by nitrogen-fixing rhizobium bacteria. Although a variety of other approaches have been used to study single plant cells or single cell types, the root hair system is perhaps unique in allowing application of the full repertoire of functional genomic and biochemical approaches. In this mini review, we summarize our published work and place this within the broader context of root biology, with a significant focus on understanding the initial events in the soybean-rhizobium interaction.

Keywords: root hair, single cell, rhizobium, soybean, systems biology

Why Root Hairs are an Excellent, Single-cell, Plant Model for Systems Biology?

A root hair is a single cell (Wan et al., 2005; Brechenmacher et al., 2009, 2012; Libault et al., 2010a; Qiao and Libault, 2013), structurally simple and tubular outgrowth of root epidermal cells (Grierson et al., 2014). Root hairs have a huge absorptive surface area (Hofer, 1991), evolved in order to allow the plant to take up water, nutrients and minerals (Minorsky, 2002). They are also a major route for plant-microbe interactions (Oldroyd, 2001). For example, legume-rhizobium interactions that lead to the formation of a new organ, the nodule, where biological N₂-fixation takes place (Oldroyd and Dixon, 2014).

In order to adapt to environmental changes and respond to morphological or developmental stimuli, plant cells have evolved complex regulatory networks integrating the response at a variety of levels; such as DNA, RNA, proteins, metabolites and small molecules (Pu and Brady, 2010). While some plant genomes can encode over 50,000 genes, it is clear that many genes are specifically



expressed in only a few organs, tissues or cell types. However, it can be technically very challenging to measure the levels of genes, proteins or metabolites in a specific cell type (Rogers et al., 2012). Hence, most studies measure these components in whole plants or tissues, resulting in an averaging of the responses occurring within all cells. In order to overcome these sampling issues, the soybean root hair was proposed as a single cell plant model (Libault et al., 2010a). The size and thickness of the soybean root allows root hair cells to be isolated easily after freezing in liquid nitrogen. The result of this procedure is pure root hair preparations in gram quantities allowing the full repertoire of functional genomic methods to be applied (Wan et al., 2005; Brechenmacher et al., 2009, 2012; Libault et al., 2010a; Nguyen et al., 2012; Qiao and Libault, 2013).

Studies of plant root hairs are not new. A literature search identified more than 1,300 articles covering various aspects of root hair biology, including studies of root hair elongation, tip growth, polarized cell expansion, endomembrane trafficking, cytoskeletal organization and cell wall modifications in model and crop plants; such as, *Arabidopsis*, *Lotus japonicus*, rice, corn, barley, and tomato (<http://www.iroothair.org/>; Foreman and Dolan, 2001; Karas et al., 2005; Yokota et al., 2009; Hossain et al., 2012; Peña et al., 2012; Kwasniewski et al., 2013; Grierson et al., 2014). However, recent advances with our methods of single cell root hair isolation in soybean (Figure 1; Libault et al., 2010a; Qiao and Libault, 2013), along with the availability of the soybean genome sequence (Schmutz et al., 2010) and high-throughput sequencing, proteomic, metabolomic and epigenetic technologies make the soybean root hair system particularly attractive for detailed, systems-level studies. The ultimate goal is to use this information for computational prediction and integration of big data sets for network analysis of plant cell function (Figure 2; Foreman and Dolan, 2001; Wan et al., 2005; Brechenmacher et al., 2009, 2012; Kwasniewski et al., 2010; Libault et al., 2010a; Peña et al., 2012; Grierson et al., 2014).

What have We Learned about Root Hairs Using System Approaches?

The study of a single cell system provides significantly higher resolution and sensitivity when various functional genomic methods are applied. This has been amply demonstrated by studies in *Arabidopsis* that characterized the transcriptome, proteome and metabolome of various root cell types (Pu and Brady, 2010; Rogers

et al., 2012; Moussaieff et al., 2013; Misra et al., 2014). In our laboratory, over the past few years, the full repertoire of functional genomic methods has been applied to studies of soybean root hairs as a single cell plant model (Wan et al., 2005; Brechenmacher et al., 2009; Libault et al., 2010b; Qiao and Libault, 2013). From these efforts, one could argue that, the soybean root hair cell is one of the best characterized cell types in plant biology.

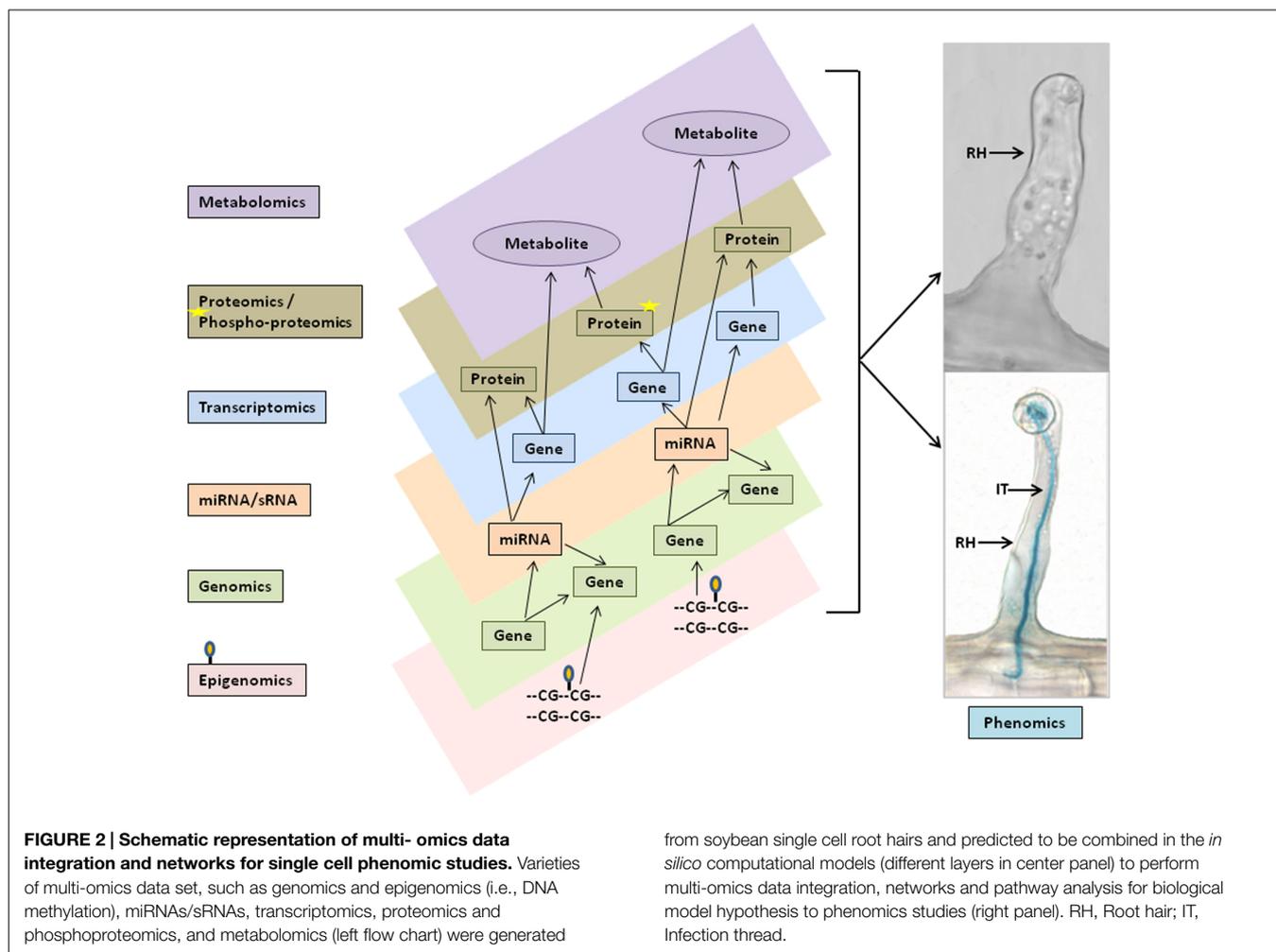
Root Hair Transcriptome

A number of genome-wide transcriptome profiling studies have been published using root hairs from several model and crop plants (Jones et al., 2006; Kwasniewski et al., 2010; Libault et al., 2010b; Lan et al., 2013; Libault, 2013; Breakspear et al., 2014). For example, in *Arabidopsis* and barley, researchers used this cell type to study transcriptional regulation mostly focused on root hair morphogenesis, cell fate, cellular growth and differentiation. Since root hairs expand by polar growth, along with pollen tubes, they serve as a model to study this distinctive growth process (Campanoni and Blatt, 2007).

However, in legumes, root hairs are also the primary site for rhizobial infection and, therefore, several studies have sought to define the early events in this infection process. Libault et al. (2010b) studied soybean root hair infection by the symbiotic bacterium *Bradyrhizobium japonicum*. This study used two combined platforms (Affymetrix and Illumina sequencing) and identified 1,973 genes that were differentially expressed in response to bacterial infection, including those involved in the initial rhizobial symbiotic signal, lipo-chitooligosaccharide (Nod) factor perception, plant defense response, modification of cell wall composition, signal transduction, basic metabolic processes, and hormonal regulation (Libault et al., 2010b). Very similar findings came from a more recent investigation in the model legume, *Medicago truncatula* (Breakspear et al., 2014). A microarray based transcriptomics approach identified hundreds of genes regulated in root hair cells in response to *Sinorhizobium meliloti* and bacterial Nod factor application. A comparison of these two studies revealed ~370 genes differentially regulated by rhizobial inoculation in the two legume species. Among genes responding in both species were those shown previously to be critical for the legume-rhizobial interaction; including, *NIN* (Nodule Inception), *PUB1* (Plant U-box Protein 1), *VPY* (Vapyrin), *RPG* (Rhizobium-directed Polar Growth), *NSP1* (Nodulation Signaling Pathway 1), *NSP2* (Nodulation signaling Pathway 2), *NPL1* (Nodulation Pectate Lyase 1), *FLOT4* (Flotillin-like protein 4), *ERN1* (Ethylene Response Factor Required for Nodulation1), *ERN2* (Ethylene Response Factor Required for Nodulation2), *NFYA1* (Nuclear transcription factor Y subunit A-1), and *NMNI* (Nucleolar/Mitochondrial protein involved in Nodulation) (Breakspear et al., 2014). These findings emphasize the utility of root hair studies to identify key genes involved in the rhizobial symbiosis.

Root Hair Proteome

While transcriptome studies are rather common, fewer studies have focused on the root hair proteome (Wan et al., 2005; Brechenmacher et al., 2009, 2012; Pang et al., 2010; Nestler et al., 2011;



Subramanian and Smith, 2013). One can argue that, while mRNA profiling provides a picture of the potential functions in the cell, only the proteome can give you a true picture of which of these functions are likely occurring. The first report on the soybean root hair proteome focused on providing a protein reference map of this single cell type (Brechenmacher et al., 2009) using two-dimensional-polyacrylamide gel electrophoresis (2D-PAGE), augmented by multidimensional protein identification technology (MudPIT). This study identified 1,492 proteins involved in basic cell metabolism, water and nutrient uptake, vesicle trafficking, and hormone and secondary metabolism. A later study, using the Accurate Mass and Time (AMT) tag approach combined with liquid chromatography-tandem mass spectrometry (LC-MS/MS), identified a total of 5,702 proteins from soybean root hair cell preparations (Brechenmacher et al., 2012). Both studies reported similar functional categories of proteins. A recent study focused on root hairs isolated from the monocot maize with proteins separated by 1-dimensional PAGE and then subjected to nano LC-MS/MS (Nestler et al., 2011). This study identified 2,573 abundant proteins in maize root hair cells. Interestingly, a comparison of the soybean (dicot) and maize (monocot) datasets identified 252 conserved proteins pointing to functionally conserved, root hair functions in these disparate species.

In order to specifically address protein changes occurring in root hairs upon rhizobial inoculation, Wan et al. (2005) utilized 2-D PAGE to separate proteins, which were then identified by matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) MS. This study identified 37 proteins including enzymes, such as chitinase and phosphoenolpyruvate carboxylase, that appeared to be specific to root hairs (relative to roots stripped of the root hairs), as well as peroxidase, phenylalanine-ammonia lyase, lectin, phospholipase D and phosphoglucomutase, whose expression changed significantly upon rhizobial inoculation. It was previously shown that these proteins played significant roles in root hair deformation, infection and legume nodulation in response to bacterial infection (Halverson and Stacey, 1985; Díaz et al., 1989; Estabrook and Sengupta-Gopalan, 1991; Cook et al., 1995; van Rhijn et al., 1998; den Hartog et al., 2001; Lepek et al., 2002; Mitra and Long, 2004; Yan et al., 2015b).

Root Hair Phosphoproteome

It is now clear that many of the initial events in Nod factor perception and the plant response to rhizobial infection involves activation of a variety of protein kinases (Antolín-Llovera et al.,

2014; Liang et al., 2014). Therefore, it is of interest to identify those proteins that are rapidly phosphorylated after rhizobial inoculation. Modern MS-based methods for phosphoproteomic analysis allow for such global analyses. For example, Rose et al. (2012) analyzed the phosphoproteome of *M. truncatula* roots after inoculation with *S. meliloti*. However, again, although this report identified a variety of interesting proteins, the results represent an average of the whole root response, not that of specific cells. Hence, our laboratory undertook a similar study to specifically examine the phosphoproteome of soybean root hair cells subsequent to rhizobial inoculation (Nguyen et al., 2012). Again, this is an ideal cell type since it is the site of the initial interaction and penetration of the plant root by the rhizobial symbiont. The root hair phosphoproteome was compared to that of roots stripped of their root hairs. In order to provide accurate quantification of peptide levels, each was labeled with an isobaric tag (eight plex) using the iTRAQ (isobaric tags for relative and absolute quantification) method, followed by phosphopeptide enrichment and LC-MS/MS analysis. A total of 1,625 unique phosphopeptides, spanning 1,659 non-redundant phosphorylation sites, were detected from 1,126 soybean phosphoproteins from both root hairs and stripped roots. Among the identified phosphopeptides, the levels of 273, belonging to 240 phosphoproteins, were found to be significantly regulated upon *B. japonicum* infection suggesting a complex network of kinase substrate and phosphatase-substrate interactions in response to rhizobial inoculation (Nguyen et al., 2012). Proteins predicted to play a role in signal transduction (e.g., protein kinases, protein phosphatases, protein phosphatase inhibitors, and G protein-related proteins) and those involved in hormone signaling were among those whose phosphorylation was specifically affected by *B. japonicum* inoculation. The identified phosphoproteins and phosphorylation site data were deposited and are available at the Plant Protein Phosphorylation Database (P³DB; <http://digbio.missouri.edu/p3db/>).

Root Hair Metabolome

Similar to transcriptomics and proteomics, responses and measurement of metabolites from multi-cellular tissues, organs or the whole plant could give misleading information if specific metabolism is confined to only a few cells or a single cell type. Until recently, a single analytical platform was not available that could measure a significant number of metabolites from a single cell (Fukushima et al., 2009). However, extant methods could be applied to soybean root hairs given our ability to isolate this cell type in a pure form and in quantity (Libault et al., 2010a). Again, the legume-rhizobium interaction was an obvious target for such analyses given the relevance of the root hair system and a variety of publications implicating specific metabolites as playing important roles. A variety of secondary products (e.g., flavonoids), as well as various hormones, were shown to be important regulators of the nodulation process (Matamoros et al., 2006; Gibson et al., 2008; Ding and Oldroyd, 2009). For example, formation of the infection thread by which rhizobia gain entry into the plant root hair cell might require a significant change in the metabolism of both the symbiont and host.

Brechenmacher et al. (2010) investigated the root hair metabolome in an effort to identify metabolites whose levels changed significantly during the first 48 h after rhizobial inoculation. Metabolites were analyzed using both gas chromatography-mass spectrometry (GC-MS) and ultra-performance liquid chromatography-quadrupole time of flight-mass spectrometry (UPLC-MS). Using these combined approaches, a total of 2,610 metabolites were identified in soybean root hair cells. Among these metabolites, 166 were found to be significantly regulated in response to rhizobial infection, including various (iso)flavonoids, amino acids, fatty acids, carboxylic acids, and various carbohydrates, indicating major metabolic changes occurring during *B. japonicum*-root hair interactions. Among these metabolites was trehalose, an α -linked disaccharide of glucose, which has been implicated in a variety of plant processes, including resistance to osmotic stress (Iturriaga et al., 2009). Trehalose levels increased significantly in root hairs upon inoculation by *B. japonicum*. However, through the use of various mutants of *B. japonicum* blocked in the synthesis of trehalose, the majority of this disaccharide could be attributed to the bacteria and did not appear to be predominantly derived from the plant host. The authors postulated that trehalose synthesis by *B. japonicum* may be important to allow the bacteria to survive what may be a stressful osmotic environment within the plant.

Root Hair Small RNAome

MicroRNAs (miRNAs) are small non-protein coding endogenous RNAs, typically 21-24 nucleotides in length. In the past decade, miRNAs have been shown to be key players controlling gene expression by transcript cleavage or translational inhibition in a wide variety of plant biological processes, including growth and development, disease, stress and plant microbe interactions (Kidner and Martienssen, 2005; Xiao and Rajewsky, 2009; Lauressergues et al., 2012; Sunkar et al., 2012; Weiberg et al., 2013). Indeed, several studies documented an important role for miRNAs in regulated gene expression during the legume nodulation process (Combiere et al., 2006; Subramanian et al., 2008; Lelandais-Brière et al., 2009; Wang et al., 2009; Joshi et al., 2010; Li et al., 2010; De Luis et al., 2012; Turner et al., 2013; Yan et al., 2013, 2015a). However, all of these reports came from small RNA profiling of roots and/or nodules and not from more specific studies of the initial infection process within root hairs. In order to target these very early stages of rhizobial-host interaction, Yan et al. (2015b) utilized *B. japonicum* infected soybean root hair single cells and roots stripped of their root hairs to generate both small RNA and degradome libraries. Sequencing of three small RNA libraries from inoculated root hairs, stripped roots and mock inoculated control samples identified a total of 114 miRNAs. Among these, 22 were found to be novel miRNAs. Comparative analysis of miRNA abundance identified 66 miRNAs that were differentially expressed between root hair and stripped roots. A total of 48 miRNAs were differentially regulated in root hairs in response to bacterial infection when compared to the un-infected control root hairs. Sequencing of a Parallel Analysis of RNA Ends (PARE; i.e., degradome) library from similar tissues revealed a total of 405 soybean mRNA targets. This method identifies

new 5'-mRNA ends presumably arising from miRNA-mediated cleavage (German et al., 2009; Zhai et al., 2014). The mRNA targets identified were predicted to encode transcription factors or proteins involved in protein modification, protein degradation and enzymes in various hormonal pathways. The root hair data set represents an important starting point for in depth analysis of the role that specific miRNAs may be playing in the legume-rhizobial symbiosis.

Bioinformatics and Data Integration for Network Analysis and Modeling

Next generation sequencing technologies have empowered researchers to conduct experiments on a whole genome scale and have the potential to completely revolutionize biological research. Big Data have been generated in all domains including transcriptomics (RNA-seq), proteomics, metabolomics, epigenomics, microRNA/smallRNA, and genomic variations, including single nucleotide polymorphisms (SNP) and insertion/deletions (InDels; Mochida and Shinozaki, 2010; Brauer et al., 2014). Most lab experiments now generate anywhere from a few GB to several TB of raw and analyzed data, and have a need to overlay different types of data and information to get a comprehensive understanding. All of these data provide valuable insights into a systems-level understanding of the biology of an organism and need to be mined in an innovative and integrative manner. This is posing a new challenge for researchers and mandating the development of comprehensive, efficient informatic platforms and web resources to facilitate data sharing and collaboration among the research community, while providing advanced techniques for multi-omics integration, computational analysis, and hypothesis generation.

Most databases and available tools can handle analysis of a single -omics data-type, but analysis of multiple omics data-types presents a major challenge. No tools are currently available that provide a systematic solution to this problem. In most cases, only by integrated analysis of the expression of mRNA, proteins, metabolites along with miRNA/sRNA, etc (Figure 2) can we draw conclusions about all the involved regulatory mechanisms. With multi-omics data integration techniques, we can gain better insight into the network modules derived and, supported by experimental evidence, utilize this information to build *in silico* computational models for understanding the underlying mechanisms and to generate new hypotheses. The *in silico* models provide templates that automate the identification and generation of network modules by incorporating differentially expressed genes, proteins, and metabolites into the function and pathway enrichment analysis. The models can be visualized using multiple layers of platforms such that data from transcriptomics, proteomics, and

metabolomics can each be visualized simultaneously on separate platform layers as shown in Figure 2. This presentation enables rapid comparisons that can be further enriched by incorporating available experimental data, literature references, and user inputs to cross-validate results. Users can start with their own hypothesis and add the multi-omics evidence stored in the associated database to continue building the modules. Such *in silico* modeling systems can provide researchers with additional avenues for generating and testing a specific hypothesis by incorporating diverse varieties of data. Bioinformatics plays a key and essential role in bringing these pieces together in the form of one-stop-shop web resources, such as the soybean knowledge base (SoyKB; Joshi et al., 2014, 2012), and facilitating the ability of researchers to get a more complete understanding of the underlying molecular mechanisms.

Conclusion

Our planet is facing increasing challenges related to environment, food security, non-renewable energy sources, water availability and overall sustainability. Plant biological research and its application in agriculture have the potential to mitigate many of these challenges. Thus, we need to have a better understanding of plant biology at a systems level, ultimately reaching the point where computational models can predict biological outcomes. Irrespective of the popularity of system biology, we are still a long way from achieving this level of understanding. Among the obstacles is the functional complexity of cellular/organismal regulatory networks, as well as our ability to deal with Big Data, especially integration of disparate data types. There is a need to simplify our systems to allow for clearer conclusions, while increasing the resolution and sensitivity of our analysis, without losing contact with real world relevance. The root hair single cell system provides one such route for the study of plant processes. The use of the diverse “omics” data sets developed from the soybean root hair as a single cell type provides an opportunity for comprehensive network analysis and ultimately will help to build an integrated predictive model. Besides the inherent interest of the root hair cell, the hope is that as an illustrative model system, the soybean root hair cell system can contribute to an overall greater understanding of plant biology leading ultimately to improvements in agriculture.

Acknowledgments

Research was funded by a grant from the USA National Science Foundation Plant Genome program (grant no. DBI-0421620 to GS). We are thankful to Joseph M. Batek, University of Missouri for aiding in the literature review. We also thank Dr. Katalin Toth for reading and revising the manuscript.

References

- Antolín-Llovera, M., Petutsching, E. K., Ried, M. K., Lipka, V., Nürnberger, T., Robatzek, S., et al. (2014). Knowing your friends and foes—plant receptor-like kinases as initiators of symbiosis or defence. *New Phytol.* 204, 791–802. doi: 10.1111/nph.13117
- Brauer, E. K., Singh, D. K., and Popescu, S. C. (2014). Next-generation plant science: putting big data to work. *Genome Biol.* 15, 301. doi: 10.1186/gb4149

- Brechenmacher, L., Lei, Z., Libault, M., Findley, S., Sugawara, M., Sadowsky, M. J., et al. (2010). Soybean metabolites regulated in root hairs in response to the symbiotic bacterium *Bradyrhizobium japonicum*. *Plant Physiol.* 153, 1808–1822. doi: 10.1104/pp.110.157800
- Breakspear, A., Liu, C., Roy, S., Stacey, N., Rogers, C., Trick, M., et al. (2014). The root hair “Infectome” of *Medicago truncatula* uncovers changes in cell cycle genes and reveals a requirement for auxin signaling in rhizobial infection. *Plant Cell* 26, 4680–4701. doi: 10.1105/tpc.114.133496

- Brechenmacher, L., Lee, J., Sachdev, S., Song, Z., Nguyen, T. H. N., Joshi, T., et al. (2009). Establishment of a protein reference map for soybean root hair cells. *Plant Physiol.* 149, 670–682. doi: 10.1104/pp.108.131649
- Brechenmacher, L., Nguyen, T. H. N., Hixson, K., Libault, M., Aldrich, J., Pasa-Tolic, L., et al. (2012). Identification of soybean proteins from a single cell type: the root hair. *Proteomics* 12, 3365–3373. doi: 10.1002/pmic.201200160
- Campanoni, P., and Blatt, M. R. (2007). Membrane trafficking and polar growth in root hairs and pollen tubes. *J. Exp. Bot.* 58, 65–74. doi: 10.1093/jxb/erl059
- Combier, J.-P., Frugier, F., de Billy, F., Boualem, A., El-Yahyaoui, F., Moreau, S., et al. (2006). MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in *Medicago truncatula*. *Genes Dev.* 20, 3084–3088. doi: 10.1101/gad.402806
- Cook, D., Dreyer, D., Bonnet, D., Howell, M., Nony, E., and VandenBosch, K. (1995). Transient induction of a peroxidase gene in *Medicago truncatula* precedes infection by *Rhizobium meliloti*. *Plant Cell* 7, 43–55. doi: 10.1105/tpc.7.1.43
- De Luis, A., Markmann, K., Cognat, V., Holt, D. B., Charpentier, M., Parniske, M., et al. (2012). Two microRNAs linked to nodule infection and nitrogen-fixing ability in the legume *Lotus japonicus*. *Plant Physiol.* 160, 2137–2154. doi: 10.1104/pp.112.204883
- den Hartog, M., Musgrave, A., and Munnik, T. (2001). Nod factor-induced phosphatidic acid and diacylglycerol pyrophosphate formation: a role for phospholipase C and D in root hair deformation. *Plant J. Cell Mol. Biol.* 25, 55–65. doi: 10.1046/j.1365-313x.2001.00931.x
- Díaz, C. L., Melchers, L. S., Hooykaas, P. J. J., Lugtenberg, B. J. J., and Kijne, J. W. (1989). Root lectin as a determinant of host-plant specificity in the *Rhizobium-legume* symbiosis. *Nature* 338, 579–581. doi: 10.1038/338579a0
- Ding, Y., and Oldroyd, G. E. D. (2009). Positioning the nodule, the hormone dictum. *Plant Signal. Behav.* 4, 89–93. doi: 10.4161/psb.4.2.7693
- Estabrook, E. M., and Sengupta-Gopalan, C. (1991). Differential expression of phenylalanine ammonia-lyase and chalcone synthase during soybean nodule development. *Plant Cell* 3, 299–308. doi: 10.1105/tpc.3.3.299
- Foreman, J., and Dolan, L. (2001). Root Hairs as a model system for studying plant cell growth. *Ann. Bot.* 88, 1–7. doi: 10.1006/anbo.2001.1430
- Fukushima, A., Kusano, M., Redestig, H., Arita, M., and Saito, K. (2009). Integrated omics approaches in plant systems biology. *Curr. Opin. Chem. Biol.* 13, 532–538. doi: 10.1016/j.cbpa.2009.09.022
- German, M. A., Luo, S., Schroth, G., Meyers, B. C., and Green, P. J. (2009). Construction of parallel analysis of RNA Ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. *Nat. Protoc.* 4, 356–362. doi: 10.1038/nprot.2009.8
- Gibson, K. E., Kobayashi, H., and Walker, G. C. (2008). Molecular determinants of a symbiotic chronic infection. *Annu. Rev. Genet.* 42, 413–441. doi: 10.1146/annurev.genet.42.110807.091427
- Grierson, C., Nielsen, E., Ketelaarc, T., and Schiefelbein, J. (2014). Root Hairs. *Arab. Book* 12:e0172. doi: 10.1199/tab.0172
- Halverson, L. J., and Stacey, G. (1985). Host Recognition in the *Rhizobium-Soybean* Symbiosis. *Plant Physiol.* 77, 621–625. doi: 10.1104/pp.77.3.621
- Hofer, R.-M. (1991). “Root hairs,” in *Plant Roots The Hidden Half*, eds Y. Waisel, A. Eshel, and U. Kafkafi (New York: Marcel Dekker), 129–148.
- Hossain, M. S., Liao, J., James, E. K., Sato, S., Tabata, S., Jurkiewicz, A., et al. (2012). *Lotus japonicus* ARPC1 is required for rhizobial infection. *Plant Physiol.* 112, 202572. doi: 10.1104/pp.112.202572
- Iturriaga, G., Suárez, R., and Nova-Franco, B. (2009). Trehalose metabolism: from osmoprotection to signaling. *Int. J. Mol. Sci.* 10, 3793–3810. doi: 10.3390/ijms10093793
- Jones, M. A., Raymond, M. J., and Smirnov, N. (2006). Analysis of the root-hair morphogenesis transcriptome reveals the molecular identity of six genes with roles in root-hair development in *Arabidopsis*. *Plant J. Cell Mol. Biol.* 45, 83–100. doi: 10.1111/j.1365-313X.2005.02609.x
- Joshi, T., Fitzpatrick, M. R., Chen, S., Liu, Y., Zhang, H., Endacott, R. Z., et al. (2014). Soybean knowledge base (SoyKB): a web resource for integration of soybean translational genomics and molecular breeding. *Nucleic Acids Res.* 42, D1245–D1252. doi: 10.1093/nar/gkt905
- Joshi, T., Patil, K., Fitzpatrick, M. R., Franklin, L. D., Yao, Q., Cook, J. R., et al. (2012). Soybean knowledge base (SoyKB): a web resource for soybean translational genomics. *BMC Genomics* 13(Suppl. 1), S15. doi: 10.1186/1471-2164-13-S1-S15
- Joshi, T., Yan, Z., Libault, M., Jeong, D.-H., Park, S., Green, P. J., et al. (2010). Prediction of novel miRNAs and associated target genes in *Glycine max*. *BMC Bioinform.* 11:S14. doi: 10.1186/1471-2105-11-S1-S14
- Karas, B., Murray, J., Gorzelak, M., Smith, A., Sato, S., Tabata, S., et al. (2005). Invasion of *Lotus japonicus* root hairless 1 by *Mesorhizobium loti* involves the nodulation factor-dependent induction of root hairs. *Plant Physiol.* 137, 1331–1344. doi: 10.1104/pp.104.057513
- Kidner, C. A., and Martienssen, R. A. (2005). The developmental role of microRNA in plants. *Curr. Opin. Plant Biol.* 8, 38–44. doi: 10.1016/j.pbi.2004.11.008
- Kwasniewski, M., Janiak, A., Mueller-Roeber, B., and Szarejko, I. (2010). Global analysis of the root hair morphogenesis transcriptome reveals new candidate genes involved in root hair formation in barley. *J. Plant Physiol.* 167, 1076–1083. doi: 10.1016/j.jplph.2010.02.009
- Kwasniewski, M., Nowakowska, U., Szumera, J., Chwialkowska, K., and Szarejko, I. (2013). iRootHair: a comprehensive root hair genomics database. *Plant Physiol.* 161, 28–35. doi: 10.1104/pp.112.206441
- Lan, P., Li, W., Lin, W.-D., Santi, S., and Schmidt, W. (2013). Mapping gene activity of *Arabidopsis* root hairs. *Genome Biol.* 14, R67. doi: 10.1186/gb-2013-14-6-r67
- Lauressergues, D., Delaux, P.-M., Formey, D., Lelandais-Brière, C., Fort, S., Cottaz, S., et al. (2012). The microRNA miR171h modulates arbuscular mycorrhizal colonization of *Medicago truncatula* by targeting NSP2. *Plant J. Cell Mol. Biol.* 72, 512–522. doi: 10.1111/j.1365-313X.2012.05099.x
- Lelandais-Brière, C., Naya, L., Sallet, E., Calenge, F., Frugier, F., Hartmann, C., et al. (2009). Genome-wide *Medicago truncatula* small RNA analysis revealed novel microRNAs and isoforms differentially regulated in roots and nodules. *Plant Cell* 21, 2780–2796. doi: 10.1105/tpc.109.068130
- Lepek, V. C., D’Antuono, A. L., Tomatis, P. E., Ugalde, J. E., Giambiagi, S., and Ugalde, R. A. (2002). Analysis of *Mesorhizobium loti* glycogen operon: effect of phosphoglucomutase (pgm) and glycogen synthase (g/gA) null mutants on nodulation of *Lotus tenuis*. *Mol. Plant-Microbe Interact.* 15, 368–375. doi: 10.1094/MPMI.2002.15.4.368
- Liang, Y., Tóth, K., Cao, Y., Tanaka, K., Espinoza, C., and Stacey, G. (2014). Lipochitooligosaccharide recognition: an ancient story. *New Phytol.* 204, 289–296. doi: 10.1111/nph.12898
- Libault, M. (2013). Decipher plant gene regulatory networks using single plant cell types. *Int. J. Plant Biol. Res.* 1, 3.
- Libault, M., Brechenmacher, L., Cheng, J., Xu, D., and Stacey, G. (2010a). Root hair systems biology. *Trends Plant Sci.* 15, 641–650. doi: 10.1016/j.tplants.2010.08.010
- Libault, M., Farmer, A., Brechenmacher, L., Drnevich, J., Langley, R. J., Bilgin, D. D., et al. (2010b). Complete transcriptome of the soybean root hair cell, a single-cell model, and its alteration in response to *Bradyrhizobium japonicum* infection. *Plant Physiol.* 152, 541–552. doi: 10.1104/pp.109.148379
- Li, H., Deng, Y., Wu, T., Subramanian, S., and Yu, O. (2010). Misexpression of miR482, miR1512, and miR1515 increases soybean nodulation. *Plant Physiol.* 153, 1759–1770. doi: 10.1104/pp.110.156950
- Matamoros, M. A., Loscos, J., Coronado, M. J., Ramos, J., Sato, S., Testillano, P. S., et al. (2006). Biosynthesis of ascorbic acid in legume root nodules. *Plant Physiol.* 141, 1068–1077. doi: 10.1104/pp.106.081463
- Minorsky, P. V. (2002). The hot and Classic. *Plant Phys.* 129, 438–439. doi: 10.1104/pp.900038
- Misra, B. B., Assmann, S. M., and Chen, S. (2014). Plant single-cell and single-cell-type metabolomics. *Trends Plant Sci.* 19, 637–646. doi: 10.1016/j.tplants.2014.05.005
- Mitra, R. M., and Long, S. R. (2004). Plant and bacterial symbiotic mutants define three transcriptionally distinct stages in the development of the *Medicago truncatula/Sinorhizobium meliloti* symbiosis. *Plant Physiol.* 134, 595–604. doi: 10.1104/pp.103.031518
- Mochida, K., and Shinozaki, K. (2010). Genomics and bioinformatics resources for crop improvement. *Plant Cell Physiol.* 51, 497–523. doi: 10.1093/pcp/pcq027
- Moussaieff, A., Rogachev, I., Brodsky, L., Malitsky, S., Toal, T. W., Belcher, H., et al. (2013). High-resolution metabolic mapping of cell types in plant roots. *Proc. Natl. Acad. Sci. U.S.A.* 110, E1232–E1241. doi: 10.1073/pnas.1302019110
- Nestler, J., Schütz, W., and Hochholdinger, F. (2011). Conserved and unique features of the maize (*Zea mays* L.) root hair proteome. *J. Proteome Res.* 10, 2525–2537. doi: 10.1021/pr200003k
- Nguyen, T. H. N., Brechenmacher, L., Aldrich, J. T., Clauss, T. R., Gritsenko, M. A., Hixson, K. K., et al. (2012). Quantitative phosphoproteomic analysis of soybean root hairs inoculated with *Bradyrhizobium japonicum*. *Mol. Cell. Proteomics MCP* 11, 1140–1155. doi: 10.1074/mcp.M112.018028
- Oldroyd, G. E. (2001). Dissecting symbiosis: developments in Nod factor signal transduction. *Ann. Bot.* 87, 709–718. doi: 10.1006/anbo.2001.1410

- Oldroyd, G. E., and Dixon, R. (2014). Biotechnological solutions to the nitrogen problem. *Curr. Opin. Biotechnol.* 26, 19–24. doi: 10.1016/j.copbio.2013.08.006
- Pang, C. -Y., Wang, H., Pang, Y., Xu, C., Jiao, Y., Qin, Y.-M., et al. (2010). Comparative proteomics indicates that biosynthesis of pectic precursors is important for cotton fiber and *Arabidopsis* root hair elongation. *Mol. Cell. Proteomics MCP* 9, 2019–2033. doi: 10.1074/mcp.M110.000349
- Peña, M. J., Kong, Y., York, W. S., and O'Neill, M. A. (2012). A galacturonic acid-containing xyloglucan is involved in *Arabidopsis* root hair tip growth. *Plant Cell* 24, 4511–4524. doi: 10.1105/tpc.112.103390
- Pu, L., and Brady, S. (2010). Systems biology update: cell type-specific transcriptional regulatory networks. *Plant Physiol.* 152, 411–419. doi: 10.1104/pp.109.148668
- Qiao, Z., and Libault, M. (2013). Unleashing the potential of the root hair cell as a single plant cell type model in root systems biology. *Front. Plant Sci.* 4:484. doi: 10.3389/fpls.2013.00484
- Rogers, E. D., Jackson, T., Moussaieff, A., Aharoni, A., and Benfey, P. N. (2012). Cell type-specific transcriptional profiling: implications for metabolite profiling. *Plant J.* 70, 5–17. doi: 10.1111/j.1365-313X.2012.04888.x
- Rose, C. M., Venkateshwaran, M., Grimsrud, P. A., Westphall, M. S., Sussman, M. R., Coon, J. J., et al. (2012). Medicago PhosphoProtein Database: a repository for *Medicago truncatula* phosphoprotein data. *Front. Plant Sci.* 3:122. doi: 10.3389/fpls.2012.00122
- Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463, 178–183. doi: 10.1038/nature08670
- Subramanian, S., Fu, Y., Sunkar, R., Barbazuk, W. B., Zhu, J.-K., and Yu, O. (2008). Novel and nodulation-regulated microRNAs in soybean roots. *BMC Genomics* 9:160. doi: 10.1186/1471-2164-9-160
- Subramanian, S., and Smith, D. L. (2013). “A proteomics approach to study soybean and its symbiont *Bradyrhizobium japonicum* –a review,” in *A Comprehensive Survey of International Soybean Research—Genetics, Physiology, Agronomy and Nitrogen Relationships*, ed. J. Board (InTech). Available at: <http://www.intechopen.com/books/a-comprehensive-survey-of-international-soybean-research-genetics-physiology-agronomy-and-nitrogen-relationships/a-proteomics-approach-to-study-soybean-and-its-symbiont-bradyrhizobium-japonicum-a-review> [accessed March 4, 2015].
- Sunkar, R., Li, Y.-F., and Jagadeeswaran, G. (2012). Functions of microRNAs in plant stress responses. *Trends Plant Sci.* 17, 196–203. doi: 10.1016/j.tplants.2012.01.010
- Turner, M., Nizampatnam, N. R., Baron, M., Coppin, S., Damodaran, S., Adhikari, S., et al. (2013). Ectopic expression of miR160 results in auxin hypersensitivity, cytokinin hyposensitivity, and inhibition of symbiotic nodule development in soybean. *Plant Physiol.* 162, 2042–2055. doi: 10.1104/pp.113.220699
- van Rhijn, P., Goldberg, R., and Hirsch, A. (1998). Lotus corniculatus nodulation specificity is changed by the presence of a soybean lectin gene. *Plant Cell* 10, 1233–1250.
- Wang, Y., Li, P., Cao, X., Wang, X., Zhang, A., and Li, X. (2009). Identification and expression analysis of miRNAs from nitrogen-fixing soybean nodules. *Biochem. Biophys. Res. Commun.* 378, 799–803. doi: 10.1016/j.bbrc.2008.11.140
- Wan, J., Torres, M., Ganapathy, A., Thelen, J., DaGue, B. B., Mooney, B., et al. (2005). Proteomic analysis of soybean root hairs after infection by *Bradyrhizobium japonicum*. *Mol. Plant Microbe Interact.* 18, 458–467. doi: 10.1094/MPMI-18-0458
- Weiberg, A., Wang, M., Lin, F.-M., Zhao, H., Zhang, Z., Kaloshian, I., et al. (2013). Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science* 342, 118–123. doi: 10.1126/science.1239705
- Xiao, C., and Rajewsky, K. (2009). MicroRNA control in the immune system: basic principles. *Cell* 136, 26–36. doi: 10.1016/j.cell.2008.12.027
- Yan, Z., Hossain, M. S., Arikait, S., Valdes-Lopez, O., Zhai, J., Wang, J., et al. (2015a). Identification of microRNAs and their mRNA targets during soybean nodule development: functional analysis of the role of miR393j-3p in soybean nodulation. *New Phytol.* doi: 10.1111/nph.13365 [Epub ahead of print].
- Yan, Z., Hossain, M. S., Valdés-López, O., Hoang, N. T., Zhai, J., Wang, J., et al. (2015b). Identification and functional characterization of soybean root hair microRNAs expressed in response to *Bradyrhizobium japonicum* infection. *Plant Biotechnol. J.* doi: 10.1111/pbi.12387
- Yan, Z., Hossain, M. S., Wang, J., Valdés-López, O., Liang, Y., Libault, M., et al. (2013). miR172 regulates soybean nodulation. *Mol. Plant Microbe Interact.* 26, 1371–1377. doi: 10.1094/MPMI-04-13-0111-R
- Yokota, K., Fukai, E., Madsen, L. H., Jurkiewicz, A., Rueda, P., Radutoiu, S., et al. (2009). Rearrangement of actin cytoskeleton mediates invasion of *Lotus japonicus* roots by *Mesorhizobium loti*. *Plant Cell* 21, 267–284. doi: 10.1105/tpc.108.063693
- Zhai, J., Arikait, S., Simon, S. A., Kingham, B. F., and Meyers, B. C. (2014). Rapid construction of parallel analysis of RNA end (PARE) libraries for Illumina sequencing. *Methods* 67, 84–90. doi: 10.1016/j.ymeth.2013.06.025

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Hossain, Joshi and Stacey. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The female gametophyte: an emerging model for cell type-specific systems biology in plant development

Marc W. Schmid^{*†}, Anja Schmidt^{*†} and Ueli Grossniklaus^{*}

Department of Plant & Microbial Biology and Zurich-Basel Plant Science Center, University of Zurich, Zurich, Switzerland

OPEN ACCESS

Edited by:

Marc Libault,
University of Oklahoma, USA

Reviewed by:

Matthew Mount Stuart Evans,
Carnegie Institution for Science, USA
Zhenzhen Qiao,
University of Oklahoma, USA

*Correspondence:

Marc W. Schmid
marcschmid@gmx.ch;
Anja Schmidt
aschmidt@botinst.uzh.ch;
Ueli Grossniklaus
grossnik@botinst.uzh.ch

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Plant Systems and Synthetic Biology,
a section of the journal
Frontiers in Plant Science

Received: 09 July 2015

Accepted: 10 October 2015

Published: 03 November 2015

Citation:

Schmid MW, Schmidt A and
Grossniklaus U (2015) The female
gametophyte: an emerging model for
cell type-specific systems biology in
plant development.
Front. Plant Sci. 6:907.
doi: 10.3389/fpls.2015.00907

Systems biology, a holistic approach describing a system emerging from the interactions of its molecular components, critically depends on accurate qualitative determination and quantitative measurements of these components. Development and improvement of large-scale profiling methods (“omics”) now facilitates comprehensive measurements of many relevant molecules. For multicellular organisms, such as animals, fungi, algae, and plants, the complexity of the system is augmented by the presence of specialized cell types and organs, and a complex interplay within and between them. Cell type-specific analyses are therefore crucial for the understanding of developmental processes and environmental responses. This review first gives an overview of current methods used for large-scale profiling of specific cell types exemplified by recent advances in plant biology. The focus then lies on suitable model systems to study plant development and cell type specification. We introduce the female gametophyte of flowering plants as an ideal model to study fundamental developmental processes. Moreover, the female reproductive lineage is of importance for the emergence of evolutionary novelties such as an unequal parental contribution to the tissue nurturing the embryo or the clonal production of seeds by asexual reproduction (apomixis). Understanding these processes is not only interesting from a developmental or evolutionary perspective, but bears great potential for further crop improvement and the simplification of breeding efforts. We finally highlight novel methods, which are already available or which will likely soon facilitate large-scale profiling of the specific cell types of the female gametophyte in both model and non-model species. We conclude that it may take only few years until an evolutionary systems biology approach toward female gametogenesis may decipher some of its biologically most interesting and economically most valuable processes.

Keywords: developmental systems biology, model systems, single cell type isolation, gametophyte, transcriptomics

1. SYSTEMS BIOLOGY: AN INTEGRATED APPROACH TO MODEL BIOLOGICAL PROCESSES WITH LARGE-SCALE DATA

Since the foundation of the Institute for Systems Biology in the year 2000 and the formal definition of systems biology at the beginning of the twenty-first century (Ideker et al., 2001; Kitano, 2002), it has been a steadily growing field of research. As an integrative approach, systems biology is markedly different from the reductionistic approach generally used in molecular biology and genetics. Powered by the central dogma of biology, where a gene is transcribed to mRNA, which is then translated into proteins, molecular biology and genetics have successfully identified genes, their functions, and the processes they are involved in. However, the implicit link of a gene to a certain function or a phenotype is an oversimplification of the underlying process. It thus frequently misses important interactions with other cellular or environmental factors (e.g., responses to environmental conditions like a temperature-dependent phenotype of a mutant). In contrast, systems biology may be described as an attempt to quantitatively and/or qualitatively describe and understand the global behavior of a biological entity, emerging from the interactions between its molecular components. Such a comprehensive understanding would allow the prediction and modeling of the biological entity, its precise control, and ultimately the targeted manipulation of a complex biological system (reviewed in Kitano, 2002; Yuan et al., 2008; Fukushima et al., 2009; Chuang et al., 2010; Katari et al., 2010; Weckwerth, 2011).

Systems biology comprises and integrates experimental studies and large-scale data sets derived from high-throughput technologies (omics), such as transcriptomics (RNA profiling), proteomics (analysis of proteins), and metabolomics (profiling of metabolites). However, also epigenetic regulatory processes based on the modification of chromatin components or DNA (epigenomics), the translation of mRNAs to proteins (translatomics), complex formation of proteins with proteins or nucleic acids (interactomics), the investigation of protein modifications, e.g., phosphorylation important for the regulation of their activity (phospho-proteomics), and the transport of ions or metabolites (fluxomics) need to be taken into account to achieve a full picture of the dynamic processes of a cell or organism (reviewed by Sheth and Thaker, 2014). One of the most crucial aspects for systems biology approaches is the comprehensiveness of the omics data (Kitano, 2002). For a given method this includes the number of items that can be measured at once (e.g., transcripts with transcriptomics). For the entire system, it is then important whether the relevant items (e.g., enzymes and metabolites) or processes (e.g., posttranslational modifications) can be accurately measured with a combination of certain methods. An additional level of complexity may be imposed by the requirement of a high spatial and/or temporal resolution. For a single, isolated cell this can refer to specific organelles, subcellular compartments, certain domains of the plasma membrane, and the stage of the cell-cycle. For an unicellular organism like yeast, this may be augmented

by studying the cell-to-cell variability within the population (Pelkmans, 2012). In multicellular organisms, each cell (type) has a specific function and position within an organ. Its role and differentiation status may be influenced by local signals as well as systemic signals originating from other organs (e.g., hormones). In addition, the temporal coordinate expands to developmental stages of the organs or the life span of the organism.

Consequently, a complete understanding at the systems level requires highly resolved, quantitative spatio-temporal data on the individual components and their interactions, and the integration of the data into models. On one hand, integration of these data with computational methods can aid to characterize previously unknown components (e.g., genes) of a system, as exemplified for yeast (Brown et al., 2006). Alternatively, the data may be used in a mathematical model describing the system and allowing the prediction of a system's behavior and the formulation of hypotheses (Stuel et al., 2007). Finally, the integration of omics data, the formulation of mathematical models, the generation of hypotheses, and the experiments are interlinked and benefit from each other. A possible extension of systems biology is the use of interspecies comparisons to, for example, elucidate the extent to which genotypic variation translates into phenotypic differences (Konstantinidis et al., 2009). Even broader, evolutionary systems biology may be recognized as an approach to describe and understand how biological systems are shaped by evolution and are steering it at the same time (reviewed in Soyer, 2012).

Prior to the understanding of a complex organism composed of many different cell and tissue types, investigations of distinct cell types can lead to an understanding of basic processes governing cellular specification, differentiation, and metabolism. To date, yeast (*S. cerevisiae*) is a widely used model system appreciated as the currently best understood cell (Boone, 2014). While evolutionary only distantly related, pathways in yeast have shown to have considerable similarities to the ones in plants, animals, and humans (Ideker et al., 2001). In addition, yeast serves for the production of food and pharmaceuticals. Due to its simplicity and its importance for biotechnology and biomedical research, yeast has shaped modern molecular biology to a great extent. Indeed, it has been a pioneering organism in systems biology (reviewed in Bostein and Fink, 2011; Österlund et al., 2012; Boone, 2014), starting from gene expression and regulatory networks discovered during early transcriptome studies and their integration with other genome-wide data, over genetic interaction networks obtained by crossing thousands of mutant strains (Costanzo et al., 2010) and modeling of gene expression as a Quantitative Trait Locus (eQTL, Brem et al., 2002), to genome-wide metabolic models. However, given the unicellularity of yeast, it can hardly serve as a developmental model for complex multicellular animals and even less so for plants. In plants, systems biology is less advanced for several reasons, including the higher complexity of most plant genomes, large gene families, the multitude of primary and secondary metabolites, and the lack of suitable *in vitro* systems or cell lines for most plant tissues. Most efforts in plant research thus require *in vivo* experiments, making the procedures generally more difficult and less suitable to

high-throughput approaches. As a consequence, data generation can be a severely limiting factor for plant systems biology. On the other hand, the results are of high relevance for the process under investigation.

Apart from the above mentioned obstacles, substantial progress in the analysis of specific cell types in plants has been made over the last decade. Facilitated by advances in high-throughput profiling technologies and methods for the isolation of individual cell types, recent studies focussed on the analysis of specific cell types or even single cells (Figure 1). To investigate cell type-specific processes in higher plants, root hairs and trichomes have been used as models, both for their physiological importance and their accessibility at the epidermal surface (for details see below; Ishida et al., 2008; Brechenmacher et al., 2009, 2010; Dai et al., 2010; Libault et al., 2010a,b; Schilmiller et al., 2010; Nestler et al., 2011; Van Cutsem et al., 2011; Dai and Chen, 2012; Rogers et al., 2012; Tissier, 2012; Qiao and Libault, 2013). In addition, starting with only a few examples at the beginning of the twenty-first century (Kehr, 2001), cell type-specific transcriptional profiling has become a robust and frequently used method. In the model plant *Arabidopsis thaliana*, novel insights into plant development and cellular responses to environmental stimuli were for example gained through studies on individual cell types of the root, root hairs, trichomes, and guard cells, and by transcriptional profiling during male and female gametogenesis (reviewed in Taylor-Teeples et al., 2011; Schmidt et al., 2012; Wuest et al., 2013). These examples clearly illustrate the importance of cell type-specific investigations for a detailed understanding of differentiation processes and environmental responses of distinct cell types. However, depending on the cell type under investigation, the currently available methods for cell isolation may be challenging, time-consuming, or limited to a subset of omics approaches (e.g., Laser-Assisted Microdissection (LAM) of rare cell types, Wuest et al., 2013). While studies focusing on specific cell types, which can be isolated in quantities high enough for the full set of omics approaches, can serve as initial models for cell type-specific systems biology in plants (Libault et al., 2010a), the ultimate goal must be that the full set of methods can be applied to any cell type of interest.

2. METHODS FOR THE ACQUISITION OF LARGE-SCALE QUANTITATIVE DATA FROM SPECIFIC CELL TYPES

Large-scale profiling of distinct cell types critically depends on the possibility to isolate these cells in sufficient purity and quantity, as well as the sensitivity and accuracy of the profiling methods. Despite the rapid improvements of established and novel tools for systems biology, the demand for fast and easily applicable methodologies for cell type-specific analyses is not yet satisfied. Further challenges are associated with the requirement for normalization and integration of different data types, and the increasing demand for platforms allowing storage and sharing of the rapidly growing amount of large-scale datasets (reviewed by Chuang et al., 2010; Katari et al., 2010; Gomez-Cabrero

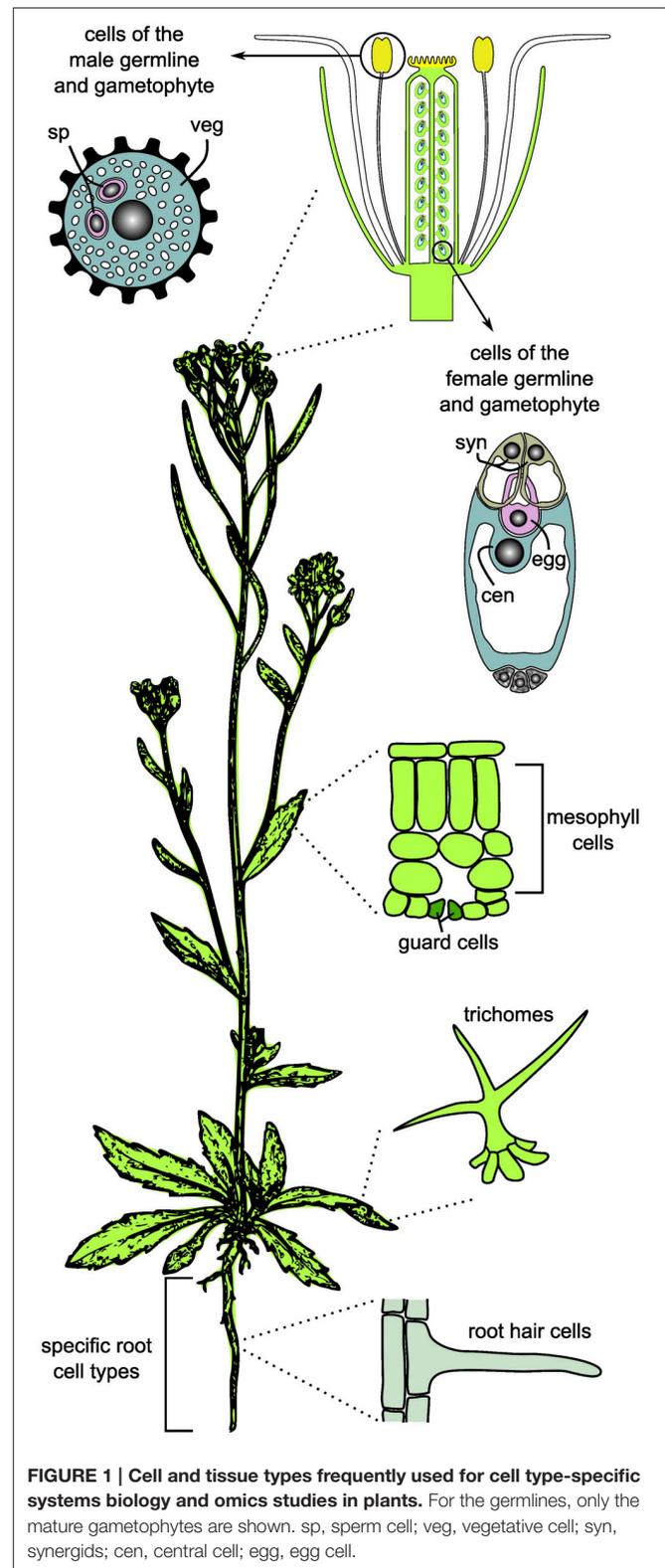


FIGURE 1 | Cell and tissue types frequently used for cell type-specific systems biology and omics studies in plants. For the germlines, only the mature gametophytes are shown. sp, sperm cell; veg, vegetative cell; syn, synergids; cen, central cell; egg, egg cell.

et al., 2014; Sheth and Thaker, 2014). In brief, three steps are of great importance for cell type-specific systems biology: (i) isolation and purification of the specific cell type, (ii) profiling

of the selected molecular compounds, and (iii) data analysis, integration, storage, and sharing. In the following sections, we will present current methods to acquire large-scale quantitative data required for systems biology. We will focus on methods allowing genome-wide cell type-specific analyses and present representative examples. For a discussion on the computational challenges in systems biology, the reader is referred to several recent reviews (Ahrens et al., 2007; Yuan et al., 2008; Fukushima et al., 2009; Chuang et al., 2010; Katari et al., 2010; Liberman et al., 2012; Fukushima et al., 2014; Gomez-Cabrero et al., 2014; Robinson et al., 2014).

2.1. Methods for the Isolation of Specific Cell Types

A few cell types in plants are exposed on the surfaces of tissues and can be collected by abrasion or mechanical detachment. Depending on the species, relatively simple mechanical isolation procedures for trichomes and root hairs enabled a large spectrum of methods. Mechanical isolation of trichomes allowed transcriptomics and metabolomics in various species (for an integrated database see Dai et al., 2010) and proteomics (Schillmiller et al., 2010; Van Cutsem et al., 2011). Another example for an exposed cell type are root hairs, for which relatively simple isolation procedures facilitated transcriptomics (Libault et al., 2010b), proteomics (Brechenmacher et al., 2009; Nestler et al., 2011), and metabolomics (Brechenmacher et al., 2010). Certain other cell types can be isolated by tissue disruption, followed by centrifugation-based methods or manual isolation of the dissociated cells under a microscope using a micropipette (eventually with a marker for the cell type of interest). Examples include specific cell types from the male or female reproductive lineages, plant mesophyll cells, and guard cells (reviewed by Dai and Chen, 2012; Schmid et al., 2012; Wuest et al., 2013). Proteomic profiling has, for example, been performed on *Brassica napus* guard cells and mesophyll cells that could be purified as protoplasts (Zhu et al., 2009).

However, for most cell types these methods are not applicable. Several methods for the isolation of specific cell types embedded in differentiated tissues have been established. Fluorescent Activated Cell Sorting (FACS) can be used to sort fluorescent cells based on their light scattering characteristics and fluorescence (reviewed by Hu et al., 2011). This method allowed high resolution transcriptional profiling of different cell types in the *Arabidopsis* root, and, more recently, proteomics (Petricka et al., 2012) and metabolite mapping of selected root cell and tissue types (Brady et al., 2007; reviewed by Benfey, 2012; Moussaieff et al., 2013). Similarly, Fluorescence-Activated Nuclei Sorting (FANS) has been established and, for example, used to isolate endosperm nuclei for profiling of RNA activity or epigenetic modifications (Weinhofer et al., 2010; Weinhofer and Köhler, 2014). Despite the great potential of FACS/FANS for plant cell type-specific systems biology, both approaches have certain limitations: They can only be applied if transgenic lines carrying cell type-specific fluorescent markers can be established, and they are thus not suitable for most non-model species. In addition, depending on the tissue type, longer enzymatic incubations are

required to digest the cell walls and to release the protoplasted cells prior to sorting (Evrard et al., 2012). Consequently, changes in, for example, the transcriptome or metabolome cannot be fully excluded. Alternatively, the INTACT method (Isolation of Nuclei TAgged in specific Cell Types) allows the isolation of nuclei expressing a biotinylated nuclear envelope protein by affinity purification with streptavidin-coated beads (Deal and Henikoff, 2011). This method is suitable to study epigenetic modifications (DNA methylation of histone modifications) and to profile the RNA within the nucleus. To study actively translated mRNAs bound to ribosomes (translatome), small epitope tags can be fused to a ribosomal protein to allow immunopurification of the ribosomes containing the mRNAs with a method named TRAP (Translating Ribosome Affinity Purification; reviewed in Bailey-Serres, 2013). Alternatively, RNAs binding to RNA binding proteins involved in the formation of ribonucleoprotein (RNP) complexes can be profiled by immunoprecipitation of an epitope-tagged protein (RNP ImmunoPurification, RIP; Bailey-Serres, 2013). It has to be noted that the analyses of transcriptome and translome abundance will not give the same results, because not all mRNAs present in a cell are actively translated at a given time point. In this respect, profiling of mRNAs bound to ribosomes gives complementary results to transcriptome profiling as the readouts are closer to the synthesis of proteins (Bailey-Serres, 2013). Similar to FACS and FANS, also INTACT, TRAP, and cell type-specific RIP require the use of transgenic lines and pre-existing knowledge about cell type-specific promoters or markers.

An alternative method not requiring any molecular knowledge is LAM (Kerk et al., 2003). Plant tissues are thereby typically fixed and embedded in paraffin wax (reviewed in Schmid et al., 2012; Wuest et al., 2013) or resin (Tucker et al., 2012; Okada et al., 2013). Thin sections of the tissues (typically between 6 and 10 μm) are subsequently mounted on metal framed plastic slides and used to isolate the cell types of interest after resolving the wax or resin and drying the tissues on the slides (Okada et al., 2013; Wuest and Grossniklaus, 2014). Alternatively, the tissue may also be embedded in optimal cutting temperature compound for cryosectioning, followed by on-slide tissue dehydration and LAM (Kelliher and Walbot, 2012, 2014). The main constraint of LAM is that harvesting sufficient material for downstream omics methods can be very time-consuming. Furthermore, the suitability for single cell isolation depends on the optical resolution in sectioned tissues and the recognizability of the cell type of interest. In addition, the physical properties of the laser beam of the instrument used can impose limitations on which cell types can be isolated (Schmid et al., 2012). Thus, the time required for collecting enough material for one sample is largely dependent on the cell type of interest. So far, the applications of LAM for cell type-specific omics have been restricted to transcriptional profiling, e.g., to study cell type-specification in the female reproductive lineage in *Arabidopsis thaliana*, *Boechera gunnisoniana*, and *Hieracium praealtum* (Wuest et al., 2010; Schmid et al., 2011, 2014; Schmid et al., 2012; Okada et al., 2013). However, other applications, such as genome wide profiling of DNA methylation, are likely feasible (see below).

2.2. Methods for Data Acquisition

2.2.1. Transcriptomics

Transcriptome profiling encompasses the identification and quantification of all expressed RNA transcripts at a given time point (mRNA, tRNA, microRNA). However, due to the frequent use of oligo-dT priming during cDNA synthesis or the hybridization to microarrays covering only coding regions of the genome, many studies are restricted to mRNAs or a subset of mRNAs. Several types of microarrays were produced and extensively used for the analyses of gene expression in different plant species, including the model plant *Arabidopsis thaliana* and different important crop species like maize, rice, and barley (reviewed in Sheth and Thaker, 2014). The Affymetrix ATH1 GeneCHIP (www.affymetrix.com), the most popular microarray for *Arabidopsis* has for example been used to profile a large variety of different tissue types (e.g., Schmid et al., 2005), specific cell types of the root isolated through FACS (Birnbaum et al., 2003), and specific cell types of the male and female reproductive lineages (reviewed in Schmidt et al., 2012). In addition to well established tools for data analysis, the wealth of publicly available datasets generated on the same platform makes commonly used microarrays a very valuable tool for systems biology (Katari et al., 2010).

Apart from microarrays, several platforms for Next Generation Sequencing (NGS) have been developed over the last years and are now routinely used for transcriptional profiling (RNA-Seq; see Mardis, 2013, for a review on NGS platforms). RNA-Seq has several advantages as compared to the use of microarrays, including a higher dynamic range, higher sensitivity, and whole-genome coverage allowing the identification of previously unknown transcripts and splice variants (reviewed in Schmidt et al., 2012). A major advantage is the applicability to non-model species, either through *de novo* assembly of the short reads into transcripts or by the use of a reference transcriptome either produced separately or taken from a public database (e.g., the ongoing effort to sequence 1000 plant transcriptomes, www.onekp.com). Examples for such an approach are the central cells of *Arabidopsis thaliana*, and cells of the female reproductive lineage in *Hieracium praealtum* and *Boechera gunnisoniana* (Schmid et al., 2012; Okada et al., 2013; Schmidt et al., 2014). Several tools for RNA-Seq data analysis are available (see Sheth and Thaker, 2014, for a selection of software tools, and Schmid and Grossniklaus, 2015, for Rcount, a count tool addressing the problem of reads aligning at multiple locations in the genome, or reads aligning at positions where two or more genes overlap). Current challenges are the increasing demand for standardized annotations of datasets and the development of computational methods allowing the integration of data from different studies using different methods and platforms. In the future, the integration of data from different species will be of great value for plant systems biology, allowing researchers to gain insights into conserved common regulatory mechanisms, environmental adaptations, and evolutionary changes.

2.2.2. Proteomics

In addition to the analysis of gene expression and actively translated mRNAs, the investigation of proteins and protein modifications (e.g., phospho-proteomics and glyco-proteomics) add additional levels of complexity. From a systems biology perspective the aim is the combination of cell type-specific proteomics with transcriptomics and metabolomics to elucidate and model regulatory networks (reviewed in Dai and Chen, 2012). In the beginning of proteomics, 2D gel electrophoresis was frequently used for separation of the proteins in a sample and to identify spots representing proteins differentially occurring in two samples (reviewed by Schulze and Usadel, 2010). However, the protein or protein mixture in one spot could only be identified by excising the spot and analysis using Mass Spectrometry (MS). To date, proteomics largely depends on the use of various MS methods in combination with different protein separation procedures. Typically, proteins are first digested with trypsin and subsequently either analyzed directly by MS or first separated by chromatography before MS. MS methods have greatly improved with the development of soft ionization methods like ElectroSpray Ionization (ESI) in solution (typically aqueous or organic solvents) or Matrix Assisted Laser Desorption Ionization (MALDI, Hollenbeck et al., 1999; Schulze and Usadel, 2010). By both methods, intact gas phase ions are generated that are introduced into mass analyzers and sorted depending on their mass-to-charge ratio, e.g., using their Time-Of-Flight (TOF, Hollenbeck et al., 1999; for a recent summary of mass analyzers see Lee et al., 2012; for a description of Orbitrap mass analyzers see Perry et al., 2008). However, detection based on peptide mass-to-charge ratios is largely qualitative and can only be used for quantification in two or more samples acquired under standardized conditions (Schulze and Usadel, 2010). Thus, stable isotope or chemical labeling is frequently applied for quantification in proteomic methods (reviewed in Schulze and Usadel, 2010). While software and algorithms for protein identification are well established, quantitative analysis remains more challenging (Schulze and Usadel, 2010; Sheth and Thaker, 2014, see Sakata and Komatsu, 2014, for a recent survey on proteomics repositories and databases).

To date, only a restricted number of plant cell types have been profiled in a cell type-specific manner by proteomics, including guard cells, mesophyll cells, trichomes, root hair cells, leaf epidermal cells, lily and rice sperm cells, different stages of pollen development in tobacco, *Arabidopsis*, and tomato, and rice egg cells (Brechenmacher et al., 2009; Grobei et al., 2009; Abiko et al., 2013; Chaturvedi et al., 2013; Ischebeck et al., 2014, and reviewed by Dai and Chen, 2012; Wuest et al., 2013). As compared to transcriptomics approaches, a larger amount of starting material is required. For example, approximately 40 μ g of protein were isolated to study the proteome during tobacco pollen development (Ischebeck et al., 2014). In addition, the amount of proteins detected is typically in the range of 10–30% of the transcripts identified from the same cell or tissue type, as exemplified by a study on *Arabidopsis* pollen, in which 3599 proteins as compared to 11,150 expressed genes were reported (Grobei et al., 2009). This quantitative difference largely reflects

the difference in the sensitivity of the methods and likely only to a smaller extent meaningful biological differences. Nevertheless, as only a few proteins have been identified in previous studies, e.g., from maize egg cells, these data reflect a great improvement (Okamoto et al., 2004), and a rapid advance since the shaping of the term proteomics in 1997 (James, 1997).

2.2.3. Protein-Protein Interactions

For studies of protein-protein interactions, the major methods used are Yeast Two-Hybrid (Y2H) screens, Affinity Purification Mass Spectrometry (AP-MS), or Bimolecular Fluorescence Complementation (BiFC) (reviewed in Zhang et al., 2010). Y2H assays take advantage of the bipartite structure of the yeast GAL4 transcriptional activator consisting of two functional domains, a transcription activation domain and a DNA-binding domain. In Y2H assays, the bait and the target protein are fused to the two functional domains of GAL4, respectively, together reconstituting the functional GAL4 protein that binds to its target promoter (UAS_{GAL4}) to activate the expression of a downstream gene encoding a selectable marker. Apart from a high false-positive rate, the use of yeast itself is a major drawback of the method. While cell type-specific cDNA libraries can be used to profile pairwise protein interactions, the system does not truly reflect the *in vivo* state of a specific plant cell (e.g., cofactors of an interaction may be missing). Several systems similar to Y2H assays have been established to specifically study membrane proteins (e.g., split-ubiquitin system; Obrdlik et al., 2004; Chen et al., 2010). For AP-MS, a bait protein is fused to an affinity tag for expression *in vivo*. The tagged protein of interest is subsequently purified as a complex with interacting proteins or other molecules and assayed by MS. This method is also associated with a relatively high false-positive rate due to protein contaminants. While the method is well-suitable for cell type-specific studies if the expression of the tagged protein is driven by a cell type-specific promoter, true omics-scale profiling can hardly be achieved, as a precondition would be the cell type-specific tagging of all proteins represented in a cell. This also holds true for BiFC, where a fluorescent protein (YFP, RFP, or GFP) is split into two non-fluorescent halves that are reconstituted to a fluorescent protein upon interaction of the bait and target proteins they are fused to (reviewed by Zhang et al., 2010). While BiFC has the advantage that spatial and temporal interactions can be resolved, it is also associated with a high false-positive rate. Consequently, methods for true cell type-specific large-scale protein-protein interaction studies in plants are lacking to date. Nonetheless, the currently available data on protein-protein interactions, as for example the recently established membrane protein interactome (Chen et al., 2012), may help to resolve certain dependencies within regulatory networks (see Sheth and Thaker, 2014, for a summary of the available databases).

2.2.4. Protein-DNA Interactions

Interactions between proteins and DNA comprises several functional aspects, for example nucleosome occupancy, specific histone modifications, or transcription factor binding. These interactions may be studied using either Chromatin

ImmunoPrecipitation (ChIP, Orlando and Paro, 1993), or DNA adenine methyltransferase IDentification (Dam-ID, van Steensel and Henikoff, 2000). In both cases, the interaction of one protein (variant) with the DNA is monitored genome-wide. During the ChIP procedure, the DNA is cross-linked by formaldehyde to bound proteins before fragmentation by sonication. Chromatin fragments are then isolated with antibodies against the protein (variant) of interest. After recovery of the co-purified DNA by reverting the cross-links, the DNA sequence can be identified using microarray hybridization or high-throughput sequencing (He et al., 2011). Protocols facilitating cell type-specific ChIP (Chromatin Affinity purification from Specific cell Types by ChIP; CAST-ChIP), without the need for purification of the cell type of interest or a protein-specific antibody, have been developed (Schauer et al., 2013). However, these protocols rely on transgenics and specific promoters. In addition, we are not aware of a report where this method has been applied in plants or used to study rare cell types.

For Dam-ID, the protein of interest is fused to an adenine-methyltransferase of *E. coli* (Dam, Greil et al., 2006). Endogenous methylation of adenine is absent in most eukaryotes. Upon expression of the fusion protein, Dam is targeted to the native binding sites of the protein fused to it. This results in a localized methylation of adenines in the GATC sequence context. These regions can then be identified using methylation-sensitive restriction enzymes and microarray hybridization or high-throughput sequencing (Greil et al., 2006; Luo et al., 2011). Tissue or cell type-specific expression of the fusion protein can be used to overcome the need for cell isolation and has been shown to be highly specific (targeted DamID, “TaDa,” Southall et al., 2013). The major disadvantages of the method are the requirement for transgenics and specific promoters, as well as the need for optimization of the expression level to avoid untargeted methylation and toxicity of the Dam fusion protein. Thus, both approaches are currently quite laborious and generally only applicable to model-species. Nonetheless, especially transcription factor binding is of great value for the study of transcriptional networks (Yuan et al., 2008). If cell type-specific data is not available, previously identified transcription factor binding motifs may still help to identify transcriptional modules (Diez et al., 2014).

2.2.5. Protein Microarrays

Protein microarrays are a promising tool for proteomics as well as for interactions of proteins with other proteins, nucleic acids, cellular surface markers, or posttranslational protein-modifications (Yang et al., 2011b; Uzoma and Zhu, 2013). Several different types of protein microarrays can therefore be distinguished. On analytical microarrays, well characterized proteins (e.g., monoclonal antibodies) are spotted to identify a specific set of proteins. Alternatively, less well characterized proteins (e.g., lysates from whole cells) are spotted on functional microarrays to test for interaction partners. Finally, proteome microarrays hold the majority of encoded proteins for an organism (Yang et al., 2011b). While the first proteome microarray for budding yeast was established in 2001 (reviewed by Uzoma and Zhu, 2013), not many applications were

reported in plants (Yang et al., 2011b; Uzoma and Zhu, 2013). Nevertheless, protein microarrays have, for example, successfully been used to study 802 transcription factors in *Arabidopsis* (almost half of all transcription factors annotated in *Arabidopsis*, Gong et al., 2008). While protein microarrays may have a high potential for applications in systems biology, they are currently still limited by high production costs and laborious production methods (e.g., large-scale cloning of open reading frames, protein purification, and production of high-affinity monoclonal antibodies, Yang et al., 2011b).

2.2.6. Metabolomics

Due to the high complexity of plant metabolites coming from both primary and secondary metabolism, the plant metabolome is highly complex. Although by far not comprehensively elucidated to date, about 200,000 different metabolites are estimated to be represented in plants (reviewed by Sheth and Thaker, 2014). While a variety of analysis platforms can in principle be applied for metabolite detection, Nuclear Magnetic Resonance (NMR) and MS are the most frequently used methods (Kueger et al., 2012; Sheth and Thaker, 2014). High resolution mapping of metabolites has recently been achieved in *Arabidopsis* roots by combining FACS with high resolution MS (Moussaieff et al., 2013). In addition, glandular trichomes have been used as model systems for large-scale metabolome analyses (Tissier, 2012). However, the major limitation of current metabolomics is the lack of a single method allowing comprehensive measurements in terms of qualitative detection, quantitation, and spatio-temporal resolution. This is the case because the metabolites differ significantly in their concentration, chemical properties, and analytical behavior. Two major strategies in metabolome profiling are the use of either targeted or untargeted MS (reviewed in Kueger et al., 2012). Targeted MS relies on previous knowledge about structures and chemical properties of the metabolites of interest and combines chromatographic separation techniques, e.g., High Pressure Liquid Chromatography (HPLC) or Gas Chromatography (GC), with MS techniques. In contrast, non-targeted analyses using MS without prior chromatographic separation is used to profile metabolites without prior knowledge about their abundance or structure. This method often only allows the determination of metabolic signatures, as the characterization of a specific metabolite, for example by NMR, is highly challenging. Therefore, a key problem is the availability of reference spectra and compounds for compound identification and annotation (Kueger et al., 2012). Thus, the need for comprehensive databases including relevant information on the compounds, e.g., spectra, and the requirement for integration of metabolome data with other large-scale omics data has been noted (Fukushima et al., 2009). Current online resources include the Golm Metabolome Database (gmd.mpimp-golm.mpg.de) and the MASSBANK Database (www.massbank.jp).

An alternative method to study, for example, metabolites at spatial resolution without the need for prior cell isolation is MALDI-MS Imaging (MSI, reviewed by Lee et al., 2012). For MSI, a suitable matrix is directly applied to thin tissue sections (e.g., 10–20 μm). The prepared tissue sections are

then rasterized with a laser-beam coupled to a high mass resolution (TOF-MS, reviewed in Kaspar et al., 2011). The spot size of the laser thereby determines the resolution. Only recently, technical improvements allowed to reach resolutions required for the analysis of single cells (<20 μm , reviewed in Kueger et al., 2012; Lee et al., 2012). MSI has rarely been used in plants for proteomics, and only few studies reported the imaging of metabolites (reviewed in Kaspar et al., 2011; Kueger et al., 2012). Examples for metabolite imaging with MSI include the measurement of wheat grain cell-wall polysaccharides (Veličković et al., 2014, 100 μm spot size), or the lipid measurements in embryos of cotton (Horn et al., 2012, 35 μm spot size). While MSI has a great potential for cell type-specific studies for plant systems biology, it needs to be noted that only thin surface layers of <1 μm are sampled by MALDI (Lee et al., 2012). However, further improvements in MSI are likely to be developed soon and adaption of these methods to plant tissues may once facilitate single-cell proteomics as well as metabolomics in a range of species.

In addition, to study the subcellular localization of specific ions or metabolites and their physiological relocation, e.g., by directed transport, a variety of molecular sensors has recently been developed. Such sensors usually depend on proteins changing their conformation upon binding of a specific substrate. Consequently, the distance between attached fluorescent proteins will change leading to an alteration in Fluorescent Resonance Energy Transfer (FRET, reviewed by Okumoto, 2012; Okumoto et al., 2012). For spatially and temporally resolved measurements, FRET can be measured by, for instance, Fluorescence Lifetime Imaging Microscopy (FLIM, reviewed by De Los Santos et al., 2015). While being very valuable tools in plant research, these techniques do not readily allow the high-throughput analysis of a large number of compounds in a plant cell and will thus not be discussed in detail in this review.

2.2.7. DNA Methylation

DNA (cytosine) methylation is a heritable epigenetic modification of the genome and is involved in various cellular and developmental processes in a wide range of species, including animals, fungi, and plants. Several methods for genome-wide profiling of the DNA cytosine methylation status have been established. These include the hybridization onto whole-genome DNA microarrays after digestion of genomic DNA with methylation-sensitive restriction enzymes, or the precipitation of methylated DNA with antibodies targeting methylated cytosines (Methylated DNA ImmunoPrecipitation, MeDIP), followed by either microarray hybridization (MeDIP-chip) or NGS (MeDIP-Seq, reviewed by Su et al., 2011; Ji et al., 2015). The current method of choice for methylome profiling is Whole-Genome Bisulfite Sequencing (WGBS). In brief, DNA is incubated with bisulfite, converting all unmethylated cytosines to uracils, which are identified as thymines during sequencing. In contrast, all methylated cytosines are protected from the conversion, remain unchanged, and are identified as cytosines during sequencing (Ji et al., 2015). Compared to the profiling of other epigenetic marks, such as histone modifications, WGBS has two major advantages. It does not require the use of transgenic plants or antibodies, and

recently developed methods allow WGBS on as little as 125 pg of DNA (Post-Bisulfite Adaptor Tagging (PBAT), Miura et al. (2012); 20 pg diluted *Arabidopsis* DNA with a modified protocol, our unpublished data). WGBS is therefore a very promising method for the profiling of specific cell types in plants.

3. SYSTEMS BIOLOGY APPROACH TOWARD PLANT DEVELOPMENT

As evident from the previous examples, plant cell type-specific systems biology is most advanced in cell types that can relatively easily be isolated in large enough amounts of suitable for any type of omics approach. For the root hairs of soybean, for example, a promising method to isolate large quantities facilitating any omics analysis has recently been described and will likely be of great use (Qiao and Libault, 2013). The method uses an ultrasound aeroponic system to enhance root hair density, followed by fixation and separation of the root hairs in liquid nitrogen. In addition, for the different cell types of the *Arabidopsis* root, FACS yields sufficient material for most omics approaches. An advantage of these systems is that due to the use of only one isolation method, the variability imposed by it can be held constant over all experiments. The use of a single method is also cost-efficient as it requires less time and resources to optimize only one method as compared to several. Due to the relatively easy sample collection and their physiological roles, roots, root hairs, and trichomes are excellent models to study responses to environmental stimuli, host-pathogen/symbiont interactions, metabolic pathways, or the dynamics of cellular specification and cell-cell communication in complex tissues. However, even the root may not be an optimal model to address fundamental questions of developmental systems biology. Its main disadvantages are the long developmental time span, starting very early during embryogenesis, and the complex interplay within and between the different cell types of the root, but also with the above-ground tissues, and biotic and abiotic environmental factors. Ideally, a developmental model system should allow an experimental coverage of the entire life-span of the organism. It would be of advantage if the organism were short-lived and comprise only a limited number of developmental stages and specialized cell and tissue types to reduce complexity and increase the affordability of comprehensive studies. For comparative analyses and evolutionary systems biology approaches, it would be further advantageous if the phylogeny of the model system included a broad range of organisms with gradual phenotypic changes, or with gain, loss, and alternative usage of modular building blocks. Finally, an ideal model system is most beneficial if its understanding can lead to direct applications in, for example, production of food or pharmaceuticals.

An intuitive model for the development of an organism is the embryo. During plant embryogenesis, the basic body organization with an apical-basal and radial pattern is established starting from a single cell, the zygote. The mature embryo already contains the progenitors of the main organizers of plant growth, the primary Shoot and Root Apical Meristems (SAM and RAM),

and the hypocotyl and cotyledons with their various tissue types (reviewed in Lau et al., 2012). However, it is thus already a relatively complex system composed of multiple cell and tissue types. Additional complexity is imposed by the different stages of embryo development, spanning the time between the one-cellular zygote and the mature embryo. An in-depth systems biological description of embryogenesis would therefore require sampling of a large variety of cell types at many time points. Nevertheless, while most transcriptional studies published so far focussed on whole tissues or entire embryos (reviewed in Palovaara et al., 2013; Zhan et al., 2015), recently, high-quality cell type-specific transcriptomes of the proembryo and the suspensor of the early stages of the *Arabidopsis* embryo were described (Slane et al., 2014).

Alternative models for the development of organisms, which are far less complex than the embryo, are the gametophytes of flowering plants: the pollen (male) and the embryo sac (female). They are typically formed from one spore (meiotic product) and, at maturity, they consist of only a few cells and cell types, including the male and female gametes, the sperm cells and the egg and central cells, respectively (reviewed in Yang et al., 2010; Twell, 2011; Schmid et al., 2015). Upon double fertilization, the egg cell and the central cell fuse with one sperm each to give rise to the embryo and endosperm, respectively. The latter nurtures the embryo and acts as storage organ for seed reserves in many species, including the cereals. The endosperm is thus the most important food and feed source.

Given the sheer amount of pollen produced by a single plant, and the relatively simple isolation procedures for some of the specific cell types of the male germline in developing pollen, multiple cell type-specific transcriptome data sets are available from different species, including *Arabidopsis thaliana*, *Oryza sativa* (rice), *Zea mays* (maize), *Lilium longiflorum* (lily), and *Plumbago zeylanica* (white leadwort) (Table 1; reviewed in Schmid et al., 2012; Anderson et al., 2013; Dukowic-Schulze et al., 2014; Kelliher and Walbot, 2014), and several cell type-specific proteomes have recently been described for tobacco, *Lilium davidii* var. *unicolor* (Lanzhou lily), and tomato (Table 1; Abiko et al., 2013; Chaturvedi et al., 2013; Zhao et al., 2013; Ischebeck et al., 2014). Due to its characteristic tip-growth, pollen tubes also serve as an excellent model to study cell elongation and mechanical properties of the cell wall (Vogler et al., 2013). However, pollen development is strikingly uniform in angiosperms (Maheshwari, 1950), and inter-species comparisons would therefore likely be more fruitful in gymnosperms, which show a remarkable variation in terms of the number of cell divisions between meiosis and the subsequent specification of the sperm cells (Fernando et al., 2010). In contrast to pollen, a plant forms much fewer female gametophytes, which are deeply embedded in the maternal floral tissue (e.g., in *Arabidopsis*, each flower contains around 50 ovules, each of which harbors only one embryo sac). Nonetheless, several cell type-specific transcriptomes (Table 2; reviewed in Schmid et al., 2012; Wuest et al., 2013, and more recent data in Anderson et al., 2013; Okada et al., 2013; Schmid et al., 2014) as well as a proteome analysis for rice egg cells (Table 2; Abiko et al., 2013) are currently available. Even though it is more difficult to collect than the pollen,

TABLE 1 | Summary of transcriptome (top) and proteome (bottom) datasets generated for specific cell types during formation of the male reproductive lineage and gametogenesis.

Species	Cell type	Profiling method	Literatures
<i>Oryza sativa</i> ssp. <i>Japonica</i>	Meiocyte	44K Agilent microarray	Tang et al., 2010
<i>Zea mays</i>	Meiocyte	RNA-Seq (Illumina)	Dukowic-Schulze et al., 2014
<i>Arabidopsis thaliana</i>	Meiocyte	RNA-Seq (SOLiD)	Yang et al., 2011a
<i>Arabidopsis thaliana</i>	Meiocyte	RNA-Seq (Illumina)	Chen et al., 2010
<i>Arabidopsis thaliana</i>	Meiocyte, UNM	CATMA microarray	Libeau et al., 2011
<i>Arabidopsis thaliana</i>	UNM	Affymetrix ATH1 microarray	Honys and Twell, 2004
<i>Oryza sativa</i> ssp. <i>Japonica</i>	UNM	Affymetrix rice genome array	Wei et al., 2010
<i>Lolium longiflorum</i>	GC	cDNA microarray	Okada et al., 2007
<i>Arabidopsis thaliana</i>	SC	Affymetrix ATH1 microarray	Borges et al., 2008
<i>Plumbago zeylanica</i>	SC	cDNA spotted microarray	Gou et al., 2009
<i>Oryza sativa</i> ssp. <i>Japonica</i>	SC	RNA-Seq (Illumina)	Anderson et al., 2013
<i>Nicotiana tabacum</i>	Meiocyte, tetrad	gel LC-MS	Ischebeck et al., 2014
	UNM, polarized UNM		
<i>Solanum lycopersicum</i> (ecotype Red Setter)	Meiocyte, tetrad, UNM, polarized UNM	gel LC-Orbitrap-MS	Chaturvedi et al., 2013
<i>Lilium davidii</i> var. <i>unicolor</i>	SC, GC	MS/MS with MALDI-TOF/TOF	Zhao et al., 2013
<i>Oryza sativa</i> ssp. <i>Nipponbare</i>	SC	LC-MS/MS	Abiko et al., 2013

In brief, pollen formation starts with a microspore mother cell (or meiocyte) which undergoes meiosis to give rise to a tetrad of reduced spores. Each of these microspores undergoes pollen mitosis I to give rise to a generative and a vegetative cell. The subsequent mitotic division of the generative cell (pollen mitosis II) results in the formation of two sperm cells (Twell, 2011). UNM, uninucleate microspore; GC, generative cell; SC, sperm cell; LC, liquid chromatography.

TABLE 2 | Summary of transcriptome (top) and proteome (bottom) datasets generated for specific cell types during formation of the female reproductive lineage and gametogenesis.

Species	Cell type	Profiling method	Literatures
<i>Arabidopsis thaliana</i>	MMC	ATH1 microarray	Schmidt et al., 2011
<i>Arabidopsis thaliana</i>	egg, cen, syn	ATH1 microarray	Wuest et al., 2010
<i>Arabidopsis thaliana</i>	cen	RNA-Seq (SOLiD)	Schmid et al., 2012
<i>Arabidopsis thaliana</i>	egg, syn	RNA-Seq (SOLiD)	Schmidt et al., 2014
<i>Oryza sativa</i> ssp. <i>Nipponbare</i>	egg, syn	44K Agilent microarray	Ohnishi et al., 2011
<i>Boechera gunnisoniana</i>	AIC, egg, cen, syn	ATH1 microarray, RNA-Seq (SOLiD)	Schmidt et al., 2014
<i>Hieracium praealtum</i>	AI	RNA-Seq (Roche 454)	Okada et al., 2013
<i>Oryza sativa</i> ssp. <i>Nipponbare</i>	egg	LC-MS/MS	Abiko et al., 2013

MMC, megaspore mother cell; AIC, apomictic initial cell; AI, aposporous initial cell; egg, egg cell; syn, synergids; cen, central cell; LC, liquid chromatography.

the embryo sac has certain developmental features rendering it a highly interesting model system for plant development: (i) high evolutionary diversity within angiosperms, (ii) syncytial development (i.e., the formation of a multinucleate cell), (iii) specification and differentiation of only three to four distinct cell types, and (iv) a process in which plants can reproduce asexually via seeds (gametophytic apomixis).

The mature embryo sacs of angiosperms generally contain at least three distinct cell types: the synergids required for pollen tube attraction and reception, and the two gametes, the egg and the central cell. An exception are, for example, the *Podostemaceae*, where the central cell seems to degenerate before pollen tube arrival, resulting in a single fertilization event (Sehgal et al., 2014). In addition, antipodal cells are

frequently present, but little is known about their function. It has been hypothesized that they might be involved in nutrient transfer from the surrounding tissues to the embryo sac (Raghavan, 1997). Despite the high functional similarity of mature embryo sacs, their formation is highly diverse across different plant taxa (Figure 2; Maheshwari, 1950; Huang and Russell, 1992; Baroux et al., 2002; Williams and Friedman, 2004). Reproductive development can be divided into two steps: megasporogenesis and megagametogenesis. Megasporogenesis comprises the formation and maturation of the initial meiotic products (megaspores) from a single selected sporophytic cell, the Megaspore Mother Cell (MMC), and is under the control of the usually diploid sporophytic genome. Megagametogenesis describes the following mitotic divisions, cellularization, cell

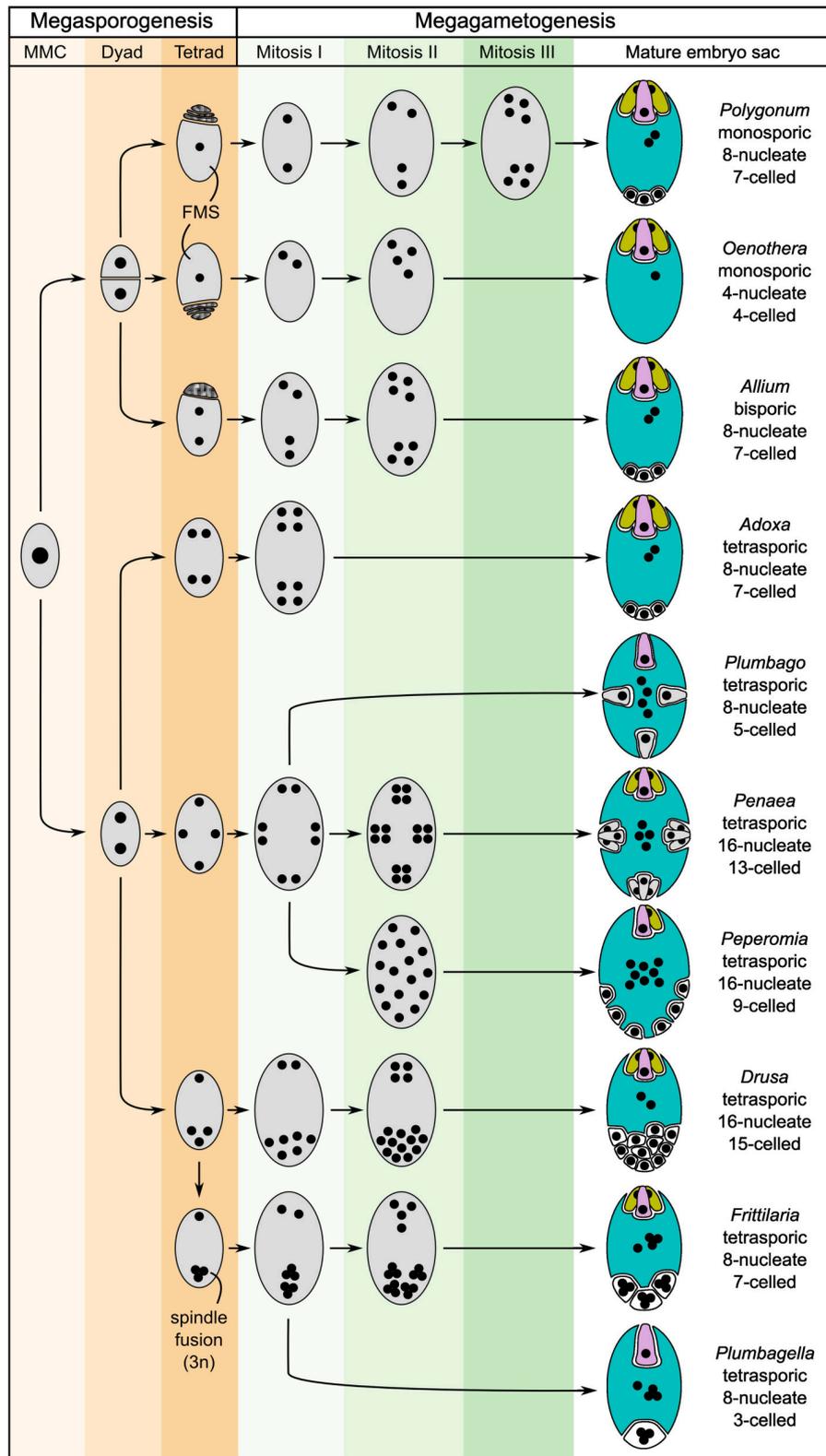


FIGURE 2 | Schematic showing several basic types of female gametophyte development in angiosperms and the structural diversity of the mature embryo sacs (after Maheshwari, 1950). The development of the female gametophyte can be divided into two steps: megasporogenesis (orange shading) and megagametogenesis (green shading). During megasporogenesis, a selected sporophytic cell, the megaspore mother cell (MMC), undergoes meiosis to give rise to

(Continued)

FIGURE 2 | Continued

spores. In most angiosperms, a tetrad of four megaspores is formed, of which three subsequently abort, leaving only one functional megaspore (FMS) to participate in megagametogenesis (e.g., *Polygonum*-type). However, a high diversity of the developmental processes of megasporogenesis and megagametogenesis has been observed in different genera, with variations, for example, including bispority and tetraspority. During megagametogenesis, the mature female gametophyte is formed through mitotic divisions, nuclear migration, and cellularization. For the mature embryo sac, the colors indicate the cell types: egg (pink), synergids (yellow), central cell (blue), and antipodal/lateral cells (white). Cells structurally similar to egg cells or synergids are drawn accordingly, but are colored gray.

specification, and maturation of the female gametophyte, which is under the control of the typically haploid genome. Both processes exhibit high diversity within angiosperms. Depending on the number of spores that survive and participate in megagametogenesis, megasporogenesis can be divided into *monosporic* (one spore), *bisporic* (two spores), and *tetrasporic* (all four spores). Further variation includes the location of the degenerating spores and the positioning of the spores in the *tetrasporic* types. Likewise, megagametogenesis can vary in the number of mitotic divisions, the arrangement of the nuclei/cells, and late divisions of individual cells after cellularization (e.g., in *Amborella*, Friedman, 2006). Comparative analysis of the structure of a wide range of embryo sacs and reconstruction of the ancestral state suggest that the embryo sacs of early angiosperms contained only four cells: two synergids, one egg cell and one central cell. It has been hypothesized that duplication of this four-celled module facilitated the emergence of the bi-nucleate central cell that, following fertilization, forms an endosperm with a maternal:paternal genome contribution ratio of 2:1 (Williams and Friedman, 2004; Friedman, 2006; Friedman and Ryerson, 2009). This unequal parental contribution to the endosperm has received a lot of attention over the last century. As a tissue protecting and nourishing the embryo, the endosperm may be subject to adaptive processes and parental conflicts (Haig and Westoby, 1989; Baroux et al., 2002).

An interesting aspect of female gametophyte development (and *tetrasporic* megasporogenesis) is the formation of a syncytium during the divisions of the nuclei prior to cellularization. In angiosperms, gametogenesis and early stages of endosperm development are the two major examples for the formation of a syncytium. In contrast, the plasmodial tapetum, for example, is formed by degeneration of the cell walls and the fusion of the resulting protoplasts (Furness and Rudall, 1998). Unlike regular cell divisions, where the positions of cells are relatively fixed due to the rigid cell wall, a syncytium allows for nuclear migration and for differentiation according to gradients of positional information. Indeed, determination of cell fate in the embryo sac of *Arabidopsis* depends on the position of the nuclei as, for example, indicated by the *Arabidopsis retinoblastoma-related1 (rbr1)* mutant, which produces supernumerary nuclei differentiating according to their position within the FG (Johnston et al., 2008; Sprunck and Groß-Hardt, 2011). However, the nature of such information is still under debate. Appealing candidates may be gradients of plant hormones, such as cytokinin or auxin. For both, a role in establishing polarity during embryo sac development has been proposed (reviewed in Schmid et al., 2015) but their role may be rather indirect (Lituiev et al., 2013). However, an alternative or complementary hypothesis can be formulated

using the analogy to the syncytial embryogenesis in *Drosophila*, where around 70% of the genes expressed during early embryogenesis show a specific subcellular localization of their mRNA in the syncytium. Interestingly, specific subcellular mRNA localization peaks around the transition from syncytial to cellular development, potentially reflecting the high demand for localization mechanisms (Lécuyer et al., 2007). Thus, a fascinating possibility is that the specific subcellular localization of mRNAs in the syncytial stage of the developing embryo sac may play a role in determining cell fate. A possibility to test this hypothesis would be to separately isolate specific subcellular regions (e.g., the two opposing poles) of the developing syncytial female gametophyte and to compare the transcriptional profiles of these regions with each other.

Another interesting variation of reproductive development is gametophytic apomixis. It refers to the process of asexual reproduction through seeds in the absence of fertilization (reviewed in Koltunow and Grossniklaus, 2003). Apomixis occurs in more than 400 plant species from around 40 genera and is likely of polyphyletic origin (Asker and Jerling, 1992; Carman, 1997). Gametophytic apomixis involves the omission or abortion of meiosis (apomeiosis) and the formation of an embryo from an unfertilized egg (parthenogenesis), while the endosperm can be formed by autonomous development of the central cell or dependent on fertilization (pseudogamy). Depending on the mechanism of the formation of the unreduced megaspore, the resulting offspring can be genetically completely identical to the mother plant without any chromosomal rearrangements. It is thereby possible to fix complex genotypes over multiple generations without a loss in heterozygosity. While gametophytic apomixis is absent in major crop plants, engineered apomictic crops would promise great potential and economical value for plant breeding and agriculture (Koltunow et al., 1995; Vielle-Calzada et al., 1996; Grossniklaus et al., 1998). From a developmental perspective, apomixis can be seen as an alteration of the sexual pathway, where certain processes are initiated too early or in the wrong cell type (Koltunow, 1993; Grossniklaus, 2001). Detailed understanding of the molecular processes and pathways governing gametogenesis during sexual and apomictic reproduction is therefore a precondition to engineer apomixis in crop plants. In evolutionary terms, apomixis is a highly interesting trait. On one hand, it allows the dispersal of seeds without the need for a sexual partner (Smith, 1978) and may therefore be advantageous for the colonization of new habitats (Tomlinson, 1966). On the other hand, the trade-off for this clonal reproduction appears to be very costly. Apomicts may accumulate deleterious mutations over many generations (Muller, 1964) and their populations are likely of low genetic variability, which reduces their potential to adapt to a changing

environment. Recent proposals, however, suggest that epigenetic variation may also contribute to adaptive potential, which may explain the ecological success of many apomicts (Hirsch et al., 2012).

Given the natural variation in sexual and apomictic species, the female gametophyte of angiosperms can be seen as an excellent model system to study fundamental developmental processes and evolutionary aspects of plant development and biology that are of high importance to agriculture. Its simple organization and the relatively few developmental stages would allow for an in-depth analysis of various species enabling evolutionary comparisons at the whole-genome level. Given the high diversity, inter-species comparisons may identify genes and genetic networks involved in the emergence of evolutionary novelties, such as the unequal genetic contribution of the two parents to the endosperm or gametophytic apomixis. Deciphering the evolutionary mechanisms underlying these processes may also provide an answer to the long-standing question, how useful research on model organisms is for crop improvement. However, the small size and inaccessibility of the cell types of developing and mature embryo sacs make the isolation and subsequent application of omics methods very difficult. Aside the challenges associated with data integration and analysis, data generation is hence a major limiting factor. In general, the main obstacle with most approaches is the number of cells required for in-depth profiling of a certain molecule (e.g., protein or metabolite). This may be overcome by either increased sensitivity of the profiling method, or through a simplified collection of a large number of cells. However, most high-throughput isolation methods (e.g., for FACS/FANS/INTACT) rely on the existence of a specific marker (i.e., a cell type-specific promoter) and the possibility to generate transgenic plants. In addition, typically a certain abundance of the cell type of interest in the sample is required for efficient sorting and purification. Given that these preconditions are generally not met by low abundant cell types of non-model organisms, it is likely that plant systems biology will profit the most from an increase in sensitivity and the development of novel profiling methods. In the following sections, we will therefore focus on a subset of omics approaches, which are readily available or which bear great future potential for routine large-scale *in vivo* profiling of specific cell types. The examples given are restricted to studies on specific cell types of the female gametophytes of angiosperms.

3.1. Transcriptome

Transcriptomics is clearly the most frequently used and currently the most robust omics approach to study female gametophyte and plant reproductive development. Following the early transcriptional profiling with low-throughput technologies [early Expressed Sequence Tag (EST) sequencing projects, reviewed in Wuest et al. (2013)], cell type-specific transcriptomes were generated for the egg cell, the central cell, the synergids, and the MMC of *Arabidopsis* (Wuest et al., 2010; Schmid et al., 2011; Schmid et al., 2012), the egg cell and the synergids for rice (Ohnishi et al., 2011; Anderson et al., 2013), all cell types of the mature embryo sac and the Apomictic Initial

Cell (AIC) of *Boechera gunnisoniana* (a close apomictic relative of *Arabidopsis thaliana* where an AIC is specified instead of a sexual MMC, Schmid et al., 2014), and the AIC of *Hieracium praealtum* (hawkweed, where the AIC is formed by an additional sporophytic cell developing adjacent to the sexual reproductive lineage, Okada et al., 2013; Table 2). Given the requirement to establish a specific gene expression profile for cell specification and differentiation, transcriptomics is also especially suitable as a first approach toward an unknown species, because it provides a comprehensive snapshot of the cellular instruction machinery. It further enables the identification of cell type-specific markers and can thus provide a basis for other approaches, like detailed molecular and mechanistic studies. The advantage of transcriptional profiling as compared to proteomic studies is the possibility to amplify the material prior to detection. Several RNA-Seq protocols allow transcriptional profiling of single cells corresponding to as little as about 10 pg of total RNA (reviewed in Head et al., 2014). This low detection limit facilitates the use of relatively low throughput isolation methods, such as LAM or manual microdissection, allowing the profiling of specific cell types of embryo sacs in model and non-model species (Okada et al., 2013; Wuest et al., 2013; Schmid et al., 2014). A current drawback of the amplification strategy is the introduction of potential quantification biases. A possible solution may be Unique Molecular Identifiers (UMI). These are short sequences with random nucleotides (e.g., 1024 different UMIs with 5 random nucleotides), which are used to label initial cDNA molecules prior to amplification. An excess of UMIs compared to the number of identical cDNAs ensures that each combination of a given UMI with a certain cDNA is unique. After amplification and sequencing, this can be used to differentiate between individual molecules in the initial cDNA pool and duplicates originating from cDNA amplification (i.e., to count molecules instead of reads, Islam et al., 2014). An interesting approach for future studies may be Fluorescent *In Situ* RNA SEQUENCING (FISSEQ), in which stably cross-linked cDNA amplicons are sequenced directly within a biological sample, thereby not only quantifying gene expression, but also detecting the subcellular localization of the transcripts (Lee et al., 2014). Improvement of this method and its adaption to plant tissues would thus undoubtedly be a major advance in cell type-specific transcriptional profiling.

3.2. Proteome and Metabolome

Proteomics and metabolomics on specific cell types is substantially more challenging than transcriptomics. A current limitation for cell type-specific proteomics is the large discrepancy between the number of detected proteins compared to the number of expressed genes, which is due to the low sensitivity of proteomics methods towards low-abundant proteins. An additional complexity arises by the presence of a wide range of post-translational modifications, such as phosphorylation or glycosylation. Apart from two early examples, identifying only the major proteins in the egg cells of maize and rice (6 and 4 proteins, Okamoto et al., 2004; Uchiumi et al., 2007), we are only aware of the recent description of the egg cell proteome in rice, where 2138 proteins were identified

using around 500 egg cells (Abiko et al., 2013; **Table 2**). In the same study, 2179 proteins were identified starting from 30,000 isolated sperm cells (**Table 1**; Abiko et al., 2013). Given the further improvements of the sensitivity of mass spectrometers, the example demonstrates that proteomics of purified cells of the female gametophyte should be possible for cases where enough material can be collected. Mechanical or manual isolation of female gametes was reported for a variety of species including barley, wheat, rape seed, maize, tobacco, *Torenia*, *Alstroemeria*, and *Arabidopsis* (Kranz et al., 1991; Holm et al., 1994; Kovács et al., 1994; Katoh et al., 1997; Tian and Russell, 1997; Sprunck et al., 2005; Hoshino et al., 2006; Okuda et al., 2009; Jullien et al., 2012). In most of these species, we anticipate that the protocols would already allow the isolation of sufficient material for MS-based proteomics. Another promising approach for future experiments may be MSI, circumventing the need for (laborious) cell purification.

3.3. Methylome

DNA cytosine methylation (5mC) plays an important role in the epigenetic regulation of plant genomes. While WGBS has not yet been reported for isolated cells of the female gametophyte, bisulfite sequencing of specific sequences has already been applied for *Arabidopsis* central cells and synergids isolated by LAM (Wöhrmann et al., 2012; You et al., 2012). It would likely be possible to combine LAM or manual microdissection with WGBS. This would thus allow methylome profiling of gametes in model as well as non-model species. Importantly, this may provide novel insights into the molecular basis underlying heterosis (Groszmann et al., 2011), characterized by superior characteristics of F1 hybrid plants as compared to their parents. While epigenetic regulatory pathways are likely important for heterosis, their precise involvement remains elusive to date (Chen, 2013). Understanding of the regulatory mechanisms governing heterosis is of great interest for plant breeding and crop production. Importantly, gametophytic development and early stages of embryogenesis are likely important for the establishment of heterosis.

4. CONCLUSION AND PERSPECTIVES

To date, cell type-specific systems biology in plants is frequently constrained by the difficulties associated with the isolation of the cell type of interest in large enough amounts. Robust and simple isolation methods exist only for a few cell types. Consequently, the comprehensive profiling of all cell types of an organism with different large-scale profiling methods, allowing the detailed understanding of all biological processes ongoing in the biological system, is still an unreached goal. While the in-depth understanding of complex organisms over their lifespan is a major aim for systems biology, the use of simple model organisms bears advantages, given the persisting technical limitations. We introduce the female gametophyte of angiosperms as an attractive model system for future systems biology approaches in plant development. Apart from its relatively simple organization, it is of great biological and agronomical importance,

for example with respect to seed production and plant breeding.

Currently, most high-throughput isolation methods with broader application (e.g., FACS/FANS/INTACT) are limited to model organisms (e.g., *Arabidopsis thaliana*, *Oryza sativa*). However, a biological system may be best understood in the context of evolution. In addition, a detailed understanding of the cellular processes in major agriculturally important species including wheat, where an additional challenge is the genome size and its hexaploid nature, are a precondition for targeted crop improvement. Such studies would thus not only be of potential applied value, but would also help to understand the common concepts and divergent mechanisms active in different species. Therefore, methods facilitating large-scale profiling of specific cell types in model as well as non-model organism are of crucial importance. Parallel high-throughput profiling of several organisms covering a phenotypic gradient, or including gain, loss, and alternative usage of modular building blocks along the phylogeny, will enable evolutionary systems biology. This approach may ultimately help to reconstruct the emergence of evolutionary novelties and to find the underlying genetic and molecular networks. Such an understanding would in turn allow the control of the underlying processes with an unprecedented resolution. In perspective, this can be an important precondition for targeted improvement of crop species, including the engineering of apomixis into crop plants.

Even though the isolation of individual cell types is currently still very challenging, the rapid technical advances observed over the past few years in, for example, transcriptional profiling, are clear indications for the tremendous improvement of large-scale profiling technologies. In this light, we emphasize methods for transcriptomics, proteomics, metabolomics, and methylomics, in which we see great future potential. However, cell type-specificity and single-cell resolution are just one step towards a more comprehensive view on developmental processes and environmental responses. Clearly, monitoring subcellular localization of molecules and their interactions will be essential to understand certain patterning processes and specific cellular functions. In analogy to the hypothesized distribution of mRNA within the syncytial female gametophyte, subcellular localization of mRNA may also occur within the cell types of the mature female gametophyte to, for example, target the proteins they encode to a specific subcellular region. In this respect, technologies based on high resolution imaging, allowing large-scale profiling without prior cell isolation, for example MSI or FISSEQ, are very promising for future applications. The growing amount of data and data types also points to the need for novel computational solutions addressing the problems of data storage, integration, and analysis (see Ahrens et al., 2007; Yuan et al., 2008; Fukushima et al., 2009; Chuang et al., 2010; Katari et al., 2010; Liberman et al., 2012; Fukushima et al., 2014; Gomez-Cabrero et al., 2014; Robinson et al., 2014). The current situation, in which data sometimes remain unpublished, are frequently poorly annotated, and widely dispersed in specialized databases, may be taken as motivation to develop integrative computational platforms specifically focussing on future data.

Considering the almost exponential growth of biological data over the last years (Ideker et al., 2001; Chuang et al., 2010), these platforms may also ignore data from the past to allow for innovative solutions. In this context, standardized data formats and annotation, easily accessible databases, powerful data mining tools, user-friendly and freely available software, as well as scalable storage platforms are the current and future demands in systems biology (Chuang et al., 2010; Gomez-Cabrero et al., 2014).

REFERENCES

- Abiko, M., Furuta, K., Yamauchi, Y., Fujita, C., Taoka, M., Isobe, T., et al. (2013). Identification of proteins enriched in rice egg or sperm cells by single-cell proteomics. *PLoS ONE* 8:e69578. doi: 10.1371/journal.pone.0069578
- Ahrens, C. H., Wagner, U., Rehrauer, H. K., Türker, C., and Schlapbach, R. (2007). Current challenges and approaches for the synergistic use of systems biology data in the scientific community. *Experientia Suppl.* 97, 277–307. doi: 10.1007/978-3-7643-7439-6/12
- Anderson, S. N., Johnson, C. S., Jones, D. S., Conrad, L. J., Gou, X., Russell, S. D., et al. (2013). Transcriptomes of isolated *Oryza sativa* gametes characterized by deep sequencing: evidence for distinct sex-dependent chromatin and epigenetic states before fertilization. *Plant J.* 76, 729–741. doi: 10.1111/tpj.12336
- Asker, S. E., and Jerling, L. (1992). *Apomixis in Plants*. London, UK: CRC Press.
- Bailey-Serres, J. (2013). Microgenomics: genome-scale, cell-specific monitoring of multiple gene regulation tiers. *Annu. Rev. Plant Biol.* 64, 293–325. doi: 10.1146/annurev-arplant-050312-120035
- Baroux, C., Spillane, C., and Grossniklaus, U. (2002). Evolutionary origins of the endosperm in flowering plants. *Genome Biol.* 3:reviews1026.1–reviews1026.5. doi: 10.1186/gb-2002-3-9-reviews1026
- Benfey, P. N. (2012). Toward a systems analysis of the root. *Cold Spring Harb. Symp. Quant. Biol.* 77, 91–96. doi: 10.1101/sqb.2012.77.014506
- Birnbaum, K., Shasha, D. E., Wang, J. Y., Jung, J. W., Lambert, G. M., Galbraith, D. W., et al. (2003). A gene expression map of the *Arabidopsis* root. *Science* 302, 1956–1960. doi: 10.1126/science.1090022
- Boone, C. (2014). Yeast systems biology: our best shot at modeling a cell. *Genetics* 198, 435–437. doi: 10.1534/genetics.114.169128
- Borges, F., Gomes, G., Gardner, R., Moreno, N., McCormick, S., Feijó, J. A., et al. (2008). Comparative transcriptomics of *Arabidopsis* sperm cells. *Plant Physiol.* 148, 1168–1181. doi: 10.1104/pp.108.125229
- Bostein, D., and Fink, G. R. (2011). Yeast: an experimental organism for 21st century biology. *Genetics* 189, 695–704. doi: 10.1534/genetics.111.130765
- Brady, S. M., Orlando, D. A., Lee, J. Y., Wang, J. Y., Koch, J., Dinneny, J. R., et al. (2007). A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science* 318, 801–806. doi: 10.1126/science.1146265
- Brechenmacher, L., Lee, J., Sachdev, S., Song, Z., Nguyen, T. H. N., Joshi, T., et al. (2009). Establishment of a protein reference map for soybean root hair cells. *Plant Physiol.* 149, 670–682. doi: 10.1104/pp.108.131649
- Brechenmacher, L., Lei, Z., Libault, M., Findley, S., Sugawara, M. J., Sadowsky, M. J., et al. (2010). Soybean metabolites regulated in root hairs in response to the symbiotic bacterium *Bradyrhizobium japonicum*. *Plant Physiol.* 153, 1808–1822. doi: 10.1104/pp.110.157800
- Brem, R. B., Yvert, G., Clinton, R., and Kruglyak, L. (2002). Genetic dissection of transcriptional regulation in budding yeast. *Science* 296, 752–755. doi: 10.1126/science.1069516
- Brown, J. A., Sherlock, G., Myers, C. L., Burrows, N. M., Deng, C., Wu, H. I., et al. (2006). Global analysis of gene function in yeast by quantitative phenotypic profiling. *Mol. Syst. Biol.* 2, 2006.0001. doi: 10.1038/msb4100043
- Carman, J. G. (1997). Asynchronous expression of duplicate genes in angiosperms may cause apomixis, bispority, tetraspority, and polyembryony. *Biol. J. Linnean Soc.* 61, 51–94. doi: 10.1111/j.1095-8312.1997.tb01778.x
- Chaturvedi, P., Ischebeck, T., Egelhofer, V., Lichtscheidl, I., and Weckwerth, W. (2013). Cell-specific analysis of the tomato pollen proteome from pollen mother cell to mature pollen provides evidence for developmental priming. *J. Proteome Res.* 12, 4892–4903. doi: 10.1021/pr400197p
- Chen, C., Farmer, A. D., Langley, R. J., Mudge, J., Crow, J. A., May, G. D., et al. (2010). Meiosis-specific gene discovery in plants: RNA-Seq applied to isolated *Arabidopsis* male meiocytes. *BMC Plant Biol.* 10:280. doi: 10.1186/1471-2229-10-280
- Chen, J., Lalonde, S., Obrdlik, P., Noorani Vatani, A., Pasra, S. A., Vilarino, C., et al. (2012). Uncovering *Arabidopsis* membrane protein interactome enriched in transporters using mating-based split ubiquitin assays and classification models. *Front. Plant Sci.* 3:124. doi: 10.3389/fpls.2012.00124
- Chen, Z. J. (2013). Genomic and epigenetic insights into the molecular bases of heterosis. *Nat. Rev. Genet.* 14, 471–482. doi: 10.1038/nrg3503
- Chuang, H. Y., Hofree, M., and Ideker, T. (2010). A decade of systems biology. *Annu. Rev. Cell Dev. Biol.* 26, 721–744. doi: 10.1146/annurev-cellbio-100109-104122
- Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E. D., Sevier, C. S., et al. (2010). The genetic landscape of a cell. *Science* 327, 425–431. doi: 10.1126/science.1180823
- Dai, S., and Chen, S. (2012). Single-cell-type proteomics: toward a holistic understanding of plant function. *Mol. Cell. Proteomics* 11, 1622–1630. doi: 10.1074/mcp.R112.021550
- Dai, X., Wang, G., Yang, D. S., Tang, Y., Broun, P., Marks, M. D., et al. (2010). TrichOME: a comparative omics database for plant trichomes. *Plant Physiol.* 152, 44–54. doi: 10.1104/pp.109.145813
- De Los Santos, C., Chang, C. W., Mycek, M. A., and Cardullo, R. A. (2015). FRAP, FLIM, and FRET: detection and analysis of cellular dynamics on a molecular scale using fluorescence microscopy. *Mol. Reprod. Dev.* 82, 587–604. doi: 10.1002/mrd.22501
- Deal, R. B., and Henikoff, S. (2011). The INTACT method for cell type-specific gene expression and chromatin profiling in *Arabidopsis thaliana*. *Nat. Protoc.* 6, 56–68. doi: 10.1038/nprot.2010.175
- Diez, D., Hutchins, A. P., and Miranda-Saavedra, D. (2014). Systematic identification of transcriptional regulatory modules from protein-protein interaction networks. *Nucleic Acids Res.* 42:e6. doi: 10.1093/nar/gkt913
- Dukowic-Szulze, S., Sundararajan, A., Mudge, J., Ramaraj, T., Farmer, A. D., Wang, M., et al. (2014). The transcriptome landscape of early maize meiosis. *BMC Plant Biol.* 14:118. doi: 10.1186/1471-2229-14-118
- Evrard, A., Bargmann, B. O., Birnbaum, K. D., Tester, M., Baumann, U., and Johnson, A. A. (2012). Fluorescence-activated cell sorting for analysis of cell type-specific responses to salinity stress in *Arabidopsis* and rice. *Methods Mol. Biol.* 913, 265–276. doi: 10.1007/978-1-61779-986-0/18
- Fernando, D. D., Quinn, C. R., Brenner, E. D., and Owens, J. N. (2010). Male gametophyte development and evolution in extant gymnosperms. *Int. J. Plant Dev. Biol.* 4, 47–63. Available online at: http://www.globalsciencebooks.info/JournalsSup/10IJPDB_4_SI1.html
- Friedman, W. E. (2006). Embryological evidence for developmental lability during early angiosperm evolution. *Mol. Syst. Biol.* 2, 2006.0001. doi: 10.1038/nature04690
- Friedman, W. E., and Ryerson, K. C. (2009). Reconstructing the ancestral female gametophyte of angiosperms: insights from *Amborella* and other ancient lineages of flowering plants. *Am. J. Bot.* 96, 129–143. doi: 10.3732/ajb.0800311

ACKNOWLEDGMENTS

Work on gametophyte development, apomixis, and epigenetic gene regulation in UG's laboratory is supported by the University of Zürich, and by grant C11.0060 from the “Staatssekretariat für Bildung und Forschung” in the framework of COST action FA0903 (to UG and AS), grant 31003A_141245 from the Swiss National Science Foundation (to UG), and the grant No 250358 (MEDEA) from the European Research Council (to UG).

- Fukushima, A., Kanaya, S., and Nishida, K. (2014). Integrated network analysis and effective tools in plant systems biology. *Front. Plant Sci.* 5:598. doi: 10.3389/fpls.2014.00598
- Fukushima, A., Kusano, M., Redestig, H., Arita, M., and Saito, K. (2009). Integrated omics approaches in plant systems biology. *Curr. Opin. Chem. Biol.* 13, 532–538. doi: 10.1016/j.cbpa.2009.09.022
- Furness, C. A., and Rudall, P. J. (1998). The tapetum and systematics in monocotyledons. *Bot. Rev.* 64, 201–239. doi: 10.1007/BF02856565
- Gomez-Cabrero, D., Abugessaisa, I., Maier, D., Teschendorff, A., Merkschlagler, M., Gisel, A., et al. (2014). Data integration in the era of omics: current and future challenges. *BMC Syst. Biol.* 8:I1. doi: 10.1186/1752-0509-8-S2-I1
- Gong, W., He, K., Covington, M., Dinesh-Kumar, S. P., Snyder, M., Harmer, S. L., et al. (2008). The development of protein microarrays and their applications in DNA-protein and protein-protein interaction analyses of *Arabidopsis* transcription factors. *Mol. Plant* 1, 27–41. doi: 10.1093/mp/ssm009
- Gou, X., Yuan, T., Wei, X., and Russell, S. D. (2009). Gene expression in the dimorphic sperm cells of *Plumbago zeylanica*: transcript profiling, diversity, and relationship to cell type. *Plant J.* 60, 33–47. doi: 10.1111/j.1365-313X.2009.03934.x
- Greil, F., Moorman, C., and Van Steensel, B. (2006). DamID: mapping of *in vivo* protein-genome interactions using tethered DNA adenine methyltransferase. *Methods Enzymol.* 410, 342–359. doi: 10.1016/S0076-6879(06)10016-6
- Grobei, M. A., Qeli, E., Brunner, E., Rehrauer, H., Zhang, R., Roschitzki, B., et al. (2009). Deterministic protein inference for shotgun proteomics data provides new insights into *Arabidopsis* pollen development and function. *Genome Res.* 19, 1786–1800. doi: 10.1101/gr.089060.108
- Grossniklaus, U. (2001). “From sexuality to apomeiosis: molecular and genetic approaches,” in *The Flowering of Apomixis: from Mechanisms to Genetic Engineering*, eds Y. Savidan, J. G. Carman, and T. Dresselhaus (Mexico, DF: CIMMYT, IRD, European Commission DG VI (FAIR)), 168–211.
- Grossniklaus, U., Koltunow, A., and van Lookeren Campagne, M. (1998). A bright future for apomixis. *Trends Plant Sci.* 3, 415–416. doi: 10.1016/S1360-1385(98)01338-7
- Groszmann, M., Greaves, I. K., Albertyn, Z. I., Scofield, G. N., Peacock, W. J., and Dennis, E. S. (2011). Changes in 24-nt siRNA levels in *Arabidopsis* hybrids suggest and epigenetic contribution to hybrid vigor. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2617–2622. doi: 10.1073/pnas.1019217108
- Haig, D., and Westoby, M. (1989). Parent-specific gene expression and the triploid endosperm. *Am. Nat.* 134, 147–155. doi: 10.1086/284971
- He, G., Elling, A. A., and Deng, X. W. (2011). The epigenome and plant development. *Annu. Rev. Plant Biol.* 62, 411–435. doi: 10.1146/annurev-arplant-042110-103806
- Head, S. R., Komori, H. K., LaMere, S. A., Whisenant, T., Van Nieuwerburgh, F., Salomon, D. R., et al. (2014). Library construction for next-generation sequencing: overviews and challenges. *Biotechniques* 56, 61–77. doi: 10.2144/000114133
- Hirsch, S., Baumberger, R., and Grossniklaus, U. (2012). Epigenetic variation, inheritance, and selection in plant populations. *Cold Spring Harb. Symp. Quant. Biol.* 77, 97–104. doi: 10.1101/sqb.2013.77.014605
- Hollenbeck, T. P., Siuzdak, G., and Blackledge, R. D. (1999). Electrospray and MALDI mass spectrometry in the identification of spermidics in criminal investigations. *J. Forensic Sci.* 44, 793–788. doi: 10.1520/JFS14553J
- Holm, P. B., Knudsen, S., Mouritzen, P., Negri, D., Olsen, F. L., and Roue, C. (1994). Regeneration of fertile barley plants from mechanically isolated protoplasts of the fertilized egg cell. *Plant Cell* 6, 531–543. doi: 10.1105/tpc.6.4.531
- Hony, D., and Twell, D. (2004). Transcriptome analysis of haploid male gametophyte development in *Arabidopsis*. *Genome Biol.* 5:R85. doi: 10.1186/gb-2004-5-11-r85
- Horn, P. J., Korte, A. R., Neogi, P. B., Love, E., Fuchs, J., Strupat, K., et al. (2012). Spatial mapping of lipids at cellular resolution in embryos of cotton. *Plant Cell* 24, 622–636. doi: 10.1105/tpc.111.094581
- Hoshino, Y., Murata, N., and Shinoda, K. (2006). Isolation of individual egg cells and zygotes in *Alstroemeria* followed by manual selection with a microcapillary-connected micropump. *Ann. Bot.* 97, 1139–1144. doi: 10.1093/aob/mcl072
- Hu, T. X., Yu, M., and Zhao, J. (2011). Techniques of cell type-specific transcriptome analysis and applications in researches of plant sexual reproduction. *Front. Biol.* 6, 31–39. doi: 10.1007/s11515-011-1090-1
- Huang, B. Q., and Russell, S. D. (1992). Female germ unit: organization, isolation, and function. *Int. Rev. Cytol.* 140, 233–293. doi: 10.1016/S0074-7696(08)61099-2
- Ideker, T., Galitski, T., and Hood, L. (2001). A new approach to decoding life: systems biology. *Annu. Rev. Genomics Hum. Genet.* 2, 343–372. doi: 10.1146/annurev.genom.2.1.343
- Ischebeck, T., Valledor, L., Lyon, D., Gingl, S., Nagler, M., Meijón, M., et al. (2014). Comprehensive cell-specific protein analysis in early and late pollen development from diploid microsporocytes to pollen tube growth. *Mol. Cell. Proteomics* 13, 295–310. doi: 10.1074/mcp.M113.028100
- Ishida, T., Kurata, T., Okada, K., and Wada, T. (2008). A genetic regulatory network in the development of trichomes and root hairs. *Annu. Rev. Plant Biol.* 59, 365–386. doi: 10.1146/annurev.arplant.59.032607.092949
- Islam, S., Zeisel, A., Joost, S., La Manno, G., Zajac, P., Kasper, M., et al. (2014). Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat. Methods* 11, 163–166. doi: 10.1038/nmeth.2772
- James, P. (1997). Protein identification in the post-genome era: the rapid rise of proteomics. *Q. Rev. Biophys.* 30, 279–331. doi: 10.1017/S0033583597003399
- Ji, L., Neumann, D. A., and Schmitz, R. J. (2015). Crop epigenomics: identifying, unlocking, and harnessing cryptic variation in crop genomes. *Mol. Plant* 8, 860–870. doi: 10.1016/j.molp.2015.01.021
- Johnston, A. J., Matveeva, E., Kirioukhova, O., Grossniklaus, U., and Grissem, W. (2008). A dynamic reciprocal RBR-PRC2 regulatory circuit controls *Arabidopsis* gametophyte development. *Curr. Biol.* 18, 1680–1686. doi: 10.1016/j.cub.2008.09.026
- Jullien, P. E., Susaki, D., Yelagandula, R., Higashiyama, T., and Berger, F. (2012). DNA methylation dynamics during sexual reproduction in *Arabidopsis thaliana*. *Curr. Biol.* 22, 1825–1830. doi: 10.1016/j.cub.2012.07.061
- Kaspar, S., Peukert, M., Svatos, A., Matros, A., and Mock, H. P. (2011). MALDI-imaging mass spectrometry - an emerging technique in plant biology. *Proteomics* 11, 1840–1850. doi: 10.1002/pmic.201000756
- Katari, M. S., Nowicki, S. D., Aceituno, F. F., Nero, D., Kelfer, J., Thompson, L. P., et al. (2010). Virtual Plant: a software platform to support systems biology research. *Plant Physiol.* 152, 500–515. doi: 10.1104/pp.109.147025
- Katoh, N., Lörz, H., and Kranz, E. (1997). Isolation of viable egg cells of rape (*Brassica napus* L.). *Zygote* 5, 31–33. doi: 10.1017/S0967199400003531
- Kehr, J. (2001). High resolution spation analysis of plant systems. *Curr. Opin. Plant Biol.* 4, 197–201. doi: 10.1016/S1369-5266(00)00161-8
- Kelliher, T., and Walbot, V. (2012). Hypoxia triggers meiotic fate acquisition in maize. *Science* 337, 345–348. doi: 10.1126/science.1220080
- Kelliher, T., and Walbot, V. (2014). Maize germinal cell initials accommodate hypoxia and precociously express meiotic genes. *Plant J.* 77, 639–652. doi: 10.1111/tpj.12414
- Kerk, N. M., Ceserani, T., Tausta, S. L., Sussex, I. M., and Nelson, T. M. (2003). Laser capture microdissection of cells from plant tissues. *Plant Physiol.* 132, 27–35. doi: 10.1104/pp.102.018127
- Kitano, H. (2002). Systems biology: a brief overview. *Science* 295, 1662–1664. doi: 10.1126/science.1069492
- Koltunow, A. M. (1993). Apomixis: embryo sacs and embryos formed without meiosis or fertilization in ovules. *Plant Cell* 5, 1425–1437. doi: 10.1105/tpc.5.10.1425
- Koltunow, A. M., Bicknell, R. A., and Chaudhury, A. M. (1995). Apomixis: molecular strategies for the generation of genetically identical seeds without fertilization. *Plant Physiol.* 108, 1345–1352.
- Koltunow, A. M., and Grossniklaus, U. (2003). Apomixis: a developmental perspective. *Annu. Rev. Plant Biol.* 54, 547–574. doi: 10.1146/annurev.arplant.54.110901.160842
- Konstantinidis, K. T., Serres, M. H., Romine, M. F., Rodrigues, J. L. M., Auchtung, J., McCue, L. A., et al. (2009). Comparative systems biology across an evolutionary gradient within the *Shewanella* genus. *Proc. Natl. Acad. Sci. U.S.A.* 106, 15909–15914. doi: 10.1073/pnas.0902000106
- Kovács, M., Barnabás, B., and Kranz, E. (1994). The isolation of viable egg cells of wheat (*Triticum aestivum* L.). *Sex. Plant Reprod.* 7, 311–312. doi: 10.1007/BF00227715
- Kranz, E., Bautor, J., and Lörz, H. (1991). *In vitro* fertilization of single, isolated gametes of maize mediated by electrofusion. *Sex. Plant Reprod.* 4, 12–16. doi: 10.1007/BF00194565

- Kueger, S., Steinhauser, D., Willmitzer, L., and Giavalisco, P. (2012). High-resolution plant metabolomics: from mass spectral features to metabolites and from whole-cell analysis to subcellular metabolite distributions. *Plant J.* 70, 39–50. doi: 10.1111/j.1365-313X.2012.04902.x
- Lau, S., Slane, D., Herud, O., Kong, J., and Jürgens, G. (2012). Early embryogenesis in flowering plants: setting up the basic body pattern. *Annu. Rev. Plant Biol.* 63, 483–506. doi: 10.1146/annurev-arplant-042811-105507
- Lécuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T., et al. (2007). Global analyses of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* 131, 174–187. doi: 10.1016/j.cell.2007.08.003
- Lee, J. H., Daugharthy, E. R., Scheiman, J., Kalhor, R., Yang, J. L., Ferrante, T. C., et al. (2014). Highly multiplexed subcellular RNA sequencing *in situ*. *Science* 343, 1360–1363. doi: 10.1126/science.1250212
- Lee, Y. J., Perdian, D. C., Song, Z., Yeung, E. S., and Nikolau, B. J. (2012). Use of mass spectrometry for imaging metabolites in plants. *Plant J.* 70, 81–95. doi: 10.1111/j.1365-313X.2012.04899.x
- Libault, M., Brechenmacher, L., Cheng, J., Xu, D., and Stacey, G. (2010a). Root hair systems biology. *Trends Plant Sci.* 15, 641–650. doi: 10.1016/j.tplants.2010.08.010
- Libault, M., Farmer, A., Brechenmacher, L., Drnevich, J., Langley, R. J., Bilgin, D. D., et al. (2010b). Complete transcriptome of the soybean root hair cell, a single-cell model, and its alteration in response to *Bradyrhizobium japonicum* infection. *Plant Physiol.* 152, 541–552. doi: 10.1104/pp.109.148379
- Libeau, P., Durand, M., Granier, F., Marquis, C., Berthomé, R., Renou, J. P., et al. (2011). Gene expression profiling of *Arabidopsis* meiocytes. *Plant Biol.* 13, 784–793. doi: 10.1111/j.1438-8677.2010.00435.x
- Lieberman, L. A., Sozzani, R., and Benfey, P. N. (2012). Integrative systems biology: an attempt to describe a simple weed. *Curr. Opin. Plant Biol.* 15, 162–167. doi: 10.1016/j.pbi.2012.01.004
- Lituiev, D. S., Krohn, N. G., Müller, B., Jackson, D., Hellriegel, B., Dresselhaus, T., et al. (2013). Theoretical and experimental evidence indicates that there is no detectable auxin gradient in the angiosperm female gametophyte. *Development* 140, 4544–4553. doi: 10.1242/dev.098301
- Luo, S. D., Shi, G. W., and Baker, B. S. (2011). Direct targets of the *D. melanogaster* DSX^F protein and the evolution of sexual development. *Development* 138, 2761–2771. doi: 10.1242/dev.065227
- Maheshwari, P. (1950). *An Introduction to the Embryology of Angiosperms*. New York, NY: McGraw-Hill.
- Mardis, E. R. (2013). Next-generation sequencing platforms. *Annu. Rev. Anal. Chem.* 6, 287–303. doi: 10.1146/annurev-anchem-062012-092628
- Miura, F., Enomoto, Y., Dairiki, R., and Ito, T. (2012). Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. *Nucleic Acids Res.* 40:e136. doi: 10.1093/nar/gks454
- Moussaieff, A., Rogachev, I., Brodsky, L., Malitsky, S., Toal, T. W., Belcher, H., et al. (2013). High-resolution metabolic mapping of cell types in plant roots. *Proc. Natl. Acad. Sci. U.S.A.* 110, E1232–E1241. doi: 10.1073/pnas.1302019110
- Muller, H. J. (1964). The relation of recombination of mutational advance. *Mutat. Res.* 106, 2–9. doi: 10.1016/0027-5107(64)90047-8
- Nestler, J., Schütz, W., and Hochholdinger, F. (2011). Conserved and unique features of the maize (*Zea mays* L.) root hair proteome. *J. Proteome Res.* 10, 2525–2537. doi: 10.1021/pr200003k
- Obrdlik, P., El-Bakkoury, M., Hamacher, T., Cappellaro, C., Vilarino, C., Fleischer, C., et al. (2004). K⁺ channel interactions detected by a genetic system optimized for systematic studies of membrane protein interactions. *Proc. Natl. Acad. Sci. U.S.A.* 101, 12242–12247. doi: 10.1073/pnas.0404467101
- Ohnishi, T., Takahashi, H., Mogi, M., Takahashi, H., Kikuchi, S., Yano, K., et al. (2011). Distinct gene expression profiles in egg and synergid cells of rice as revealed by cell type-specific microarrays. *Plant Physiol.* 155, 881–891. doi: 10.1104/pp.110.167502
- Okada, T., Hu, Y., Tucker, M. R., Taylor, J. M., Johnson, S. D., Spriggs, A., et al. (2013). Enlarging cells initiating apomixis in *Hieracium praealtum* transition to an embryo sac program prior to entering mitosis. *Plant Physiol.* 163, 216–231. doi: 10.1104/pp.113.219485
- Okada, T., Singh, M. B., and Bhalla, P. L. (2007). Transcriptome profiling of *Lilium longiflorum* generative cells by cDNA microarray. *Plant Cell Rep.* 26, 1045–1052. doi: 10.1007/s00299-006-0300-9
- Okamoto, T., Higuchi, K., Shinkawa, T., Isobe, T., Lörz, H., Koshiba, T., et al. (2004). Identification of major proteins in maize egg cells. *Plant Cell Physiol.* 45, 1406–1412. doi: 10.1093/pcp/pch161
- Okuda, S., Tsutsui, H., Shiina, K., Sprunck, S., Takeuchi, H., Yui, R., et al. (2009). Defensin-like polypeptide LUREs are pollen tube attractants secreted from synergid cells. *Nature* 458, 357–361. doi: 10.1038/nature07882
- Okumoto, S. (2012). Quantitative imaging using genetically encoded sensors for small molecules in plants. *Plant J.* 70, 108–117. doi: 10.1111/j.1365-313X.2012.04910.x
- Okumoto, S., Jones, A., and Frommer, W. B. (2012). Quantitative imaging with fluorescent biosensors. *Annu. Rev. Plant Biol.* 63, 663–706. doi: 10.1146/annurev-arplant-042110-103745
- Orlando, V., and Paro, R. (1993). Mapping *Polycomb*-repressed domains in the *bithorax* complex using *in vivo* formaldehyde cross-linked chromatin. *Cell* 75, 1187–1198. doi: 10.1016/0092-8674(93)90328-N
- Österlund, T., Nookaew, I., and Nielsen, J. (2012). Fifteen years of large scale metabolic modeling of yeast: developments and impacts. *Biotechnol. Adv.* 30, 979–988. doi: 10.1016/j.biotechadv.2011.07.021
- Palovaara, J., Saiga, S., and Weijers, D. (2013). Transcriptomics approaches in the early *Arabidopsis* embryo. *Trends Plant Sci.* 18, 514–521. doi: 10.1016/j.tplants.2013.04.011
- Pelkmans, L. (2012). Using cell-to-cell variability – a new era in molecular biology. *Science* 336, 425–426. doi: 10.1126/science.1222161
- Perry, R. H., Cooks, R. G., and Noll, R. J. (2008). Orbitrap mass spectrometry: instrumentation, ion motion and applications. *Mass Spectrom. Rev.* 27, 661–699. doi: 10.1002/mas.20186
- Petricka, J. J., Schauer, M. A., Megraw, M., Breakfield, N. W., Thompson, J. W., Georgiev, S., et al. (2012). The protein expression landscape of the *Arabidopsis* root. *Proc. Natl. Acad. Sci. U.S.A.* 109, 6811–6818. doi: 10.1073/pnas.1202546109
- Qiao, Z., and Libault, M. (2013). Unleashing the potential of the root hair cell as a single plant cell type model in root systems biology. *Front. Plant Sci.* 4:484. doi: 10.3389/fpls.2013.00484
- Raghavan, V. (1997). *Molecular Embryology of Flowering Plants*. Cambridge, UK: Cambridge University Press.
- Robinson, S. W., Fernandes, M., and Husi, H. (2014). Current advances in systems and integrative biology. *Comput. Struct. Biotechnol. J.* 11, 35–46. doi: 10.1016/j.csbj.2014.08.007
- Rogers, E. D., Jackson, T., Moussaieff, A., Aharoni, A., and Benfey, P. N. (2012). Cell type-specific transcriptional profiling: implications for metabolite profiling. *Plant J.* 70, 5–17. doi: 10.1111/j.1365-313X.2012.04888.x
- Sakata, K., and Komatsu, S. (2014). Plant proteomics: from genome sequencing to proteome databases and repositories. *Methods Mol. Biol.* 1072, 29–42. doi: 10.1007/978-1-62703-631-3/3
- Schauer, T., Schwalie, P. C., Handley, A., Margulies, C. E., Flicek, P., and Ladurner, A. G. (2013). CAST-ChIP maps cell-type-specific chromatin states in the *Drosophila* central nervous system. *Cell Rep.* 5, 271–282. doi: 10.1016/j.celrep.2013.09.001
- Schillmiller, A. L., Miner, D. P., Larson, M., McDowell, E., Gang, D. R., Wilkerson, C., et al. (2010). Studies of a biochemical factory: tomato trichome deep expressed sequence tag sequencing and proteomics. *Plant Physiol.* 153, 1212–1223. doi: 10.1104/pp.110.157214
- Schmid, M., Davison, T. S., Henz, S. R., Pape, U. J., Demar, M., Vingron, M., et al. (2005). A gene expression map of *Arabidopsis thaliana* development. *Nat. Genet.* 37, 501–506. doi: 10.1038/ng1543
- Schmid, M. W., and Grossniklaus, U. (2015). Rcount: simple and flexible RNA-Seq read counting. *Bioinformatics* 31, 436–437. doi: 10.1093/bioinformatics/btu680
- Schmid, M. W., Schmidt, A., Klostermeier, U. C., Barann, M., Rosenstiel, P., and Grossniklaus, U. (2012). A powerful method for transcriptional profiling of specific cell types in eukaryotes: laser-assisted microdissection and RNA sequencing. *PLoS ONE* 7:e29685. doi: 10.1371/journal.pone.0029685
- Schmidt, A., Schmid, M. W., and Grossniklaus, U. (2012). Analysis of plant germline development by high-throughput RNA profiling: technical advances and new insights. *Plant J.* 70, 18–29. doi: 10.1111/j.1365-313X.2012.04897.x
- Schmidt, A., Schmid, M. W., and Grossniklaus, U. (2015). Plant germline formation: molecular insights define common concepts and illustrate developmental flexibility in apomictic and sexual reproduction. *Development* 142, 229–241. doi: 10.1242/dev.102103

- Schmidt, A., Schmid, M. W., Klostermeier, U. C., Qi, W., Guthörl, D., Sailer, C., et al. (2014). Apomictic and sexual germline development differ with respect to cell cycle, transcriptional, hormonal and epigenetic regulation. *PLoS Genet.* 10:e1004476. doi: 10.1371/journal.pgen.1004476
- Schmidt, A., Wuest, S. E., Vijverberg, K., Baroux, C., Kleen, D., and Grossniklaus, U. (2011). Transcriptome analysis of the *Arabidopsis* megaspore mother cell uncovers the importance of RNA helicases for plant germ line development. *PLoS Biol.* 9:e1001155. doi: 10.1371/journal.pbio.1001155
- Schulze, W. X., and Usadel, B. (2010). Quantitation in mass-spectrometry-based proteomics. *Annu. Rev. Plant Biol.* 61, 491–516. doi: 10.1146/annurev-arplant-042809-112132
- Sehgal, A., Mann, N., and Mohan Ram, H. Y. (2014). Structural and developmental variability in the female gametophyte of *Griffithella hookeriana*, *Polypleurum stylosum*, and *Zeylanidium lichenoides* and its bearing on the occurrence of single fertilization in Podostemaceae. *Plant Reprod.* 27, 205–223. doi: 10.1007/s00497-014-0252-0
- Sheth, B. P., and Thaker, V. S. (2014). Plant systems biology: insights, advances and challenges. *Planta* 240, 33–54. doi: 10.1007/s00425-014-2059-5
- Slane, D., Kong, J., Berendzen, K. W., Kilian, J., Henschen, A., Kolb, M., et al. (2014). Cell type-specific transcriptome analysis in the early *Arabidopsis thaliana* embryo. *Development* 141, 4831–4840. doi: 10.1242/dev.116459
- Smith, J. M. (1978). *The Evolution of Sex*. Cambridge, UK: Cambridge University Press.
- Southall, T. D., Gold, K. S., Egger, B., Davidson, C. M., Caygill, E. E., Marshall, O. J., et al. (2013). Cell-type-specific profiling of gene expression and chromatin binding without cell isolation: assaying RNA pol II occupancy in neural stem cells. *Dev. Cell* 26, 101–112. doi: 10.1016/j.devcel.2013.05.020
- Soyer, O. S. (2012). *Evolutionary Systems Biology, Advances in Experimental Medicine and Biology*, Vol. 751. Heidelberg: Springer.
- Sprunck, S., Baumann, U., Edwards, K., Langridge, P., and Dresselhaus, T. (2005). The transcript composition of egg cells changes significantly following fertilization in wheat (*Triticum aestivum* L.). *Plant J.* 41, 660–672. doi: 10.1111/j.1365-3113X.2005.02332.x
- Sprunck, S., and Groß-Hardt, R. (2011). Nuclear behavior, cell polarity, and cell specification in the female gametophyte. *Sex. Plant Reprod.* 24, 123–136. doi: 10.1007/s00497-011-0161-4
- Su, Z., Han, L., and Zhao, Z. (2011). Conservation and divergence of DNA methylation in eukaryotes: new insights from single base-resolution DNA methylomes. *Epigenetics* 6, 134–140. doi: 10.4161/epi.6.2.13875
- Süel, G. M., Kulkarni, R. P., Dworkin, J., Garcia-Ojalvo, J., and Elowitz, M. B. (2007). Tunability and noise dependence in differentiation dynamics. *Science* 315, 1716–1719. doi: 10.1126/science.1137455
- Tang, X., Zhang, Z. Y., Zhang, W. J., Zhao, X. M., Li, X., Zhang, D., et al. (2010). Global gene profiling of laser-captured pollen mother cells indicates molecular pathways and gene subfamilies involved in rice meiosis. *Plant Physiol.* 154, 1855–1870. doi: 10.1104/pp.110.161661
- Taylor-Teeple, M., Ron, M., and Brady, S. M. (2011). Novel biological insights revealed from cell type-specific expression profiling. *Curr. Opin. Plant Biol.* 14, 1–7. doi: 10.1016/j.pbi.2011.05.007
- Tian, H. Q., and Russell, S. D. (1997). Micromanipulation of male and female gametes of *Nicotiana tabacum*: I. isolation of gametes. *Plant Cell Rep.* 16, 555–560. doi: 10.1007/s002990050278
- Tissier, A. (2012). Glandular trichomes: what comes after expressed sequence tags? *Plant J.* 70, 51–68. doi: 10.1111/j.1365-3113X.2012.04913.x
- Tomlinson, J. (1966). The advantages of hermaphroditism and parthenogenesis. *J. Theor. Biol.* 11, 54–58. doi: 10.1016/0022-5193(66)90038-5
- Tucker, M. R., Okada, T., Hu, Y., Scholefield, A., Taylor, J. M., and Koltunow, A. M. (2012). Somatic small RNA pathways promote the mitotic events of megagametogenesis during female reproductive development in *Arabidopsis*. *Development* 139, 1399–1404. doi: 10.1242/dev.075390
- Twell, D. (2011). Male gametogenesis and germline specification in flowering plants. *Sex. Plant Reprod.* 24, 149–160. doi: 10.1007/s00497-010-0157-5
- Uchiumi, T., Shinkawa, T., Isobe, T., and Okamoto, T. (2007). Identification of the major protein components of rice egg cells. *J. Plant Res.* 120, 575–579. doi: 10.1007/s10265-007-0095-y
- Uzoma, I., and Zhu, H. (2013). Interactome mapping: using protein microarray technology to reconstruct diverse protein networks. *Genomics Proteomics Bioinform.* 11, 18–28. doi: 10.1016/j.gpb.2012.12.005
- Van Cutsem, E., Simonart, G., Degand, H., Faber, A. M., Morsomme, P., and Boutry, M. (2011). Gel-based and gel-free proteomic analysis of *Nicotiana tabacum* trichomes identifies proteins involved in secondary metabolism and in the (a)biotic stress response. *Proteomics* 11, 440–454. doi: 10.1002/pmic.201000356
- van Steensel, B., and Henikoff, S. (2000). Identification of *in vivo* DNA targets of chromatin proteins using tethered Dam methyltransferase. *Nat. Biotechnol.* 18, 424–428. doi: 10.1038/74487
- Veličković, D., Ropartz, D., Guillon, F., Saulnier, L., and Rogniaux, H. (2014). New insights into the structural and spatial variability of cell-wall polysaccharides during wheat grain development, as revealed through MALDI mass spectrometry imaging. *J. Exp. Bot.* 65, 2079–2091. doi: 10.1093/jxb/eru065
- Vielle-Calzada, J. P., Crane, C., and Stelly, D. M. (1996). Apomixis – the asexual revolution. *Science* 274, 1322–1323. doi: 10.1126/science.274.52.91.1322
- Vogler, H., Draeger, C., Weber, A., Felekis, D., Eichenberger, C., Routier-Kierzkowsky, A. L., et al. (2013). The pollen tube: a soft shell with a hard core. *Plant J.* 73, 617–627. doi: 10.1111/tpj.12061
- Weckwerth, W. (2011). Green systems biology – from single genomes, proteomes and metabolomes to ecosystems research and biotechnology. *J. Proteomics* 75, 284–305. doi: 10.1016/j.jprot.2011.07.010
- Wei, L. Q., Xu, W. Y., Deng, Z. Y., Su, Z., Xue, Y., and Wang, T. (2010). Genome-scale analysis and comparison of gene expression profiles in developing and germinated pollen in *Oryza sativa*. *BMC Genomics* 11:338. doi: 10.1186/1471-2164-11-338
- Weinhofer, I., Hehenberger, E., Roszak, P., Hennig, L., and Köhler, C. (2010). H3K27me3 profiling of the endosperm implies exclusion of *Polycomb* group protein targeting by DNA methylation. *PLoS Genet.* 6:e1001152. doi: 10.1371/journal.pgen.1001152
- Weinhofer, I., and Köhler, C. (2014). Endosperm-specific chromatin profiling by fluorescence-activated nuclei sorting and ChIP-on-chip. *Methods Mol. Biol.* 1112, 105–115. doi: 10.1007/978-1-62703-773-0/7
- Williams, J. H., and Friedman, W. E. (2004). The four-celled female gametophyte of *Illicium* (Illiciaceae; Austrobaileyales): implications for understanding the origin and early evolution of monocots, eumangoliids, and eudicots. *Am. J. Bot.* 91, 332–351. doi: 10.3732/ajb.91.3.332
- Wöhrmann, H. J. P., Gagliardini, V., Raissig, M. T., Wehrle, W., Arand, J., Schmidt, A., et al. (2012). Identification of a DNA methylation-independent imprinting control region at the *Arabidopsis* *MEDEA* locus. *Genes Dev.* 26, 1837–1850. doi: 10.1101/gad.195123.112
- Wuest, S. E., and Grossniklaus, U. (2014). Laser-assisted microdissection applied to floral tissues. *Methods Mol. Biol.* 1110, 329–344. doi: 10.1007/978-1-4614-9408-9/19
- Wuest, S. E., Schmid, M. W., and Grossniklaus, U. (2013). Cell-specific expression profiling of rare cell types as exemplified by its impact on our understanding of female gametophyte development. *Curr. Opin. Plant Biol.* 16, 41–49. doi: 10.1016/j.pbi.2012.12.001
- Wuest, S. E., Vijverberg, K., Schmidt, A., Weiss, M., Gheyselincx, J., Lohr, M., et al. (2010). *Arabidopsis* female gametophyte gene expression map reveals similarities between plant and animal gametes. *Curr. Biol.* 20, 506–512. doi: 10.1016/j.cub.2010.01.051
- Yang, H., Lu, P., Wang, Y., and Ma, H. (2011a). The transcriptome landscape of *Arabidopsis* male meiocytes from high-throughput sequencing: the complexity and evolution of the meiotic process. *Plant J.* 65, 503–516. doi: 10.1111/j.1365-3113X.2010.04439.x
- Yang, L., Guo, S., Li, Y., Zhou, S., and Tao, S. (2011b). Protein microarrays for systems biology. *Acta Biochim. Biophys. Sin.* 43, 161–171. doi: 10.1093/abbs/gmq127
- Yang, W. C., Shi, D. Q., and Chen, Y. H. (2010). Female gametophyte development in flowering plants. *Annu. Rev. Plant Biol.* 61, 89–108. doi: 10.1146/annurev-arplant-042809-112203

- You, W., Tyczewska, A., Spencer, M., Daxinger, L., Schmid, M. W., Grossniklaus, U., et al. (2012). Atypical DNA methylation of genes encoding cysteine-rich peptides in *Arabidopsis thaliana*. *BMC Plant Biol.* 12:51. doi: 10.1186/1471-2229-12-51
- Yuan, J. S., Galbraith, D. W., Dai, S. Y., Griffin, P., and Stewart, C. N. (2008). Plant systems biology comes of age. *Trends Plant Sci.* 13, 165–171. doi: 10.1016/j.tplants.2008.02.003
- Zhan, J., Thakare, D., Ma, C., Lloyd, A., Nixon, N. M., Arakaki, A. M., et al. (2015). RNA sequencing of laser-capture microdissected compartments of the maize kernel identifies regulatory modules associated with endosperm cell differentiation. *Plant Cell* 27, 513–531. doi: 10.1105/tpc.114.135657
- Zhang, Y., Gao, P., and Yuan, J. S. (2010). Plant protein-protein interaction network and interactome. *Curr. Genomics* 11, 40–46. doi: 10.2174/138920210790218016
- Zhao, X., Yang, N., and Wang, T. (2013). Comparative proteomic analysis of generative and sperm cells reveals molecular characteristics associated with sperm development and function specialization. *J. Proteome Res.* 12, 5058–5071. doi: 10.1021/pr400291p
- Zhu, M., Dai, S., McClung, S., Yan, X., and Chen, S. (2009). Functional differentiation of *Brassica napus* guard cells and mesophyll cells revealed by comparative proteomics. *Mol. Cell. Proteomics* 8, 752–766. doi: 10.1074/mcp.M800343-MCP200

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Schmid, Schmidt and Grossniklaus. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Methods to isolate a large amount of generative cells, sperm cells and vegetative nuclei from tomato pollen for “omics” analysis

Yunlong Lu^{1,2}, Liqin Wei¹ and Tai Wang^{1*}

¹ Key Laboratory of Plant Molecular Physiology, Institute of Botany, Chinese Academy of Sciences, Beijing, China,

² University of Chinese Academy of Sciences, Beijing, China

OPEN ACCESS

Edited by:

Sixue Chen,
University of Florida, USA

Reviewed by:

David Twell,
University of Leicester, UK
Mengmeng Zhu,
The Pennsylvania State University,
USA

*Correspondence:

Tai Wang,
Key Laboratory of Plant Molecular
Physiology, Institute of Botany,
Chinese Academy of Sciences,
20 Nanxincun, Xiangshan, Haidianqu,
Beijing 100093, China
twang@ibcas.ac.cn

Specialty section:

This article was submitted to
Plant Systems and Synthetic Biology,
a section of the journal
Frontiers in Plant Science

Received: 14 April 2015

Accepted: 16 May 2015

Published: 02 June 2015

Citation:

Lu Y, Wei L and Wang T (2015)
Methods to isolate a large amount
of generative cells, sperm cells
and vegetative nuclei from tomato
pollen for “omics” analysis.
Front. Plant Sci. 6:391.
doi: 10.3389/fpls.2015.00391

The development of sperm cells (SCs) from microspores involves a set of finely regulated molecular and cellular events and the coordination of these events. The mechanisms underlying these events and their interconnections remain a major challenge. Systems analysis of genome-wide molecular networks and functional modules with high-throughput “omics” approaches is crucial for understanding the mechanisms; however, this study is hindered because of the difficulty in isolating a large amount of cells of different types, especially generative cells (GCs), from the pollen. Here, we optimized the conditions of tomato pollen germination and pollen tube growth to allow for long-term growth of pollen tubes *in vitro* with SCs generated in the tube. Using this culture system, we developed methods for isolating GCs, SCs and vegetative cell nuclei (VN) from just-germinated tomato pollen grains and growing pollen tubes and their purification by Percoll density gradient centrifugation. The purity and viability of isolated GCs and SCs were confirmed by microscopy examination and fluorescein diacetate staining, respectively, and the integrity of VN was confirmed by propidium iodide staining. We could obtain about 1.5 million GCs and 2.0 million SCs each from 180 mg initiated pollen grains, and 10 million VN from 270 mg initiated pollen grains germinated *in vitro* in each experiment. These methods provide the necessary preconditions for systematic biology studies of SC development and differentiation in higher plants.

Keywords: *Solanum lycopersicum*, generative cell, sperm cell, vegetative nuclei, isolation, Percoll density gradient centrifugation

Introduction

During the development of sperm cells (SCs, male gamete) from microspores in higher plants, the microspore generated from diploid microsporocytes via meiosis first undergoes asymmetric mitosis to produce a larger vegetative cell (VC) and a smaller generative cell (GC) embedded in the VC. Thereafter, the VC exits the cell cycle and has potential to generate a polarly growing pollen tube; the GC enters further mitosis to produce two SCs for double fertilization (McCormick, 1993; Twell et al., 1998; Twell, 2011). Depending on the plant species, GC mitosis occurs before anthesis or in growing pollen tubes; therefore, released mature pollen at anthesis is tricellular in some species such as *Oryza sativa*, *Zea mays*, and *Arabidopsis thaliana* (Berger and Twell, 2011) or bicellular in other species such as *Lilium brownii* and *Solanum lycopersicum*. This development

process involves a set of fine-tuned molecular and cellular events and the coordination of these events, such as cell cycle regulation, cell differentiation and fate determination, genome stability, and epigenetic reprogramming. Although genetic studies have functionally identified many important genes involved in plant SC development, such as *DUO1*, *DUO3*, *DAZ1*, and *DAZ2* (Borg et al., 2009, 2011, 2014; Brownfield et al., 2009a,b; Twell, 2011), the mechanisms underlying these events and their interconnections remain a major challenge for plant science. Systematic “omics” studies of the development process are essential for understanding the mechanisms.

“Omics” studies of pollen from several plants including *Arabidopsis* and rice have provided insights into the molecular mechanisms of pollen development (Rutley and Twell, 2015). During postmeiotic development from microspores, pollen express a set of specific transcripts; the total number of transcripts expressed is decreased, but the proportion of pollen-specific or preferential transcripts is increased (Honys and Twell, 2004; Wang et al., 2008; Wei et al., 2010). The composition and expression profile of miRNAs expressed in developing pollen differs from those in sporophytes, and novel and non-conserved known miRNAs are the main contributors to the difference (Wei et al., 2011). In pollen, small RNA displays cell-specific activity: working by translational repression in the SC, and by cleavage-induced mRNA turnover in the VC (Grant-Downton et al., 2013). The small RNA from the VC are strongly implicated in gene silencing in SCs (Slotkin et al., 2009; Grant-Downton et al., 2013). This indicates reprogramming of gene expression during pollen development and the importance of epigenetic signals in this reprogramming. In addition, proteomics and metabolomics studies have revealed the importance of presynthesized proteins during pollen maturation in pollen function (Holmes-Davis et al., 2005; Dai et al., 2006), and difference in proteomes and metabolic pathways between mature and germinated pollen (Dai et al., 2007; Obermeyer et al., 2013). These studies also revealed many important candidate genes for further understanding the molecular control of pollen development by functionally dissecting these candidates.

Recent studies have isolated SCs from tricellular pollen of rice and *Arabidopsis* and analyzed the transcriptome of SCs (Borges et al., 2008; Russell et al., 2012). The transcriptome of the SC was significantly different from that of the pollen grain, which is consistent with the SC being only a little part of the pollen grain that is mainly represented by the VC. SC-preferential transcripts showed a prominent functional skew toward epigenetic regulation, DNA repair, and cell cycles (Borges et al., 2008; Russell et al., 2012). Small RNA-mediated DNA methylation in SCs is associated with epigenetic inheritance, transposon silencing and paternal imprinting (Borges et al., 2008; Calarco et al., 2012). Further systematic “omics” analysis of molecular programs for SC development from its precursors, the GC and microspore, is essential to understand the mechanism of SC development. To achieve this goal, we need to establish a condition to isolate GCs and SCs from the pollen of a species. Because the GC occurs at a short time window *in vivo* and develops asynchronously in different flowers in rice and

Arabidopsis, isolating a large amount of GCs at high purity from developing pollen of these species for “omics” analysis is difficult.

Tomato is another model plant to study pollen development (Twell et al., 1990, 1991; Muschietti et al., 1994; Filichkin et al., 2004) and can be an excellent model to achieve the above target because (1) its genome has been sequenced (Sato et al., 2012) and (2) its mature pollen is bicellular. This feature of pollen indicates the possibility to isolate GCs from pollen grains or just-germinated pollen grains (JPGGs) and to isolate SCs from pollen tubes with SCs formed from GCs via mitosis.

In this study, we optimized the conditions of pollen germination and pollen tube growth to allow for long-term growth of pollen tubes *in vitro* with SCs generated in the tube. Using this culture system, we developed efficient protocols to isolate a large amount of GCs, SCs, and vegetative cell nuclei (VN) at high purity to satisfy the demands of “omics” study.

Materials and Methods

Plants Growth and Pollen Collection

Tomato (*S. lycopersicum*) plants (Heinz 1706) were grown in the greenhouse under long-day conditions (14 h light/10 h dark) at 25 ~ 35°C. During anthesis, anthers from opened flowers were collected and dried for 10 h at 28°C in an electrothermal drying closet, then placed into a colander (85 mm × 50 mm) with mesh at 63- μ m-pore size; mature pollen was released and collected by shaking the colander vigorously. Pollen grains were used immediately or stored in a 1.5 mL tube with 5~10 particles of Silica gel Rubin (Sigma, 85815) at -20°C.

Pollen Germination *In Vitro* and Morphologic Observation

Mature pollen grains (60 mg) were pre-hydrated in a Petri dish (60 mm × 15 mm), which was covered with gauze and then placed in a large Petri dish (150 mm × 25 mm) with 50 mL saturated Na₂HPO₄ at 25°C for 4~8 h. This device only permitted gauze contact this solution, and prohibited pollen grains contact the gauze and solution directly. Hydrated pollen grains were incubated in 100 mL germination medium (20 mM MES, 3 mM Ca(NO₃)₂, 1 mM KCl, 0.8 mM MgSO₄, 1.6 mM boric acid, 24% PEG 4000, 2.5% sucrose, pH 6.0; osmotic pressure, 1253.33 ± 2.33 mOsmol/kg H₂O) in a Petri dish (150 × 25 mm) at 25°C in the dark with shaking at 90 rpm (Tang et al., 2002; Zhao et al., 2013). During germination, 1 mL medium was took out at regular intervals, centrifuged to collect germinating pollen grains, then transferred to 1 mL Carnoy's fluid (three parts of absolute ethyl alcohol, one part of acetic acid) for treatment of 30 min. Thereafter, these treated pollen grains or tubes were stained with 4',6-diamino-2-phenylindole (DAPI; Molecular Probes) and observed under a microscope (Zeiss Axio Imager A1).

Isolation of GCs

An improved two-step osmotic shock was used to release GCs from JPGGs and all procedures were performed at room

temperature (Zhao et al., 2013). Two aliquots of pollen grains 90 mg each were germinated in 100 mL germination medium as described above for ~20 min until the pollen tubes emerged but were shorter than the diameter of the grains. JPGs were harvested through a Büchner funnel (100 mm in diameter) with 11- μ m hydrated nylon mesh (Millipore, NY1100010) with the help of an aspirator pump, then rinsed with osmotic shock solution (15.3% sucrose, 1% bovine serum albumin [BSA], 531.33 ± 3.84 mOsmol/kg H₂O) to clean germination medium, which would affect the result of osmotic shock. The collected JPGs were immediately transferred to 80 mL fresh osmotic shock solution and incubated for 10 min to burst tubes and release GCs. Cell debris was removed by sieving the mixture through a hydrated 11- μ m nylon mesh. The filtrate containing GCs was equally divided into two centrifugation tubes, and centrifuged at 850 g for 4 min to collect GCs. To avoid loss of GCs, we retained 10 ml of the supernatant in each tube after centrifugation, then added 10 mL of isolation buffer 1 (IB1; 20 mM MES-KOH, 20 mM NaCl, 10 mM EDTANa₂, 1 mM spermidine, 0.3 mM spermine, 2 mM DTT, 18% sucrose and 1% BSA, pH6.0) to suspend GCs. The suspension was supplemented with stock solution of cellulase “Onozuka” R-10 (Yakult) and macerozyme R-10 (Yakult; 0.4% each in IB2; IB2; 10 mM MES-KOH, 10 mM NaCl, 5 mM EDTANa₂, 0.5 mM spermidine, 0.15 mM spermine, 1 mM DTT, 18% sucrose and 1% BSA, pH6.0) to a final concentration of 0.04% each enzyme, mixed gently and incubated for 15 min without shaking, and centrifuged at 850 g for 3 min to collect GCs, followed by a washing with IB2. Collected GCs were further purified on 23/32% Percoll density gradient (2 mL 23% Percoll and 3 mL 32% Percoll in IB2) by horizontal centrifugation at 1000 g for 40 min. After centrifugation, GCs partitioned at the interface of 23% and 32% Percoll were collected with use of a glass pipette and washed twice with 3 \times volume IB2 followed by centrifugation at 950 g for 3 min each. The viability of isolated GCs was examined by fluorescein diacetate (FDA) staining. The purified GCs were snap-frozen in liquid nitrogen and stored at -80°C .

Isolation of SCs

Sperm cells were isolated under room temperature as described by Xu et al. (2002) with modifications. In brief, three aliquots of pollen grains 60 mg each were cultured in 100 mL germination medium as described above for 10 h. After germination, medium was removed by use of hydrated 100- μ m nylon mesh (Millipore, NY1H00010), and pollen tubes were washed with osmotic shock solution as for isolation of GCs, immediately transferred to low-osmotic enzyme solution (0.4% cellulase “Onozuka” R-10 and 0.2% macerozyme R-10 in osmotic shock solution), and incubated for 5 min to release SCs. Cell debris and ungerminated pollen grains were removed by use of hydrated 11- μ m nylon mesh, and SCs in the filtrate were collected and washed as described in isolation of GCs. Thereafter, collected SCs were purified on 5 mL 23% Percoll gradient in IB2 by horizontal centrifugation at 1000 g for 30 min. SCs were enriched to the upper surface of 23% Percoll gradient and harvested by use of a glass pipette and washed twice with 3 \times volume IB2 followed by centrifugation at 950 g for 3 min each. The viability of isolated

SCs was examined by FDA staining. The purified SCs were snap-frozen in liquid nitrogen and stored at -80°C .

Isolation of VN

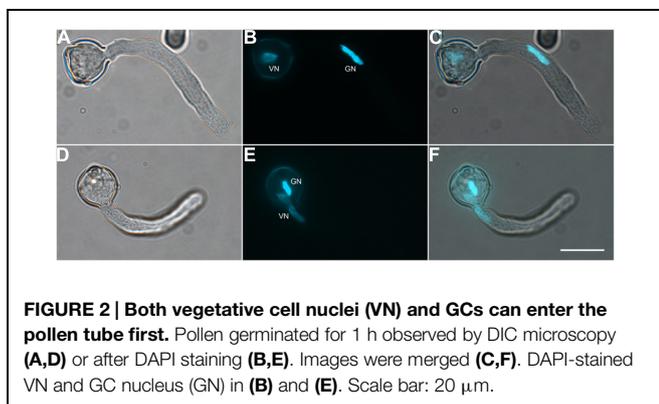
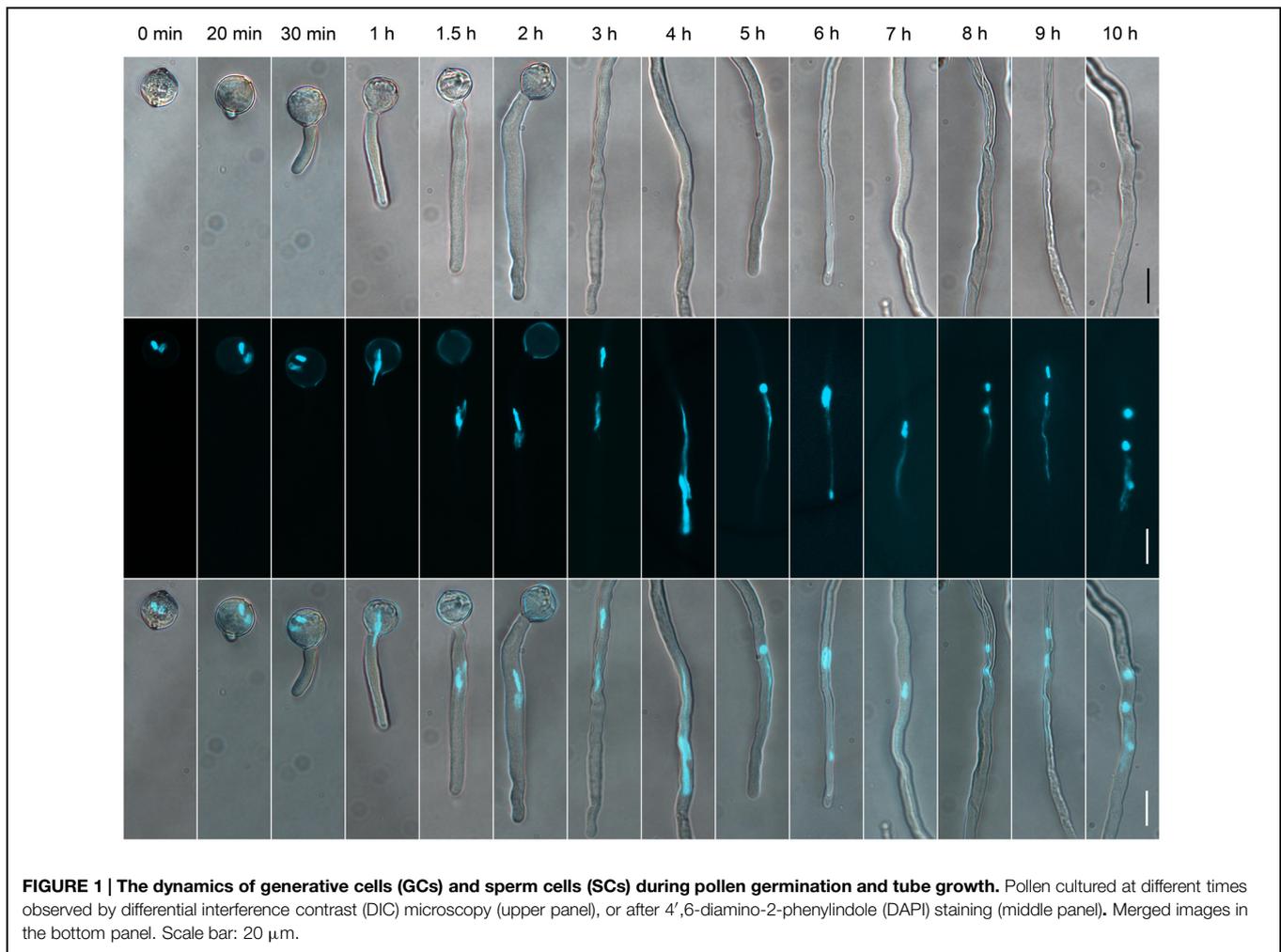
All operations were performed on ice or at 4°C unless otherwise specified, and all solutions were pre-cooled on ice or at 4°C . Three aliquots of pollen grains 90 mg each were cultured in 100 mL germination medium as described above for 1.5 h. Pollen tubes were collected with a 20- μ m hydrated nylon mesh (Millipore, NY2009000) at room temperature, rinsed with wash buffer (10 mM MOPS-NaOH, 2.5% sucrose, 9.5% mannitol, 5 mM EDTANa₂, 1% BSA, pH7.2), and treated with 12 mL enzyme solution (0.5% cellulase “Onozuka” R-10, and 0.3% macerozyme R-10 in wash buffer) for 5 min to release VN. After removal of ungerminated pollen grains and cell debris with use of 20- μ m hydrated nylon mesh, the filtrate containing VN was divided into four equal parts, loaded onto the surface of 3 mL 10% Percoll gradient in wash buffer each, then centrifuged at 1500 g for 30 min. VN in the upper surface of the gradient were collected by use of a glass pipette, snap-frozen in liquid nitrogen, then stored at -80°C .

Results

Dynamics of GCs and SCs During Culture *In Vitro*

Our experiments showed that low-temperature (-20°C) stored tomato pollen grains without prehydration germinated *in vitro* at low germination rate (Supplementary Table S1). To guarantee a high proportion of synchronously germinated tomato pollen grains after low-temperature storage and long-term growth of pollen tubes to allow generation of SCs in the tube *in vitro*, we optimized the pre-hydration condition of the stored pollen grains and culture condition of pollen tubes. We prehydrated low-temperature-stored tomato pollen grains using the saturated solution of Na₂HPO₄ to rescue the germination activity. Prehydration with the saturated solution for 4–8 h significantly increased germination rate of the pollen grains (Supplementary Table S1). Our culture conditions allowed for *in vitro* growth of pollen tubes for > 10 h (Figure 1).

To determine the suitable time of pollen tube growth for isolating GCs, SCs, and VN, we examined their dynamics during pollen germination and tube growth by DAPI staining. A bulge appeared at the germination aperture of hydrated pollen grains on culture for 20 min, and the bulge emerged as a morphologically visible pollen tube with a length shorter than or equal to the diameter of the pollen grain during 30-min culture (Figure 1). With increased culture time, GCs and VN began to move into the tube at 1 h and completely entered the tube at 1.5 h. We used DAPI staining to determine the movement order of VN and GCs (Figure 2). Among 126 surveyed pollen grains, for 68, VN entered the tube first, and for 58, GCs entered first. Therefore, during tomato pollen tube growth, VN and GCs may move into the tube in a random order. Furthermore, for GCs, 77.8% completed mitosis to generate SCs at 8 h, 84.2% at 9 h, and 92.4% at 10 h (Supplementary Table S2).



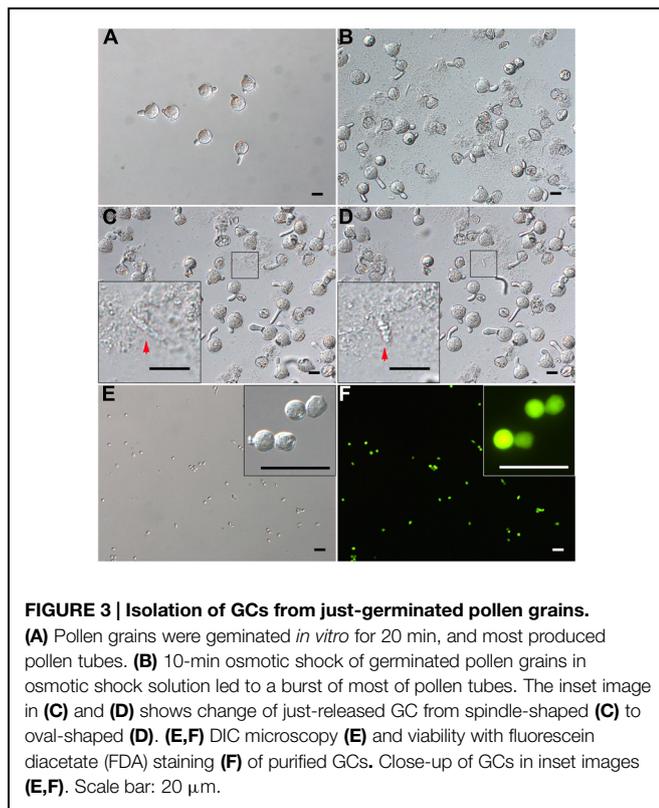
Release and Purification of GCs

Tomato pollen grains could not burst directly with osmotic shock and also could not germinate in a sucrose solution alone (data not shown). So, we developed a modified two-step method. We incubated prehydrated tomato pollen grains in germination medium for 20 min, when a bulge appeared at the aperture (Figures 1 and 3A), then osmotically shocked

the JGPGs, which were sensitive to the low-osmotic shock (Figure 3B).

When JGPGs were transferred into low-osmotic solution, the tube burst, and GCs, along with the cytoplasm, were emitted (Figures 3C,D). The just-released GCs underwent a change from spindle- to oval-shaped (Figures 3C,D). This change may be associated with the microtubule cytoskeleton, which was dynamic in response to environmental conditions and is important to determine the shape of GCs (Zhou et al., 1990). When incubated in the low-osmotic solution for 10 min, 62.9% (680/1082) of pollen tubes burst (Figure 3B). The released GCs were intact in the low-osmotic solution up to 1 h, but most GCs appeared to break after 1 h (data not shown). To maintain GC integrity, we added IB1 into the filtrate containing GCs to neutralize the low osmotic shock in a short time (<1 h). The suspension was used for purifying GCs.

Furthermore, we treated the suspension with cellulase and macerozyme at a low concentration, which had no effect on viability of GCs but increased the efficiency in removing cell debris at subsequent gradient centrifugation. GCs were enriched at the interface of 23% and 32% Percoll gradient, with cell debris on the upper interface of 23% Percoll. The isolated GCs



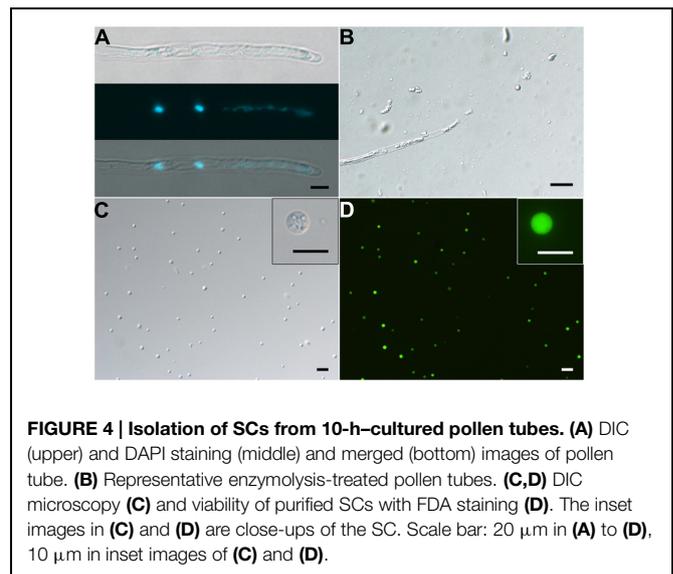
were viable on FDA staining (**Figures 3E,F**) and had no VN contamination, as confirmed by propidium iodide (PI) staining (viable GCs cannot be stained by PI; Supplementary Figures S1A,B). Finally, we obtained about 1.5 million GCs from 180 mg of initiated mature pollen grains (about 18 million grains, in that 0.1 million tomato pollen grains is about 1 mg).

Release and Purification of SCs

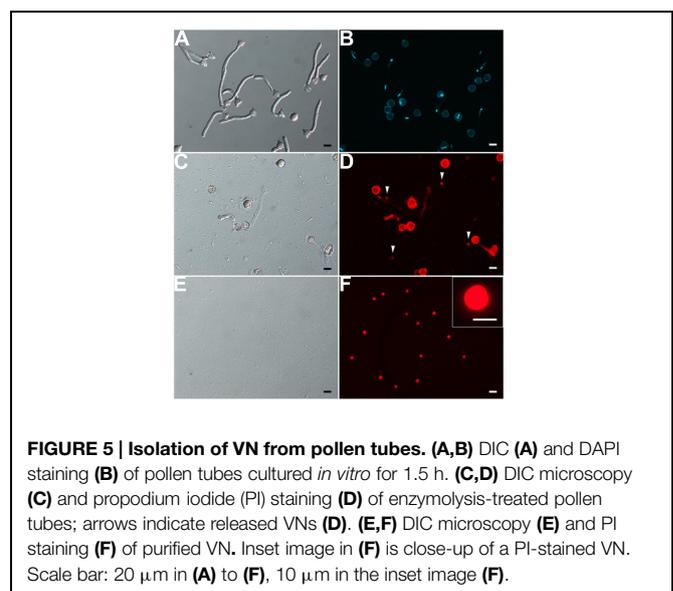
Successful isolation of SCs from *in vitro*-cultured pollen tubes depends on the formation of SCs in the growing tubes. We isolated SCs from 10-h-cultured pollen tubes, in which GCs had completed mitosis to generate SCs (see above, **Figure 4A**; Supplementary Table S2). To decrease the possible contamination of GCs, we collected long pollen tubes using a large pore-size nylon mesh (100 μm), which allowed ungerminated pollen grains and shorter pollen tubes to pass through. We found that osmotic shock alone did not burst the long tubes efficiently (data not shown), and a modified low osmotic solution with cellulase and macerozyme was efficient to burst the tube (**Figure 4B**). After removal of cell debris, SCs in filtrate could be enriched with a layer of 23% Percoll. Finally, we obtained about 2 million viable SCs at high purity from 180 mg initiated pollen grains (**Figures 4C,D**), with no VN contamination (Supplementary Figures S1C,D).

Release and Purification of VN

Our results showed that VN was fragile and disrupted quickly as released to medium at room temperature. Repeated pipetting also



led to its complete disruption (Supplementary Figures S2A,B,C). Therefore, no VN contamination was present in isolated GCs and SCs (see above). We solved the bottleneck of VN isolation by (1) keeping all operations at 4°C or on ice, (2) avoiding pipetting as much as possible, and (3) using 1.5-h-cultured pollen tubes in which VN had moved into the tube (**Figures 1** and **5A,B**), for easier release of VN (**Figures 5C,D**). Furthermore, additional washing as well as passing through Percoll on gradient centrifugation disrupted VN (Supplementary Figures S2A,D), so we used only a hydrated nylon net filter to remove pollen grains and cell debris and then enriched VN by using a layer of 10% Percoll. These measures allowed for isolation of VN without GC contamination (**Figures 5E,F**). Using this protocol, we obtained 10 million VN from 270 mg initiated pollen grains.



Discussion

We have optimized the conditions allowing for growth of pollen tubes for more than 10 h and generation of SCs in tubes, as well as conditions affecting rupture of pollen grains (tubes) and release of cytoplasm, GCs, SCs and VN into medium. Finally, we developed methods to isolate GCs, SCs, and VN from JPGs and 1.5-h- and 10-h-cultured pollen tubes, respectively (Figure 6). These methods allowed for isolating large amounts of GCs, SCs, and VN at high purity.

Culture Conditions for Low-Temperature-Stored Pollen Grains and Long-Term-Cultured Pollen Tubes

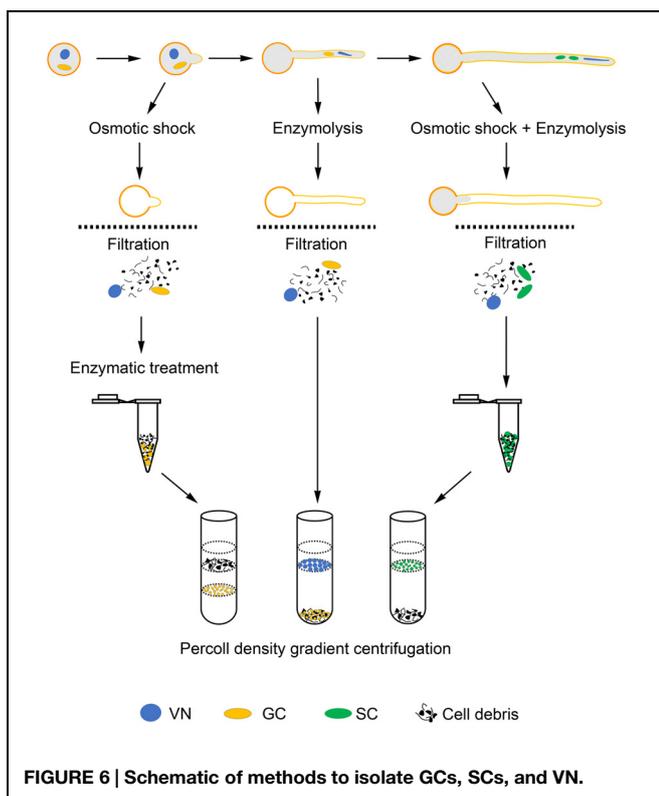
Previous study established the condition for *in vitro* germination of fresh pollen grains from tomato (Tang et al., 2002), but under the condition, low-temperature-stored tomato pollen grains did not germinate well (Supplementary Table S1). Generally, pre-hydration is required for rescuing the viability of low-temperature-stored pollen grains, such as from *Rosa*, *Pistacia vera* L., *Gladiolus* sp. and *Brassica rapa* (Visser et al., 1977; Golan-Goldhirsh et al., 1991; Rajasekharan et al., 1994; Sato et al., 1998). The prehydration was usually actualized with water or saturated salt solution, which generated a fixed relative humidity in a chamber at a certain temperature. We found that a saturated solution of Na_2HPO_4 was suitable for prehydrating tomato pollen grains. The saturated solution was previously used to prehydrate *B. rapa* pollen (Sato et al., 1998), and could result in 95% relative humidity in a chamber at 25°C, and tomato pollen grains rescued

under this condition germinated synchronously (Figure 3A), which was important for synchronous pollen tube growth and GC division. Pollen density also was an important factor affecting germination in tomato, and increased density led to increased germination percentage (Supplementary Figure S3), which agrees with observations in other species such as *Arabidopsis* (Boavida and McCormick, 2007), *Nicotiana*, *B. oleracea*, and *Betula pendula* (Roberts et al., 1983; Jahnen et al., 1989; Pasonen and Käpylä, 1998; Chen et al., 2000). Furthermore, we evaluated the effect of loaded pollen grain amount in a given volume germination medium on the integrity of pollen tubes during long-term culture (>3 h) by examining cell debris in the medium (Supplementary Figure S4). Cell debris was barely observed with ≤ 4 mg pollen grains used (Supplementary Figures S4A–D), and substantial with 5 mg pollen grains (Supplementary Figure S4E). Thus, the amount of pollen grains in a given volume germination medium is crucial to the integrity of long-term-cultured pollen tubes, and 3 mg pollen grains loaded in 5 mL germination medium was appropriate for long-term culture of pollen tubes.

Methods to Release GCs, SCs, and VN

Four major methods were used previously to break pollen grains or tubes: mechanical grinding, one-step and two-step osmotic shock and enzyme digestion (Russell, 1986, 1991; Zhou, 1988; Russell et al., 1990; Theunis et al., 1991; Chaboud and Perez, 1992; Xu et al., 2002; Borges et al., 2008; Zhao et al., 2013). The mechanical grinding had relatively low efficiency for breaking tomato pollen grains and produced a large amount of debris, which interfered with further purification. One-step or two-step osmotic shock is usually used to break pollen grains (tubes) and release target cells. The former breaks pollen grains and releases target cells simultaneously using a low-osmotic solution (Russell, 1986). The latter first makes the grain germinate in a sucrose solution, then release target cells under low-osmotic shock with a diluted sucrose solution (Zhou, 1988; Wu and Zhou, 1991). Tomato pollen grains were insensitive to osmotic shock of water, which is similar to pollen grains from *Vicia faba*, *B. napus*, and *L. davidii* var. *unicdor* (Zhou, 1988; Taylor et al., 1991; Zhao et al., 2013), thus suggesting a complicated mechanism of pollen grain burst for different species and individual methods needed for different species.

We choose the two-step osmotic shock to release GCs. Components of osmotic shock solution affected the appearance of released cytoplasm, which affected the following purification of GCs. The released cytoplasm appeared not to conglutinate when the solution had only 15.3% sucrose (Supplementary Figure S5A) but appeared to conglutinate with shock solution containing MES or MOPS regardless of concentration (Supplementary Figures S5B–I). Acidic pH could aggravate this situation (Supplementary Figures S5D,J,L). The phenomenon of cytoplasm clumping was also observed under high concentration of CaCl_2 in a previous study of isolating SCs from pollen tubes of *N. tabacum* (Tian and Russell, 1997). We chose sucrose and BSA as the components of the osmotic shock solution. The use of BSA in all solutions



except the germination medium aimed to protect GCs against damage because we found that GCs could remain intact in the shock solution with BSA up to 1 h but for only a few minutes in a shock solution without BSA; GCs from several species such as *V. faba* could remain intact in simple sucrose shock solution until being purified (Zhou, 1988).

However, pollen tubes cultured *in vitro* for 10 h, which we used to isolate SCs, were not as sensitive as pollen tubes cultured for a short time to osmotic shock described above. So we developed a modified shock solution containing cellulase and macroenzyme. The two enzymes are usually applied to digest hemicelluloses, cellulose and pectin, the main components of the pollen tube wall (Taylor and Hepler, 1997; Cheung and Wu, 2008; Rounds and Bezanilla, 2013). This modified shock solution broke the long pollen tubes and released SCs efficiently (Figure 4B). Why pollen tubes had different sensitivity to osmotic shock with long- and short-term culture needs further studies.

In contrast to reports of GCs and SCs, we have limited reports of VN isolation (Wever and Takats, 1971; LaFountain and Mascarenhas, 1972; Ueda and Tanaka, 1994; Borges et al., 2012; Calarco et al., 2012). These reports did not describe the stability of isolated VN or the effect of environmental conditions on the stability. We found that VN was fragile *in vitro* (Supplementary Figure S2) and easily ruptured at room temperature but was relatively stable at 4°C. However, at low temperature, pollen tubes were not sensitive to osmotic shock. In this situation, enzyme digestion was found efficient to break pollen tubes and release VN but depended on a suitable buffer. We found that pH value was a crucial factor of the buffer. The pH value affected the appearance of the released cytoplasm: conglutination appeared at acidic pH and not at alkaline pH (Supplementary Figure S6).

Measures to Guarantee the Purity of GCs, SCs, and VN

Previous reports mainly described isolation of SCs from tricellular pollen or GCs from bicellular pollen (Russell, 1986; Dupuis et al., 1987; Zhou, 1988; Southworth and Knox, 1989; Yang and Zhou, 1989; Xu et al., 2002; Engel et al., 2003), or along with VN (Borges et al., 2012). For these cases, possible contaminants were VN for isolated SCs or GCs and vice versa. Such purity evaluation was relatively simple. Most previous works evaluated the purity of isolated SCs or GCs based on their morphologic features observed on microscopy (Zhou, 1988; Southworth and Knox, 1989; Yang and Zhou, 1989; Xu et al., 2002; Engel et al., 2003). Only the study of *Arabidopsis* used SC- and VN-specific markers to estimate the purity of isolated SCs and VN because of available markers for this species (Borges et al., 2008, 2012).

An evaluation of the purity of isolated GCs, SCs, or VN from a species was relatively lacking in early studies. We found that a combination of PI staining and differential interference contrast (DIC) microscope observation was efficient to evaluate VN contamination of GCs or SCs and vice versa. VN was not observed on DIC microscopy but were detectable with PI staining. In contrast, viable GCs and SCs were easily observed on DIC microscopy but undetectable with PI staining. Using

the combination, we did not find VN contamination in isolated GCs and SCs or GC and SC contamination in isolated VN (Supplementary Figure S1, Figures 5E,F). We considered several measures to eliminate this contamination in designing methods to isolate GCs, SCs, and VN. (1) GCs or SCs were stable and isolated at room temperature, but under this temperature, VN was fragile and broke quickly. (2) GCs, SCs and VN had different density on Percoll gradient and could be enriched with different gradient ingredients. These measures could eliminate VN contamination in isolated GCs and SCs and vice versa (Supplementary Figure S1, Figures 5E,F).

A major challenge in methodology was how to get SCs at high purity. Here, besides the use of high synchronous pollen tubes, of which 92.4% generated SCs (Supplementary Table S2), and the measures above, we collected long pollen tubes by using a large-pore mesh (100 μm in pore diameter), which excluded ungerminated pollen grains (diameter of about 20 μm) and short pollen tubes. Thus, the methods could obtain SCs at high purity. However, tools to distinguish GCs and SCs are lacking because of their similar morphology under light microscope in tomato and other most species and lack of molecular markers.

Conclusion

Saturated Na₂HPO₄ solution was suitable for pre-hydration of low-temperature-stored tomato pollen grains, and the prehydrated pollen grains germinated synchronously. The loaded amount of 0.6 mg pollen grains per mL allowed pollen tubes to grow for more than 10 h, and more than 92% GCs completed mitosis to generate SCs. GCs or SCs were stable and could be isolated at room temperature, whereas under the same temperature, VN was fragile and broke quickly *in vitro*. GCs, SCs, and VN had different density on Percoll gradient, and could be enriched with different gradient ingredients. Thus, we have established methods to isolate GCs and VN from just-germinated pollen grains and 1.5-h-cultured pollen tubes, respectively, and SCs from 10-h-cultured pollen tubes. Using these methods, we could obtain 1.5 million GCs and 2 million SCs each from 180 mg initiated pollen grains, and 10 million VN from 270 mg initiated pollen grains, for higher productivity as compared with previous reports of other species.

Acknowledgments

We thank Xin Zhao for help in developing these methods, and Fan Yang, Bo Yu, Yunyun Song, Yuxia Liu for help in collecting tomato pollen grains. This work was supported by the Chinese Ministry of Science and Technology (grant no. 2012CB910504).

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2015.00391/abstract>

References

- Berger, F., and Twell, D. (2011). Germline specification and function in plants. *Annu. Rev. Plant Biol.* 62, 461–484. doi: 10.1146/annurev-arplant-042110-103824
- Boavida, L. C., and McCormick, S. (2007). Temperature as a determinant factor for increased and reproducible *in vitro* pollen germination in *Arabidopsis thaliana*. *Plant J.* 52, 570–582. doi: 10.1111/j.1365-313X.2007.03248.x
- Borg, M., Brownfield, L., Khatab, H., Sidorova, A., Lingaya, M., and Twell, D. (2011). The R2R3 MYB transcription factor DUO1 activates a male germline-specific regulon essential for sperm cell differentiation in *Arabidopsis*. *Plant Cell* 23, 534–549. doi: 10.1105/tpc.110.081059
- Borg, M., Brownfield, L., and Twell, D. (2009). Male gametophyte development: a molecular perspective. *J. Exp. Bot.* 60, 1465–1478. doi: 10.1093/jxb/Ern355
- Borg, M., Rutley, N., Kagale, S., Hamamura, Y., Gherghinoiu, M., Kumar, S., et al. (2014). An EAR-dependent regulatory module promotes male germ cell division and sperm fertility in *Arabidopsis*. *Plant Cell* 26, 2098–2113. doi: 10.1105/tpc.114.124743
- Borges, F., Gardner, R., Lopes, T., Calarco, J. P., Boavida, L. C., Slotkin, R. K., et al. (2012). FACS-based purification of *Arabidopsis* microspores, sperm cells and vegetative nuclei. *Plant Methods* 8:44. doi: 10.1186/1746-4811-8-44
- Borges, F., Gomes, G., Gardner, R., Moreno, N., McCormick, S., Feijó, J. A., et al. (2008). Comparative transcriptomics of *Arabidopsis* sperm cells. *Plant Physiol.* 148, 1168–1181. doi: 10.1104/pp.108.125229
- Brownfield, L., Hafidh, S., Borg, M., Sidorova, A., Mori, T., and Twell, D. (2009a). A plant germline-specific integrator of sperm specification and cell cycle progression. *PLoS Genet.* 5:e1000430. doi: 10.1371/journal.pgen.1000430
- Brownfield, L., Hafidh, S., Durbarray, A., Khatab, H., Sidorova, A., Doerner, P., et al. (2009b). *Arabidopsis* DUO POLLEN3 is a key regulator of male germline development and embryogenesis. *Plant Cell* 21, 1940–1956. doi: 10.1105/tpc.109.066373
- Calarco, J. P., Borges, F., Donoghue, M. T., Van Ex, F., Jullien, P. E., Lopes, T., et al. (2012). Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. *Cell* 151, 194–205. doi: 10.1016/j.cell.2012.09.001
- Chaboud, A., and Perez, R. (1992). Generative cells and male gametes: isolation, physiology, and biochemistry. *Int. Rev. Cytol.* 140, 205–232. doi: 10.1016/S0074-7696(08)61098-0
- Chen, Y. F., Matsubayashi, Y., and Sakagami, Y. (2000). Peptide growth factor phytosulfokine- α contributes to the pollen population effect. *Planta* 211, 752–755. doi: 10.1007/s004250000370
- Cheung, A. Y., and Wu, H. M. (2008). Structural and signaling networks for the polar cell growth machinery in pollen tubes. *Annu. Rev. Plant Biol.* 59, 547–572. doi: 10.1146/annurev-arplant.59.032607.092921
- Dai, S. J., Chen, T. T., Chong, K., Xue, Y. B., Liu, S. Q., and Wang, T. (2007). Proteomics identification of differentially expressed proteins associated with pollen germination and tube growth reveals characteristics of germinated *Oryza sativa* pollen. *Mol. Cell. Proteomics* 6, 207–230. doi: 10.1074/mcp.M600146-MCP200
- Dai, S., Li, L., Chen, T., Chong, K., Xue, Y., and Wang, T. (2006). Proteomic analyses of *Oryza sativa* mature pollen reveal novel proteins associated with pollen germination and tube growth. *Proteomics* 6, 2504–2529. doi: 10.1002/pmic.200401351
- Dupuis, I., Roedel, P., Matthys-Rochon, E., and Dumas, C. (1987). Procedure to isolate viable sperm cells from corn (*Zea mays* L.) pollen grains. *Plant Physiol.* 85, 876–878. doi: 10.1104/pp.85.4.876
- Engel, M. L., Chaboud, A., Dumas, C., and McCormick, S. (2003). Sperm cells of *Zea mays* have a complex complement of mRNAs. *Plant J.* 34, 697–707. doi: 10.1046/j.1365-313X.2003.01761.x
- Filichkin, S. A., Leonard, J. M., Monteros, A., Liu, P. P., and Nonogaki, H. (2004). A novel endo- β -mannanase gene in tomato LeMAN5 is associated with anther and pollen development. *Plant Physiol.* 134, 1080–1087. doi: 10.1104/pp.103.035998
- Golan-Goldhirsh, A., Schmidhalter, U., Müller, M., and Oertli, J. J. (1991). Germination of *Pistacia vera* L. pollen in liquid medium. *Sex. Plant Reprod.* 4, 182–187. doi: 10.1007/BF00190002
- Grant-Downton, R., Kourmpetli, S., Hafidh, S., Khatab, H., Le Trionnaire, G., Dickinson, H., et al. (2013). Artificial microRNAs reveal cell-specific differences in small RNA activity in pollen. *Curr. Biol.* 23, R599–R601. doi: 10.1016/j.cub.2013.05.055
- Holmes-Davis, R., Tanaka, C. K., Vensel, W. H., Hurkman, W. J., and McCormick, S. (2005). Proteome mapping of mature pollen of *Arabidopsis thaliana*. *Proteomics* 5, 4864–4884. doi: 10.1002/pmic.200402011
- Honys, D., and Twell, D. (2004). Transcriptome analysis of haploid male gametophyte development in *Arabidopsis*. *Genome Biol.* 5, R85. doi: 10.1186/Gb-2004-5-11-R85
- Jahnen, W., Lush, W. M., and Clarke, A. E. (1989). Inhibition of *in vitro* pollen tube growth by isolated S-glycoproteins of *Nicotiana glauca*. *Plant Cell* 1, 501–510. doi: 10.1105/tpc.1.5.501
- LaFountain, K. L., and Mascarenhas, J. P. (1972). Isolation of vegetative and generative nuclei from pollen tubes. *Exp. Cell Res.* 73, 233–236. doi: 10.1016/0014-4827(72)90125-5
- McCormick, S. (1993). Male gametophyte development. *Plant Cell* 5, 1265–1275. doi: 10.2307/3869779
- Muschietti, J., Dircks, L., Vancanneyt, G., and McCormick, S. (1994). Lat52 protein is essential for tomato pollen development: pollen expressing antisense *Lat52* RNA hydrates and germinates abnormally and cannot achieve fertilization. *Plant J.* 6, 321–338. doi: 10.1046/j.1365-313X.1994.06030321.x
- Obermeyer, G., Fagner, L., Lang, V., and Weckwerth, W. (2013). Dynamic adaption of metabolic pathways during germination and growth of lily pollen tubes after inhibition of the electron transport chain. *Plant Physiol.* 162, 1822–1833. doi: 10.1104/pp.113.219857
- Pasonen, H. L., and Käpylä, M. (1998). Pollen-pollen interactions in *Betula pendula* *in vitro*. *New Phytol.* 138, 481–487. doi: 10.1046/j.1469-8137.1998.00135.x
- Rajasekharan, P. E., Rao, T. M., Janakiram, T., and Ganeshan, S. (1994). Freeze preservation of gladiolus pollen. *Euphytica* 80, 105–109. doi: 10.1007/Bf00039304
- Roberts, I. N., Gaude, T. C., Harrod, G., and Dickinson, H. G. (1983). Pollen-stigma interactions in *Brassica oleracea*; a new pollen germination medium and its use in elucidating the mechanism of self incompatibility. *Theor. Appl. Genet.* 65, 231–238. doi: 10.1007/Bf00308074
- Rounds, C. M., and Bezanilla, M. (2013). Growth mechanisms in tip-growing plant cells. *Annu. Rev. Plant Biol.* 64, 243–265. doi: 10.1146/annurev-arplant-050312-120150
- Russell, S. D. (1986). Isolation of sperm cells from the pollen of *Plumbago zeylanica*. *Plant Physiol.* 81, 317–319. doi: 10.1104/pp.81.1.317
- Russell, S. D. (1991). Isolation and characterization of sperm cells in flowering plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 42, 189–204. doi: 10.1146/annurev-arplant.42.1.189
- Russell, S. D., Cresti, M., and Dumas, C. (1990). Recent progress on sperm characterization in flowering plants. *Physiol. Plant.* 80, 669–676. doi: 10.1034/j.1399-3054.1990.800427.x
- Russell, S. D., Gou, X. P., Wong, C. E., Wang, X. K., Yuan, T., Wei, X. P., et al. (2012). Genomic profiling of rice sperm cell transcripts reveals conserved and distinct elements in the flowering plant male germ lineage. *New Phytol.* 195, 560–573. doi: 10.1111/j.1469-8137.2012.04199.x
- Rutley, N., and Twell, D. (2015). A decade of pollen transcriptomics. *Plant Reprod.* 28, 73–89. doi: 10.1007/s00497-015-0261-7
- Sato, S., Katoh, N., Iwai, S., and Hagimori, M. (1998). Establishment of reliable methods of *in vitro* pollen germination and pollen preservation of *Brassica rapa* (syn. *B. campestris*). *Euphytica* 103, 29–33. doi: 10.1023/A:1018381417657
- Sato, S., Tabata, S., Hirakawa, H., Asamizu, E., Shirasawa, K., Isobe, S., et al. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485, 635–641. doi: 10.1038/Nature11119
- Slotkin, R. K., Vaughn, M., Borges, F., Tanurdžić, M., Becker, J. D., Feijó, J. A., et al. (2009). Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell* 136, 461–472. doi: 10.1016/j.cell.2008.12.038
- Southworth, D., and Knox, R. B. (1989). Flowering plant sperm cells: isolation from pollen of *Gerbera jamesonii* (Asteraceae). *Plant Sci.* 60, 273–277. doi: 10.1016/0168-9452(89)90177-5
- Tang, W. H., Ezcurra, I., Muschietti, J., and McCormick, S. (2002). A cysteine-rich extracellular protein, LAT52, interacts with the extracellular domain of the pollen receptor kinase LePRK2. *Plant Cell* 14, 2277–2287. doi: 10.1105/TPc.003103

- Taylor, L. P., and Hepler, P. K. (1997). Pollen germination and tube growth. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 48, 461–491. doi: 10.1146/annurev.arplant.48.1.461
- Taylor, P. E., Kenrick, J., Blomstedt, C. K., and Knox, R. B. (1991). Sperm cells of the pollen tubes of *Brassica*: ultrastructure and isolation. *Sex. Plant Reprod.* 4, 226–234. doi: 10.1007/BF00190009
- Theunis, C. H., Pierson, E. S., and Cresti, M. (1991). Isolation of male and female gametes in higher plants. *Sex. Plant Reprod.* 4, 145–154. doi: 10.1007/BF00189998
- Tian, H. Q., and Russell, S. D. (1997). Micromanipulation of male and female gametes of *Nicotiana tabacum* L. Isolation of gametes. *Plant Cell Rep.* 16, 555–560. doi: 10.1007/BF01142323
- Twell, D. (2011). Male gametogenesis and germline specification in flowering plants. *Sex. Plant Reprod.* 24, 149–160. doi: 10.1007/s00497-010-0157-5
- Twell, D., Park, S. K., and Lalanne, E. (1998). Asymmetric division and cell-fate determination in developing pollen. *Trends Plant Sci.* 3, 305–310. doi: 10.1016/S1360-1385(98)01277-1
- Twell, D., Yamaguchi, J., and McCormick, S. (1990). Pollen-specific gene expression in transgenic plants: coordinate regulation of two different tomato gene promoters during microsporogenesis. *Development* 109, 705–713.
- Twell, D., Yamaguchi, J., Wing, R. A., Ushiba, J., and McCormick, S. (1991). Promoter analysis of genes that are coordinately expressed during pollen development reveals pollen-specific enhancer sequences and shared regulatory elements. *Genes Dev.* 5, 496–507. doi: 10.1101/Gad.5.3.496
- Ueda, K., and Tanaka, I. (1994). The basic proteins of male gametic nuclei isolated from pollen grains of *Lilium longiflorum*. *Planta* 192, 446–452. doi: 10.1007/BF00198582
- Visser, T., De Vries, D. P., Welles, G. W. H., and Scheurink, J. A. M. (1977). Hybrid tea-rose pollen. I. Germination and storage. *Euphytica* 26, 721–728. doi: 10.1007/Bf00021697
- Wang, Y., Zhang, W. Z., Song, L. F., Zou, J. J., Su, Z., and Wu, W. H. (2008). Transcriptome analyses show changes in gene expression to accompany pollen germination and tube growth in *Arabidopsis*. *Plant Physiol.* 148, 1201–1211. doi: 10.1104/pp.108.126375
- Wei, L. Q., Xu, W. Y., Deng, Z. Y., Su, Z., Xue, Y. B., and Wang, T. (2010). Genome-scale analysis and comparison of gene expression profiles in developing and germinated pollen in *Oryza sativa*. *BMC Genomics* 11:338. doi: 10.1186/1471-2164-11-338
- Wei, L. Q., Yan, L. F., and Wang, T. (2011). Deep sequencing on genome-wide scale reveals the unique composition and expression patterns of microRNAs in developing pollen of *Oryza sativa*. *Genome Biol.* 12:R53. doi: 10.1186/Gb-2011-12-6-R53
- Wever, G. H., and Takats, S. T. (1971). Isolation and separation of S-competent and S-incompetent nuclei from *Tradescantia* pollen grains. *Exp. Cell Res.* 69, 29–32. doi: 10.1016/0014-4827(71)90306-5
- Wu, X., and Zhou, C. (1991). A comparative study on methods for isolation of generative cell in various angiosperm species. *Acta Biol. Exp. Sin.* 24, 15–23.
- Xu, H. P., Weterings, K., Vriezen, W., Feron, R., Xue, Y. B., Derksen, J., et al. (2002). Isolation and characterization of male-germ-cell transcripts in *Nicotiana tabacum*. *Sex. Plant Reprod.* 14, 339–346. doi: 10.1007/s00497-002-0128-6
- Yang, H. Y., and Zhou, C. (1989). Isolation of viable sperms from pollen of *Brassica napus*, *Zea mays* and *Secale cereale*. *Chin. J. Bot.* 1, 80–84.
- Zhao, X., Yang, N., and Wang, T. (2013). Comparative proteomic analysis of generative and sperm cells reveals molecular characteristics associated with sperm development and function specialization. *J. Proteome Res.* 12, 5058–5071. doi: 10.1021/pr400291p
- Zhou, C. (1988). Isolation and purification of generative cells from fresh pollen of *Vicia faba* L. *Plant Cell Rep.* 7, 107–110. doi: 10.1007/BF00270116
- Zhou, C., Zee, S. Y., and Yang, H. Y. (1990). Microtubule organization of in situ and isolated generative cells in *Zephyranthes grandiflora* Lindl. *Sex. Plant Reprod.* 3, 213–218. doi: 10.1007/BF00202877

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Lu, Wei and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Re-analysis of RNA-seq transcriptome data reveals new aspects of gene activity in Arabidopsis root hairs

Wenfeng Li^{1,2} and Ping Lan^{2*}

¹ Collaborative Innovation Center of Sustainable Forestry in Southern China of Jiangsu Province, College of Biology and the Environment, Nanjing Forestry University, Nanjing, China, ² State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing, China

OPEN ACCESS

Edited by:

Marc Libault,
University of Oklahoma, USA

Reviewed by:

Jedrzez Jakub Szymanski,
Weizmann Institute of Science, Israel
Chuang Ma,
Northwest Agricultural and Forestry
University, China

*Correspondence:

Ping Lan,
Institute of Soil Science, Chinese
Academy of Sciences, 71# East
Beijing Road, Nanjing 210008, China
plan@issas.ac.cn

Specialty section:

This article was submitted to
Plant Systems and Synthetic Biology,
a section of the journal
Frontiers in Plant Science

Received: 31 January 2015

Accepted: 25 May 2015

Published: 08 June 2015

Citation:

Li W and Lan P (2015) Re-analysis of
RNA-seq transcriptome data reveals
new aspects of gene activity in
Arabidopsis root hairs.
Front. Plant Sci. 6:421.
doi: 10.3389/fpls.2015.00421

Root hairs, tubular-shaped outgrowths from root epidermal cells, play important roles in the acquisition of nutrients and water, interaction with microbe, and in plant anchorage. As a specialized cell type, root hairs, especially in Arabidopsis, provide a pragmatic research system for various aspects of studies. Here, we re-analyzed the RNA-seq transcriptome profile of Arabidopsis root hair cells by Tophat software and used Cufflinks program to mine the differentially expressed genes. Results showed that *ERD14*, *RIN4*, *AT5G64401* were among the most abundant genes in the root hair cells; while *ATGSTU2*, *AT5G54940*, *AT4G30530* were highly expressed in non-root hair tissues. In total, 5409 genes, with a fold change greater than two-fold (FDR adjusted $P < 0.05$), showed differential expression between root hair cells and non-root hair tissues. Of which, 61 were expressed only in root hair cells. One hundred and thirty-six out of 5409 genes have been reported to be “core” root epidermal genes, which could be grouped into nine clusters according to expression patterns. Gene ontology (GO) analysis of the 5409 genes showed that processes of “response to salt stress,” “ribosome biogenesis,” “protein phosphorylation,” and “response to water deprivation” were enriched. Whereas only process of “intracellular signal transduction” was enriched in the subset of 61 genes expressed only in the root hair cells. One hundred and twenty-one unannotated transcripts were identified and 14 of which were shown to be differentially expressed between root hair cells and non-root hair tissues, with transcripts XLOC_000763, XLOC_031361, and XLOC_005665 being highly expressed in the root hair cells. The comprehensive transcriptomic analysis provides new information on root hair gene activity and sets the stage for follow-up experiments to certify the biological functions of the newly identified genes and novel transcripts in root hair cell morphogenesis.

Keywords: root hair, novel transcript, RNA-seq, co-expression, Arabidopsis

Introduction

Root hairs provide a remarkably tractable system for various aspects of studies, such as development, cell biology, and physiology, particularly in *Arabidopsis thaliana* (Dolan et al., 1998; Ryan et al., 2001; Grebe, 2012; Grierson et al., 2014). Over the last 20 years, the mechanisms underlying root hair morphogenesis have been extensively investigated, and “how and where to build a root hair” has been getting more comprehensive (Grebe, 2012; Grierson et al., 2014). The fate of epidermal cells is determined in a position-dependent manner, cells spanning the cleft of two underlying cortical cells, namely “H position,” will form hair cells; while cells presenting over a single cortical cell, called “N position,” will stay as non-hair cells (Grebe, 2012; Grierson et al., 2014). Molecular genetics studies have shown that only 0.0625% (21 out of 33,602 genes) of *Arabidopsis* genes are involved in the root cell patterning formation (Grierson et al., 2014). Among them, WEREWOLF (WER), MYB23, MYC1, TRANSPARENT TESTA GLABRA (TTG), GLABRA 3 (GL3)/ENHANCER OF GLABRA 3 (EGL3), and GL2 are critical positive regulators for non-hair cell differentiation through the inhibition of *RHD6* expression (Galway et al., 1994; Di Cristina et al., 1996; Masucci and Schiefelbein, 1996; Lee and Schiefelbein, 1999; Bernhardt et al., 2003, 2005; Kang et al., 2009; Bruex et al., 2012; Pesch et al., 2013). *GL2* itself is regulated by the regulatory complex TTG-GL3/EGL3/MYC1-WER/MYB23 (Grebe, 2012; Grierson et al., 2014). Whereas CAPRICE (CPC), TRIPTYCHON (TRY), ENHANCER OF TRY AND CPC1 (ETC1) have been proven to be positive regulators determining the cell fate of root hair (Wada et al., 1997, 2002; Schellmann et al., 2002; Kirik et al., 2004; Tominaga-Wada and Wada, 2014). In addition, some upstream genes, such as *SCRAMBLED* (*SCM*), *HISTONE DEACETYLASE 18* (*HD 18*), and *JACKDAW* (*JKD*) have been identified and well-documented as critical elements in the cell patterning (Kwak et al., 2005, 2014; Xu et al., 2005; Kwak and Schiefelbein, 2007, 2008; Hassan et al., 2010; Liu et al., 2013; Kwak et al., 2014). Although being defined as a root hair, whether a cell could finally become a root hair is relied on many internal and external factors (Grierson et al., 2014). More than 45 genes including *ROOT HAIR DEFECTIVE 6* (*RHD6*), *ROOT HAIR DEFECTIVE 2* (*RHD2*), *EXPANSIN A7* (*EXPA7*), and *EXPANSIN A18* (*EXPA18*) have been proved to be involved in root hair morphogenesis by molecular genetics studies (Grierson et al., 2014). These genes coordinately regulate the processes of Rop-GTPase re-localization and subsequently mediated signaling, vesicle trafficking, cell wall reassembly, establishment of ion gradients, reorganization of cytoskeleton (actin and microtubule), and producing and homoeostasis-maintaining of reactive oxygen species (Ishida et al., 2008; Grierson et al., 2014).

During the past 10 years, with the emergence of microarray technology coupled with advanced computational methods, vast transcriptome analyses at genome-wide level have been performed in *Arabidopsis* either by comparing transcriptional profiles of root hair-defective mutants compared to those of wild type plants, or by direct exploration of root hair-specific transcriptional profiles from root hair protoplasts based

on fluorescence-activated cell sorting (FACS) platform, with hundreds to thousands of genes being identified either involved in root hair morphogenesis or in root hair response to abiotic stresses (Jones et al., 2006; Brady et al., 2007; Dinneny et al., 2008; Gifford et al., 2008; Bruex et al., 2012; Hill et al., 2013; Kwasniewski et al., 2013; Simon et al., 2013; Becker et al., 2014; Niu et al., 2014; Tanaka et al., 2014; Wilson et al., 2015). From these omics datasets, supported by previous molecular genetics studies (Ishida et al., 2008; Grebe, 2012; Ryu et al., 2013; Grierson et al., 2014), a subset of 208 “core” epidermal genes has been identified and a gene regulatory network involved in root epidermis cell differentiation in *Arabidopsis* has been established (Bruex et al., 2012), which provides an advantage model to study the roles of both single and duplicate genes in a specific gene network (Simon et al., 2013). However, several technical limitations of microarrays, such as limited gene probes present in the chip, narrow dynamic range of gene expression changes, as well as incapability to distinguish homologous genes with high similarity, have failed to show the dynamicity and genome-wide range of transcriptional profiling of root hairs. Fortunately, next-generation sequencing technology has overcome such weaknesses and enabled us to explore whole transcriptomes at single-base resolution in a cost-effective manner. It has also enabled us to accurately quantify gene expression and identify unannotated transcripts and splicing isoforms via advanced computational methods (Trapnell et al., 2012).

In our previous study, the paired-end reads were separately matched to *Arabidopsis* genome in each biological repeat using BLAT program (Kent, 2002) and the differentially expressed genes were then identified in each replication using custom-made software RACKJ (Lan et al., 2013), which results in 1617 differentially expressed genes between root hairs (herein referred as RH) and non-root hair tissues (all root tissues except root hairs; herein referred as NRH). However, it must be noticed that although BLAT is a very effective tool for doing nucleotide alignments between mRNA and genomic DNA, it was slow and not very accurate for mapping RNA-seq reads to the genome. In addition, BLAT is not designed for the alignment of paired-end reads. RACKJ was initiated for identification of splicing isoforms. It was employed to identify differentially genes in each biological repeat via Z-score analysis. Therefore, additional information could be revealed by re-analysis of the RNA-seq data using advanced pipeline. Tophat-Cufflinks pipeline are free, open-source software tools for gene discovery and comprehensive expression analysis of RNA-seq data (Trapnell et al., 2012). Tophat was initiated specially for RNA-seq data analysis (Trapnell et al., 2009), which enables both single-end and paired-end reads to align to huge genomes using the ultra-high-throughput short read aligner Bowtie, and then analyzes the mapping results to identify splice junctions between exons (Trapnell et al., 2009, 2012). The Cufflinks pipeline contains four programs which enables us to perform not only accurate quantification of the known gene expression but also identification and quantification of any previously unannotated transcripts with or without biological repeats (Trapnell et al., 2012). In this study, we extended previous study by re-analyzing RNA-seq data using Tophat-Cufflinks

pipeline aimed to provide additional information on root hair gene activity. We revealed more than five thousands of genes that were differentially expressed between RH and NRH, with more than 4000 genes only being reported in the present study. Moreover, a subset of 14 previously unannotated transcripts was identified as to be differentially expressed between RH and NRH. The comprehensive transcriptomic analysis expands our knowledge in root hair gene activity and sets the stage for follow-up experiments on the biological functions of the newly identified genes and novel transcripts in root hair morphogenesis.

Results

Digital Information on Gene Expression in Root Hairs and Non-root Hair Tissues at Genome-wide Level

In previous study, using transgenic plants carrying *Expansin7* (*EXP7*) promoter fused to GFP as materials (Cho and Cosgrove, 2002), coupled with FACS technique, Arabidopsis root hair protoplasts were harvested and the transcriptome profiling has been explored by RNA-seq from two biological repeats. The RNA-seq data was subsequently analyzed using BLAT program (Kent, 2002) and the differentially expressed genes were further mined by custom-made software RACKJ (Lan et al., 2013). In the present study, the RNA-seq data were re-analyzed by aligning the paired-end reads to Arabidopsis Genome released in TAIR10 via Tophat program (Trapnell et al., 2009). The differentially expressed genes between RH and NRH were subsequently identified using Cufflinks pipeline (Trapnell et al., 2010, 2012). Results showed that a total of 19,743 and 19,660 genes were confidently identified (with status “OK”) in RH and NRH, respectively (Table S1 in the Supplementary Material). Of which, an overlap of 19,600 genes were expressed both in RH and NRH. In the RH, *ERD14*, *RIN4*, and *AT5G64401* were among the most abundant genes, with RPKM (Reads Per Kb per Million reads) value more than 1500 (Table 1). Among the 30 highest abundant genes, those encoding arabinogalactan proteins were the most enriched, and four of which were highly expressed in RH. Genes encoding glutathione S-transferases, dehydrins, thioredoxins, and proline-rich extension-like proteins were among the second enriched group in RH, with at least two members detected from each gene family (Table 1). Among the 30 highest abundant genes in NRH were arabinogalactan protein-encoding genes and genes encoding glutathione S-transferases and dehydrins (Table S2 in the Supplementary Material). A comparison of the 30 highest abundant genes in RH and NRH resulted in an overlap of 14 genes. Of which, genes encoding arabinogalactan proteins, glutathione S-transferases and dehydrins were the most enriched (Table S3 in the Supplementary Material).

Gene Ontology (GO) analysis of the top 30 most abundant genes in RH and NRH revealed that stress related processes involved in “cold acclimation,” “response to water deprivation,” “toxin catabolic process,” and “aluminum ion transport” were enriched both in RH and NRH (Table S4 in the Supplementary Material). Other processes of “responsive to oxidative stress,” “response to microbial phytotoxin,” “defense response to fungus,”

TABLE 1 | List of the 30 most highly expressed transcripts in root hairs.

AGI	Annotation	RPKM
AT1G76180	ERD14, Dehydrin family protein	4562.78
AT5G63270	RPM1-interacting protein 4 (RIN4) family protein	4112.99
AT5G64401	Unknown protein	3479.22
AT3G15450	Aluminum induced protein with YGL and LRDR motifs	3385.68
AT5G65207	Unknown protein	3333.24
AT1G67785	Unknown protein	3258.96
AT2G22470	AGP2, ATAGP2, arabinogalactan protein 2	3220.46
AT5G54940	Translation initiation factor SUI1 family protein	3070.01
AT5G42980	ATH3, ATTRX3, ATTRXH3, TRX3, TRXH3, thioredoxin 3	2714.9
AT1G20440	AtCOR47, COR47, RD17, cold-regulated 47	2699.7
AT2G23120	Late embryogenesis abundant protein, group 6	2371.51
AT1G17190	ATGSTU26, GSTU26, glutathione S-transferase tau 26	2344.36
AT5G40730	AGP24, ATAGP24, arabinogalactan protein 24	2027.53
AT1G20450	ERD10, LTI29, LTI45, Dehydrin family protein	2026.31
AT1G66580	RPL10C, SAG24, senescence associated gene 24	2021.75
AT2G15970	ATCOR413-PM1, cold regulated 413 plasma membrane 1	1992.6
AT5G44020	HAD superfamily, subfamily IIIB acid phosphatase	1990.44
AT3G08580	AAC1, ADP/ATP carrier 1	1968.22
AT5G20230	ATBCB, BCB, BCB, SAG14, blue-copper-binding protein	1947.38
AT5G11740	AGP15, ATAGP15, arabinogalactan protein 15	1943.79
AT1G45145	ATH5, ATTRX5, LIV1, TRX5, thioredoxin H-type 5	1928.02
AT3G54580	Proline-rich extensin-like family protein	1903.3
AT1G80920	J8, Chaperone DnaJ-domain superfamily protein	1780.59
AT2G24850	TAT, TAT3, tyrosine aminotransferase 3	1764.74
AT3G28550	Proline-rich extensin-like family protein	1748.86
AT5G64310	AGP1, ATAGP1, arabinogalactan protein 1	1739.09
AT2G29450	ATGSTU5, GSTU5, glutathione S-transferase tau 5	1729.43
AT2G40000	ATHSPRO2, HSPRO2, ortholog of sugar beet HS1 PRO-1 2	1611.9
AT3G18780	ACT2, DER1, ENL2, LSR2, actin 2	1579.69
AT5G15650	ATRGP2, RGP2, reversibly glycosylated polypeptide 2	1568.93

and “aluminum ion transport” were more enriched in RH than in NRH. In contrast, genes involved in “response to cold” and “serine–isocitratylase pathway” were more pronounced

in NRH than in RH (Table S4 in the Supplementary Material).

Differentially Expressed Genes Identified Between Root Hairs and Non-root Hair Tissues Using Cufflinks Pipeline

The differentially expressed genes between RH and NRH were identified using Cuffdiff algorithm in Cufflinks pipeline with following parameters: for a given gene, (1) the FDR (false discovery rate) adjusted p -value (that is q -value) must be less than 0.05; (2) a fold change between RH and NRH is greater than two-fold; (3) the RPKM value of each gene must be more than one in either of the samples. Subsequently, a total of 5409 genes were identified as differentially expressed genes between RH and NRH (Table S5 in the Supplementary Material). Of which, 2596 genes were significantly greater expressed in NRH than in RH; while abundance of the other 2813 genes was markedly higher in RH than in NRH (Table S6 in the Supplementary Material). Of the 2813 genes, a subset of 61 genes was only detected in RH (Table S7 in the Supplementary Material). Among the rest of 2752 (excluding 61 genes from 2813 genes), the most up-regulated genes were those encoding proline-rich (extensin-like) proteins, extensins, expansins such as EXP7 and EXP18, arabinogalactan proteins, xyloglucan endotransglucosylase/hydrolase, peroxidase superfamily protein, and others. Some other genes including *COBL9*, *COW1*, *LRX1*, *IRE*, *IRC1*, and others, which were reported to be required for or associated with root hair development and growth, were also found highly induced in RH.

Comparison of the 1617 differentially expressed genes identified in previous study (Lan et al., 2013) to the 5409 genes identified in the present study led to an overlap of 1259 genes (Table S8 in the Supplementary Material). Seventy-seven percent (1259 out of 1617) of the differentially expressed genes identified in previous study have been determined in the present analysis. By contrast, only 23% (1259 out of 5409) of the differentially expressed genes identified in the present study have been found in previous analysis, and 77% (4150 out of 5409) of the additional genes were only discovered in the present study using Tophat-Cufflinks pipeline. Several of these additional genes, such as *COBL9*, *RHS15*, and *RHS* have been reported to be associated with root hair formation; some of them were among the most up-regulated genes in RH. List of the top 100 most up-regulated genes in the present study can be found in **Table 2** for detailed information.

In addition, among the 1617 differentially expressed genes identified in previous study, a subset of 635 genes was shown up-regulated in RH. Among them, 580 genes were found up-regulated in RH in present study, i.e., 91% (580 out of 635) genes up-regulated in previous study were also identified by Tophat-Cufflinks pipeline.

Differential Go Analysis of Differentially Expressed Genes

Differential GO analysis of the 5409, 4150, and 1259 genes, which were differentially expressed genes identified in this study, newly identified in this study and identified by both present and previous study, respectively, were performed.

Results showed that processes of “response to salt stress,” “ribosome biogenesis,” “protein phosphorylation,” “embryo development ending in seed dormancy,” and “response to water deprivation” were most enriched ($P \leq 3.01E-10$) in the total 5409 differentially expressed genes (Table S9 in the Supplementary Material). In the 4150 subset, processes of “protein phosphorylation,” “embryo development ending in seed dormancy,” “response to water deprivation,” “response to chitin,” “cytokinesis,” “intracellular signal transduction,” “DNA replication,” “transport,” and “microtubule-based movement” were enriched ($P \leq 5.21E-7$). Whereas protein synthesis related processes of “ribosome biogenesis” and “translation,” and root hair related processes of root hair cell differentiation and development were underrepresented ($P \geq 0.55$). By contrast, protein synthesis and root hair related processes as well as processes of “response to salt stress,” “response to cold,” and “response to cadmium ion” were dramatically overrepresented ($P \leq 2.97E-11$) in the 1259 overlapping genes (Table S9 in the Supplementary Material).

GO analysis of the top 100 most induced genes in RH revealed that processes of “plant-type cell wall organization,” “root hair cell differentiation,” “trichoblast differentiation,” “response to oxidative stress,” “unidimensional cell growth,” “protein phosphorylation,” and “root hair cell tip growth” were enriched; while only the process of “intracellular signal transduction” was shown significantly ($P < 0.01$) in the subset of 61 genes expressed only in RH (Table S10 in the Supplementary Material).

Co-Expression Network Construction and Module Identification in RH

MACCU program (Lin et al., 2011) was used to calculate the Pearson correlation coefficients of any two genes based on the 300 root-related arrays which were manually identified as previously described, and gene pairs with a threshold value of ≥ 0.83 were selected to build co-expression networks (Lin et al., 2011). The threshold value was selected for individual co-expression network mainly based on the GO enrichment analysis of genes involved in the network (Lin et al., 2011). Briefly, first a series of threshold values from 0.7 to 0.9 were employed to select gene pairs for co-expression networks. Then, we applied a series of GO enrichment analysis of the genes corresponding to individual co-expression network and looked for the threshold with the best enrichments of GO categories ($P < 1E-03$) among the input genes. Cytoscape (<http://www.cytoscape.org>) program was applied to visualize the co-expression relationships among genes and the tool of NetworkAnalyzer was employed to extract connected components (sub-network). In the present study, we mainly focused on the up-regulated genes in RH and on finding novel modules, when compared to the previous study. To this end, the co-expression analysis of the 635 up-regulated genes previously identified was performed. Results showed that a network comprising of 122 nodes from 124 genes and 367 edges (correlations between genes) was constructed. This network can be divided into one large and 12 small components (sub-networks), with the large one consisting of 93 nodes from 94 genes and 349 edges (**Figure S1**). Using MCODE program, two modules containing 20 and nine genes were extracted from

TABLE 2 | List of the 100 most up-regulated genes in root hairs (RH) compared to non-root hair tissues (NRH).

AGI	Annotation	RH (RPKM)	NRH (RPKM)	Fold-change (log2)
AT2G24980	Proline-rich extensin-like family protein	282.55	0.145177	-10.9265
AT5G06630	proline-rich extensin-like family protein	350.094	0.263724	-10.3745
AT1G12560	ATEXP7, expansin A7	390.921	0.365196	-10.064
AT4G40090	AGP3, arabinogalactan protein 3	1002.28	1.07251	-9.86808
AT3G62680	ATPRP3, PRP3, proline-rich protein 3	419.939	0.453138	-9.85601
AT5G04960	Plant invertase/pectin methylesterase inhibitor	179.825	0.201765	-9.7997
AT3G09925	Pollen Ole e 1 allergen and extensin family protein	685.787	0.79408	-9.75426
AT4G25820	ATXTH14	758.459	0.933845	-9.66567
AT5G06640	Proline-rich extensin-like family protein	300.574	0.375472	-9.6448
AT3G54590	ATHRGP1, HRGP1, hydroxyproline-rich glycoprotein	882.994	1.10396	-9.64358
AT5G67400	RHS19, root hair specific 19	456.332	0.570824	-9.64282
AT4G13390	Proline-rich extensin-like family protein	346.903	0.436798	-9.63335
AT4G02270	RHS13, root hair specific 13	828.721	1.12794	-9.52106
AT4G00680	ADF8, actin depolymerizing factor 8	479.401	0.676228	-9.46951
AT1G62980	ATEXP18, expansin A18	208.435	0.345746	-9.23567
AT5G57530	AtXTH12	125.269	0.209592	-9.22322
AT5G35190	Proline-rich extensin-like family protein	467.479	0.808864	-9.17479
AT2G29620	Unknown protein	35.8692	0.067347	-9.05692
AT1G30870	Peroxidase superfamily protein	413.75	0.811178	-8.99453
AT1G12040	LRX1, leucine-rich repeat/extensin 1	184.771	0.367391	-8.97421
AT5G05500	Pollen Ole e 1 allergen and extensin family protein	467.413	0.954079	-8.93637
AT5G11440	CID5, IPD1, CTC-interacting domain 5	141.179	0.300056	-8.87808
AT1G48930	AtGH9C1, GH9C1, glycosyl hydrolase 9C1	212.097	0.450867	-8.87781
AT1G54970	ATPRP1, PRP1, RHS7, proline-rich protein 1	197.916	0.420925	-8.87711
AT2G41970	Protein kinase superfamily protein	232.38	0.551853	-8.71799
AT2G47540	Pollen Ole e 1 allergen and extensin family protein	213.419	0.513428	-8.69931
AT5G22410	RHS18, root hair specific 18	106.322	0.292415	-8.50621
AT2G47360	Unknown protein	62.0735	0.184463	-8.3945
AT2G30670	NAD(P)-binding Rossmann-fold superfamily protein	112.11	0.339629	-8.36674
AT3G10710	RHS12, root hair specific 12	57.5381	0.189001	-8.24998
AT2G45890	ATROPGEF4, RHS11	101.066	0.368687	-8.09869
AT5G49270	COBL9	80.4778	0.313345	-8.00469
AT5G22555	Unknown protein	196.529	0.860094	-7.83603
AT5G40860	Unknown protein	136.684	0.628081	-7.76568
AT3G54580	Proline-rich extensin-like family protein	1903.3	8.85952	-7.74706
AT4G09990	Protein of unknown function (DUF579)	189.187	0.903108	-7.7107
AT1G08090	ATNRT2.1, nitrate transporter 2:1	52.9332	0.266549	-7.63363
AT3G60330	AHA7, HA7, H(+)-ATPase 7	310.855	1.56648	-7.63257
AT2G33460	RIC1, ROP-interactive CRIB motif-containing protein 1	58.361	0.294145	-7.63233
AT4G26010	Peroxidase superfamily protein	461.842	2.38604	-7.59664
AT3G07070	Protein kinase superfamily protein	39.4694	0.207821	-7.56925
AT2G46860	AtPPa3, PPa3, pyrophosphorylase 3	111.429	0.593574	-7.55249
AT3G47040	Glycosyl hydrolase family protein	30.0456	0.168941	-7.47449
AT4G01110	Unknown protein	62.5961	0.387963	-7.33401
AT1G08990	PGSIP5, plant glycogenin-like starch initiation protein 5	62.0058	0.388489	-7.31838
AT5G51270	U-box domain-containing protein kinase family protein	31.5351	0.198535	-7.31142
AT5G58010	LRL3, LJRHL1-like 3	273.512	1.80494	-7.24351
AT3G49960	Peroxidase superfamily protein	159.373	1.06339	-7.22759
AT4G34580	COW1, SRH1	218.104	1.47609	-7.20709
AT3G51350	Eukaryotic aspartyl protease family protein	38.8901	0.274518	-7.14636
AT4G38390	RHS17, root hair specific 17	36.6331	0.260206	-7.13735

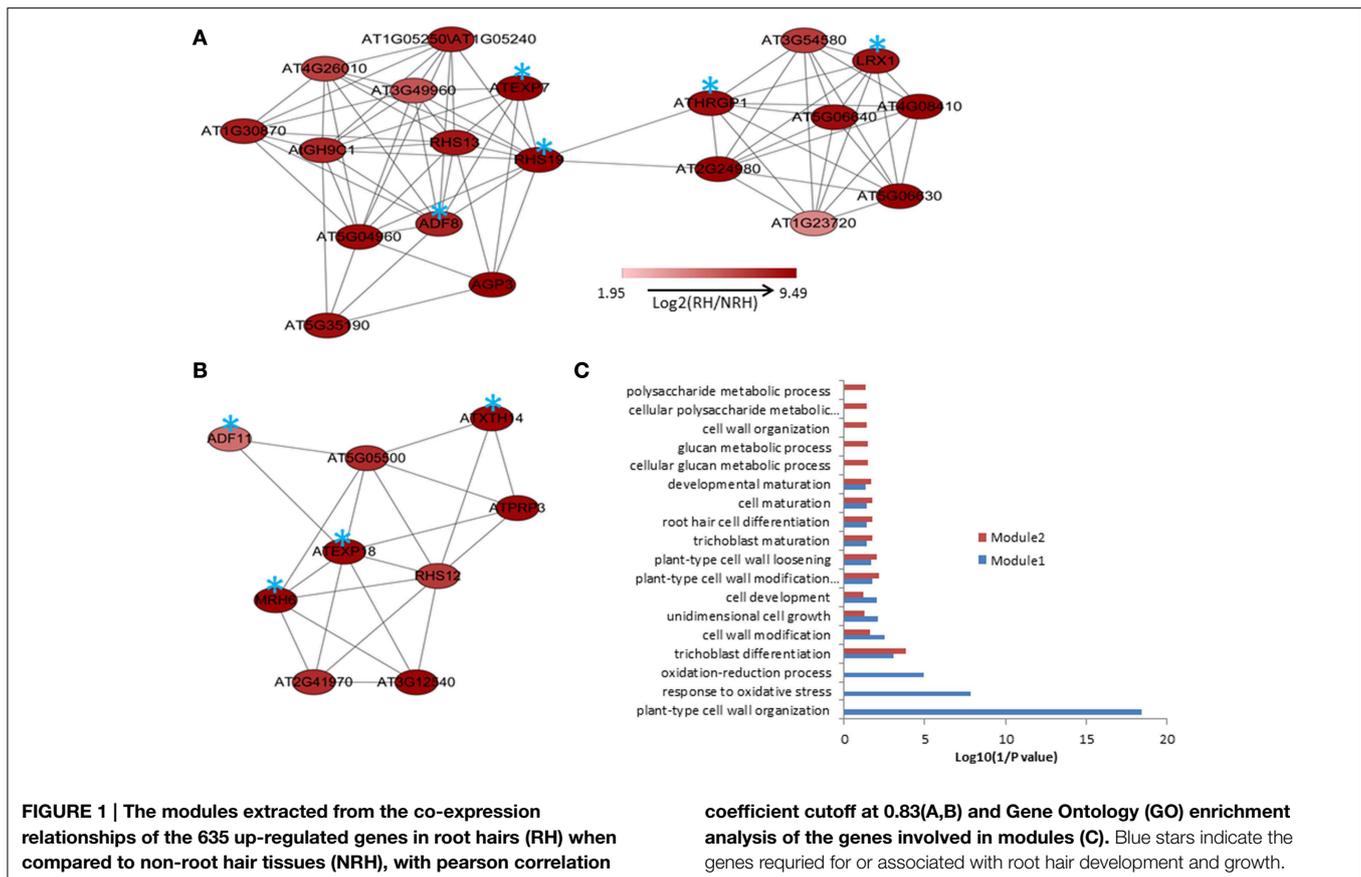
(Continued)

TABLE 2 | Continued

AGI	Annotation	RH (RPKM)	NRH (RPKM)	Fold-change (log2)
AT1G70460	RHS10, root hair specific 10	88.1011	0.641957	-7.10054
AT5G62310	IRE	64.0072	0.516494	-6.95334
AT4G29180	RHS16, root hair specific 16	55.3506	0.448309	-6.94796
AT5G17820	Peroxidase superfamily protein	1305.01	10.998	-6.89067
AT1G27740	RSL4, root hair defective 6-like 4	254.49	2.16234	-6.87888
AT1G09170	P-loop nucleoside triphosphate hydrolases	24.029	0.205787	-6.86748
AT4G30320	CAP superfamily protein	89.776	0.843779	-6.73332
AT5G49870	Mannose-binding lectin superfamily protein	31.8792	0.309214	-6.68786
AT1G51860	Leucine-rich repeat protein kinase family protein	8.1714	0.0812391	-6.65226
AT5G65160	tetratricopeptide repeat (TPR)-containing protein	74.2567	0.744421	-6.64026
AT4G02830	Unknown protein	34.9084	0.352575	-6.6295
AT5G01280	Proline-rich family protein (TAIR:AT3G09000.1)	52.2871	0.534523	-6.61206
AT5G61550	U-box domain-containing protein kinase family protein	35.9846	0.372456	-6.59417
AT4G25220	RHS15, root hair specific 15	24.1657	0.254417	-6.56962
AT2G17890	CPK16, calcium-dependent protein kinase 16	8.41945	0.0946467	-6.47503
AT2G45750	S-adenosyl-L-methionine-dependent methyltransferases	142.026	1.61829	-6.45554
AT4G25090	Riboflavin synthase-like superfamily protein	62.3366	0.714573	-6.44686
AT1G07795	Unknown protein	47.4163	0.554259	-6.41868
AT4G34380	Transducin/WD40 repeat-like superfamily protein	8.24351	0.0984524	-6.38769
AT5G25810	Tny, Integrase-type DNA-binding superfamily protein	21.4022	0.257057	-6.37953
AT2G37820	Cysteine/Histidine-rich C1 domain family protein	17.8879	0.221005	-6.33876
AT1G34760	GF14 OMICRON, GRF11, RHS5	50.7197	0.636707	-6.31577
AT4G25940	ENTH/ANTH/VHS superfamily protein	19.5987	0.246851	-6.31097
AT1G12550	D-isomer specific 2-hydroxyacid dehydrogenase	81.9584	1.03544	-6.30658
AT5G21080	Uncharacterized protein	30.2492	0.410728	-6.20257
AT5G65090	BST1, DER4, MRH3, DNase I-like superfamily protein	41.273	0.593343	-6.12019
AT4G25110	AtMC2, MC2, metacaspase 2	25.4734	0.366441	-6.11927
AT3G18450	PLAC8 family protein	26.2681	0.378257	-6.1178
AT1G01750	ADF11, actin depolymerizing factor 11	723.49	10.7564	-6.07171
AT3G21340	Leucine-rich repeat protein kinase family protein	83.7404	1.26439	-6.04941
AT2G30660	ATP-dependent caseinolytic (Clp) protease/crotonase	24.6728	0.377473	-6.0304
AT1G10385	Vps51/Vps67 (components of vesicular transport) protein	8.71123	0.133365	-6.02942
AT4G14780	Protein kinase superfamily protein	11.7601	0.181057	-6.02132
AT3G54870	ARK1, CAE1, MRH2	48.081	0.745398	-6.01131
AT3G47050	Glycosyl hydrolase family protein	10.9459	0.183671	-5.89713
AT2G38500	2-Oxoglutarate (2OG) and Fe(II)-dependent oxygenase	91.0314	1.52791	-5.89674
AT3G07900	O-fucosyltransferase family protein	32.1397	0.546227	-5.87871
AT5G15600	SP1L4, SPIRAL1-like4	148.905	2.58613	-5.84745
AT1G53680	ATGSTU28, GSTU28, glutathione S-transferase TAU 28	304.142	5.34615	-5.8301
AT5G42785	Unknown protein	80.0565	1.42305	-5.81396
AT2G20030	RING/U-box superfamily protein	8.41905	0.150287	-5.80787
AT2G03360	Glycosyltransferase family 61 protein	6.24031	0.112152	-5.79809
AT5G12050	Unknown protein	184.136	3.34175	-5.78402
AT4G22217	Arabidopsis defensin-like protein	349.124	6.59131	-5.72703
AT4G21200	ATGA2OX8, GA2OX8, gibberellin 2-oxidase 8	11.7754	0.22266	-5.72479
AT4G08450	Disease resistance protein (TIR-NBS-LRR class) family	7.34767	0.139961	-5.71419
AT1G18420	Aluminum activated malate transporter family protein	55.8597	1.08416	-5.68716
AT1G04280	P-loop containing nucleoside triphosphate hydrolases superfamily protein	192.645	3.74287	-5.68566
AT2G17590	Cysteine/Histidine-rich C1 domain family protein	4.31448	0.0844839	-5.67437

the large component, respectively (Figures 1A,B). GO analysis showed that processes of “plant-type cell wall organization,” “response to oxidative stress,” “oxidation–reduction process,”

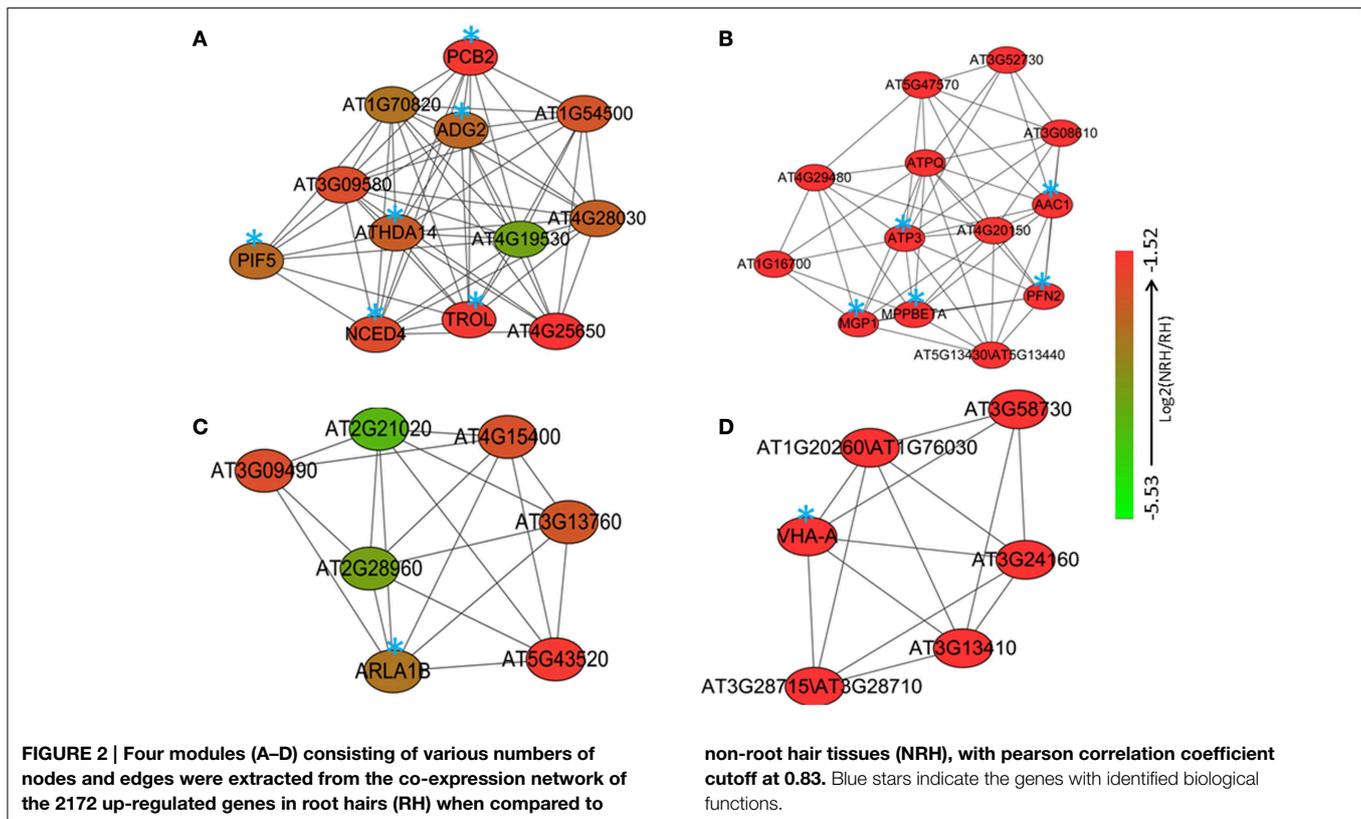
and “trichoblast differentiation” were enriched ($P < 0.001$) in module1; while only process of “trichoblast differentiation” was enriched in module 2 (Figure 1C). To know whether this



network and modules were presented in the present study, analysis of the 580 up-regulated overlapping genes showed that nearly the same co-expression network was found except that four genes were not included (nodes labeled in blue stars in **Figure S1**).

Co-expression analysis was then performed on the 2172 up-regulated genes identified only in the present study. A subset of 264 out of 2172 genes (12%) was co-expressed at the cutoff of 0.83. This network contained 260 nodes from 264 genes and 589 edges, which can be divided into five large (>10 nodes), eight middle (3–10 nodes), and 20 small (2 nodes) sub-networks (**Figure S2**). The largest sub-network contained 70 nodes from 74 genes, and the second largest sub-network contained 44 genes and the third one contained 29 genes, respectively (**Figure S2**). The other two large sub-networks contained 20 and 16 genes, respectively (**Figure S2**). GO analysis showed that processes of energy-related metabolism, such as “ATP synthesis coupled proton transport,” “photorespiration,” “response to salt stress,” “mitochondrial electron transport, ubiquinol to cytochrome c,” “response to cadmium ion,” and “proton transport” were enriched in the largest sub-network (Table S11 in the Supplementary Material). Other enriched processes were mainly related to stress responses such as cellular response to cold, cold acclimation, response to wounding, response to chitin, hyperosmotic salinity response, response to

karrikin, and response to UV-B. These enriched processes were mainly distributed in the sub-networks of 3, 4, and 7 (Table S11 in the Supplementary Material). No ($P = 1$) and low ($0.01 > P > 0.001$) enriched processes were found in the sub-networks of 9, and 2, 5, 6, and 8, respectively (Table S11 in the Supplementary Material). Four functional modules were extracted from the network, which contains 12, 13, 7, and 6 nodes and various edges, respectively (**Figure 2**). GO analysis showed that processes of “glycogen biosynthetic process,” “photosynthetic electron transport in photosystem II,” “histone deacetylation,” “red, far-red light phototransduction,” “defense response signaling pathway, resistance gene-dependent,” and “ethylene biosynthetic process” were enriched in the module 1 (**Figure 3**), and processes of “ATP synthesis coupled proton transport,” “mitochondrial electron transport, ubiquinol to cytochrome c,” “purine nucleotide transport,” “oxidation-reduction process,” and “actin polymerization or depolymerization” were enriched in the module 2, respectively. In the module 3, besides the process of “ATP synthesis coupled proton transport” which was overrepresented, other processes of “glucose mediated signaling pathway,” “Golgi organization,” and “proton transport” were enriched. While signaling related processes of “small GTPase mediated signal transduction,” “photosynthesis, light reaction,” and “intracellular signal transduction” were enriched in the module 4 (**Figure 3**).



Analysis of Root Hair Regulatory Element in the Differentially Expressed Genes

Existence of Root Hair Regulatory Element (RHE) cis-element sequence “WHHDTGNNN(N)KCACGWH” (where W = A/T, H = A/T/C, D = G/T/A, K = G/T, and N = A/T/C/G) in the 5409 differentially expressed genes was investigated as previously described (Won et al., 2009). Screening within 3000 bp upstream of the start codon (Hereafter named as -3000 bp) resulted in 201 RHE hits from 194 genes, with few genes carrying two or more RHEs (Table S12 in the Supplementary Material). Among the 201 RHE hits, RHE patterns of “AAAGTG TAGACGAT,” “ATCTGGCTTCACGTT,” and “TTCGTGAGTTTCAAATA” were relatively enriched. Subsequently, screening within introns identified 43 genes with one RHE in different intron positions (Table S13 in the Supplementary Material). Eighty nine genes were found to contain one RHE in the CDS (Encoding DNA Sequence) regions, with the sequences of “TCCATGGAAGTCACGAT,” and “TTTATGGCTGGCACGTA” being pronounced among the hits (Table S14 in the Supplementary Material). AT2G31350, encoding glyoxalase 2-5, was shown to contain two RHEs in -3000 bp region and the first intron, respectively (Table 3). Four genes AT1G18460, AT1G18470, AT2G33320, and AT3G45530 were found to harbor one RHE in -3000 bp regions and another one in the CDS regions (Table 3). Another three genes AT3G19050, AT4G03500, and AT5G27680 were shown to carry RHEs in both introns and CDS regions but not in the -3000 bp region (Table 3).

Identification of Conserved Root Epidermal Genes and Associated Co-expression Network

With the attempt to identify conserved root epidermal genes, the set of 208 “core” root epidermal genes was derived from previous report (Bruex et al., 2012), and was compared with the 5409 differentially expressed genes in this study. Comparison resulted in an overlap of 136 genes (Table S15 in the Supplementary Material), which could be grouped into nine clusters according to expression patterns (Figure S3). One hundred and twenty three out of 136 genes were annotated as hair genes, but only 27 of the 123 genes carry RHEs in their promoters (Table S15 in the Supplementary Material). To obtain the conserved root epidermal gene-specific co-expression network, the 136 genes were loaded as baits with the rest of 5409 differentially expressed genes (preys) for subsequent co-expression analysis at correlation coefficient cutoff of 0.83. The final network composing of 122 nodes (genes) and 306 edges was generated after discarding edges only linked to two preys. Of the 122 genes, 50 and 72 were from baits and preys, respectively. This network can be further divided into one large and three small clusters (Figure S4). GO analysis of the bait genes involved in the network showed that processes of “trichoblast differentiation,” “plant-type cell wall organization,” and “root hair elongation” were most enriched (Figure S4), while processes of “plant-type cell wall organization” and “oxidation–reduction process” were overrepresented in the prey genes (Figure S5). One module was extracted from the network, which contains 15 nodes and 80 edges (Figure 4A). GO analysis showed that processes of “plant-type cell wall

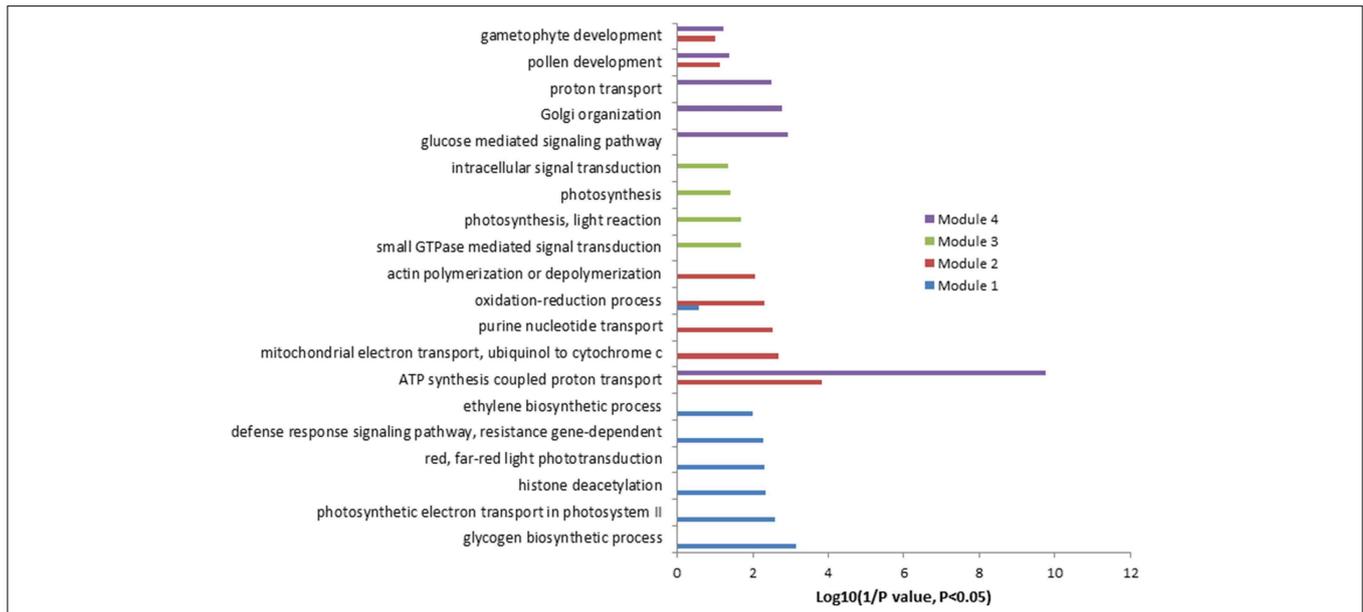
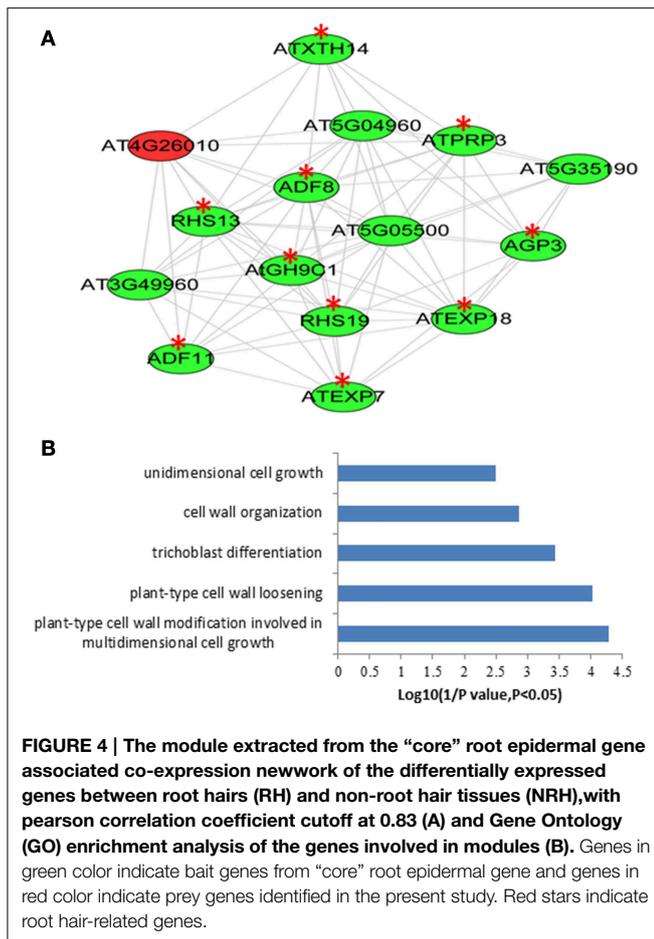


FIGURE 3 | Gene Ontology (GO) enrichment analysis of the genes involved in different modules mentioned in Figure 2.

TABLE 3 | Distribution of RHE motif in the differentially expressed genes between root hairs and non-root tissues.

AGI	Annotation	Matching Positions		Hit pattern (5'–3')	RH (RPKM)	NRH (RPKM)	Fold-change (log ₂)
		Start	End				
RHEs WITHIN 3000 bp UPSTREAM THE START CODON							
AT1G18460	Alpha/beta-hydrolases superfamily protein	2170	2154	AACGTGAACACCATGGA	142.28	60.44	–1.24
AT1G18470	Transmembrane fragile-X-F-associated protein	831	815	AACGTGAAACACATGTT	104.71	48.45	–1.11
AT2G31350	GLX2-5, glyoxalase 2-5	1950	1966	ACTATGTGGATCACGTT	163.77	15.16	–3.43
AT2G33320	Calcium-dependent lipid-binding (CaLB domain) family protein	2495	2479	AACGTGAAAAACATAGA	9.96	3.39	–1.56
AT3G45530	Cysteine/Histidine-rich C1 domain family protein	344	328	AACGTGAAAACCAAAAA	3.22	0.09	–5.1
RHEs WITHIN INTRONS							
AT2G31350.1-1	GLX2-5, glyoxalase 2-5	161	145	TACGTGATGATCATTTT	163.77	15.16	–3.43
AT3G19050.1-19	POK2, phragmoplast orienting kinesin 2	25	41	TTCTTGTGCATCACGTA	0.47	4.6	3.29
AT4G03500.1-1	Ankyrin repeat family protein	856	840	TACGTGCTAAGCAAATT	16.3	2.45	–2.73
AT5G27680.1-8	RECQSIM, RECQ helicase SIM	72	56	TACGTGCTATTCAAATT	0.18	2.15	3.59
RHEs WITHIN CDS REGIONS							
AT1G18460.1	Alpha/beta-Hydrolases superfamily protein	1435	1451	AACATGTGTTTCACGTT	142.28	60.44	–1.24
AT1G18470.1	Transmembrane Fragile-X-F-associated protein	466	482	TCCATGGTGTTCACGTT	104.71	48.45	–1.11
AT2G33320.1	Calcium-dependent lipid-binding (CaLB domain) family protein	254	270	TCCGTGATGTTCACGTT	9.96	3.39	–1.56
AT3G19050.1	POK2, phragmoplastorienting kinesin 2	2555	2571	ATTTTGAGCCGCACGAA	0.47	4.6	3.29
AT3G45530.1	Cysteine/Histidine-rich C1 domain family protein	1442	1458	TCCATGGAAGTCACGAT	3.22	0.09	–5.1
AT4G03500.1	Ankyrin repeat family protein	1723	1739	TTTATGGCTGGCACGTA	16.3	2.45	–2.73
AT5G27680.1	RECQSIM, RECQ helicase SIM	1241	1257	ATTTTGGTTCTCACGAT	0.18	2.15	3.59



modification involved in multidimensional cell growth,” “plant-type cell wall loosening,” and “trichoblast differentiation” were most enriched in this module (Figure 4B).

Identification of Unannotated Transcripts

To identify previously unannotated transcripts which are differentially expressed between RH and NRH, we first assembled a new transcript on the basis of annotated transcript reference (TAIR10_GFF3_genes_gff) using Cuffmerge algorithm in the Cufflinks pipeline (Trapnell et al., 2012). Subsequently, the differentially expressed previously unannotated transcripts were analyzed using Cuffdiff program (Trapnell et al., 2012). Results showed that a total of 121 novel transcripts were identified, and 14 out of 121 unannotated transcripts were differentially expressed between RH and NRH, with transcripts XLOC_000763, XLOC_031361, and XLOC_005665 being the most expressed genes in RH (Table 4). XLOC_005665 is of particular interest, which was highly expressed in RH (Figure 5) and deduced a small peptide with 59 amino acids.

Discussion

Root hairs in Arabidopsis have been intensively studied in various respects and close to 100 genes involved in the cell

fate determination and root hair formation have been identified, which provides numerous advantages for basic studies of development, cell biology, and physiology (Grierson et al., 2014).

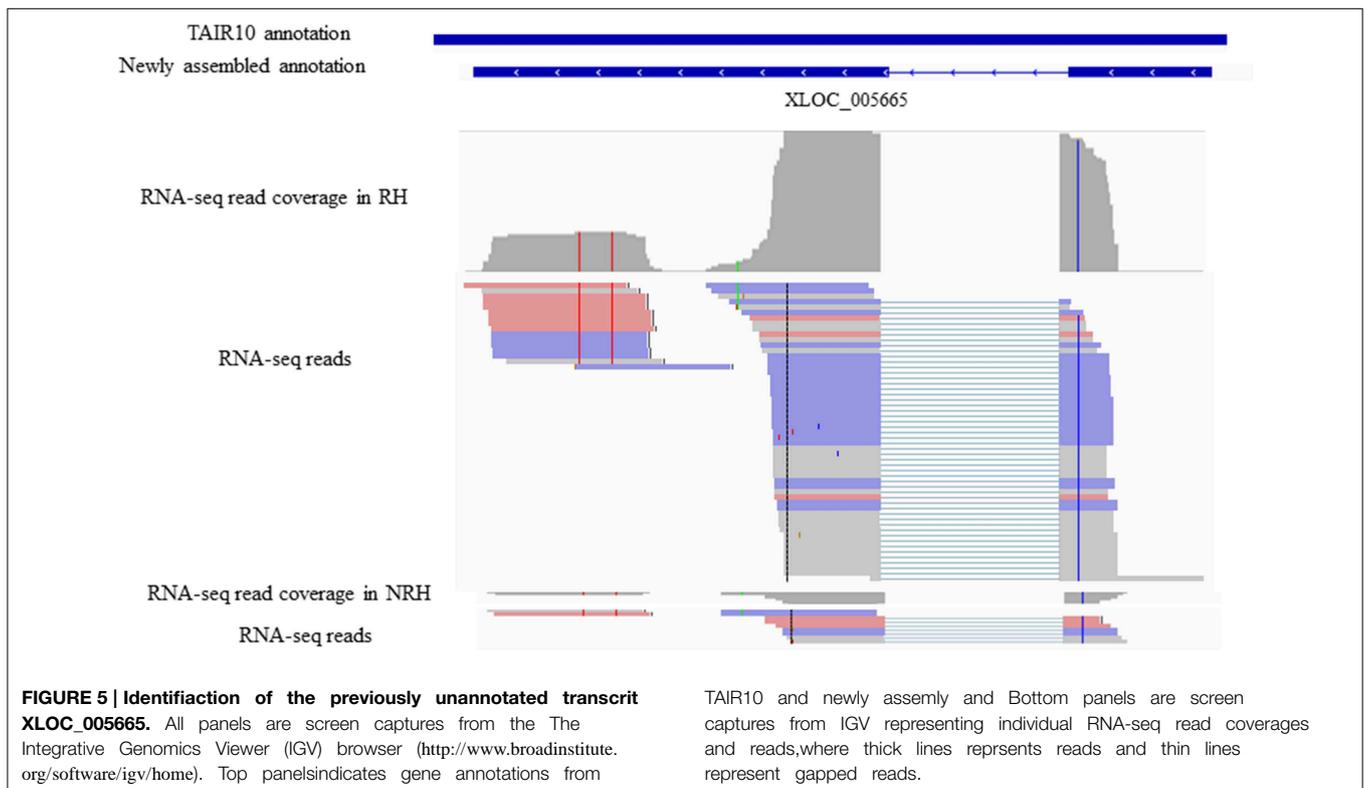
In the last decade, high-throughput transcriptome analysis, used as alternate approaches differing from traditional molecular genetic analysis, have been adopted extensively to explore genes potentially involved in root hair morphogenesis at genome-wide in Arabidopsis (Birnbaum et al., 2003; Jones et al., 2006; Brady et al., 2007; Dinneny et al., 2008; Gifford et al., 2008; Bruex et al., 2012; Lan et al., 2013). In the current study, RNA-seq data sets were re-analyzed by Tophat-Cufflinks pipeline, and several new aspects of root hair gene expression were presented. First, RNA-seq technique facilitated obtaining the global “digital” transcriptional information on root hair genes (Table S1 in the Supplementary Material). Of the 19,743 genes detected in RH, *ERD14*, *RIN4*, *AT5G64401*, and others were among the most abundant transcripts (Table 1). *ERD14* and its homologous *ERD10* were previously isolated from a cDNA library of Arabidopsis plants hydrated for 1 h and induced by ABA treatment and dehydration (Kiyosue et al., 1994). In this study, both *ERD14* and *ERD10* were shown highly expressed in RH, and were up-regulated in RH compared to NRH (Table 1 and Table S2 in the Supplementary Material). This suggests that *ERD14* and *ERD10* might be important in root hair morphogenesis or in response to abiotic stresses. *RIN4* (RPM1-interacting protein 4) is first reported to interact with *Pseudomonas syringae* type III effector or molecules, and is required for RPM1-mediated resistance in Arabidopsis (Mackey et al., 2002). Further study showed that *RIN4* can interact with *AHA1* and *AHA2* both *in vitro* and *in vivo*, thus regulating plasma membrane (PM) H(+)-ATPases activity. PM H(+)-ATPase activation/ inactivation can regulate the opening or closure of stomata, thereby controls bacterial entry into the leaf (Liu et al., 2009). *AHA2* has been reported to be a major regulator controlling the rhizosphere acidification in response to Fe deficiency (Santi and Schmidt, 2009). Taken together, it is possible that *RIN4* also plays important roles in root hair morphogenesis and response to Fe deficiency by regulating (PM) H(+)-ATPases activity mediated by *AHA2*. The third highest expressed gene in root hairs was *AT5G64401* which encodes a small peptide with unknown function (Table 1).

In previous study, a subset of 1617 genes showed differential expression between RH and NRH (Lan et al., 2013). In this study, the abundance of the 5409 genes was revealed to be changed significantly (Table S5) by Tophat-Cufflinks pipeline. Comparison of these two sets (5409 vs. 1617) resulted in an overlapped 1259 genes. We showed that additional 4150 genes were differentially expressed between RH and NRH. Genes like *COBL9* (Jones et al., 2006) and *RHS15* (Won et al., 2009; Bruex et al., 2012), which were reported to be required for or associated with root hair development and growth, were only determined in this study (Table 2). Moreover, 1/3 of cell-type patterning genes, such as *ECTOPIC ROOT HAIR2* (*ERH2/POM1*), *ECTOPIC ROOT HAIR3* (*ERH3*), *GLABRA3* (*GL3*), *ROOTHAIRLESS2* (*RHL2*), *SCRAMBLED* (*SCM/SUB*), *TRANSPARENT TESTA GLABRA2* (*TTG2*), and *WEREWOLF* (*WER*), 63% (31 out of 49) of root hair morphogenesis-related genes (Grierson et al.,

TABLE 4 | List of the 14 differentially expressed unannotated transcripts between root hairs (RH) and non-root hair tissues (NRH).

Gene_id	Locus	RH (RPKM)	NRH (RPKM)	Fold-change (log2)	q-values
XLOC_000763	Chr1:5374818–5375862	23.2677	1.33563	−4.12273	0.00083917
XLOC_007957	Chr2:565024–566158	0	1.08437	inf	0.00083917
XLOC_013179	Chr3:872193–873509	0	2.1933	inf	0.00083917
XLOC_031350	Chrchloroplast:72259–88953	0.46923	5.2598	3.48665	0.00083917
XLOC_031361	Chrmitchondria:286411–363534	386.671	71.7074	−2.43091	0.00201438
XLOC_007668	Chr1:29208757–29210196	0	1.59459	inf	0.00346354
XLOC_007865	Chr1:17233557–17234225	0	0.583177	inf	0.00427705
XLOC_017836	Chr3:12558721–12560070	0	3.49622	inf	0.00427705
XLOC_008519	Chr2:4871295–4871915	0	1.02029	inf	0.0065152
XLOC_012898	Chr2:18797715–18799154	0	0.445751	inf	0.00997437
XLOC_026701	Chr5:20540459–20541185	1.4363	7.97303	2.47277	0.0321607
XLOC_005665	Chr1:12271087–12271679	13.8534	0.950421	−3.86553	0.033759
XLOC_007866	Chr1:17292409–17293002	48.231	16.1677	−1.57685	0.0422255
XLOC_026377	Chr5:17958314–17959838	5.63137	17.0462	1.59789	0.0482593

"inf" indicates no ratio.



2014), and 45% (five out of 11) of genes related to hormone action affecting root hair development (Grierson et al., 2014) have been identified as differentially expressed genes between RH and NRH (Table S5). This study well-complements and extends the previous study by adding new information on root hair genes' numbers and activity. Several highly up-regulated genes in RH, which were not reported previously, deserve further investigation.

Co-expression analysis, which is based on the concept that genes with coordinated expression pattern under diverse

conditions are often functionally related (Eisen et al., 1998). This concept allows us to filter and select genes of unknown functions for experimental validation and functional predictions as their co-expression is related to genes of known functions (Aoki et al., 2007; Usadel et al., 2009). Not only did we identified modules from previous study (Figure 1 and Figure S1), but also revealed some new modules by the co-expression analysis of the subset of 2172 up-regulated genes in RH (Figure 2 and Figure S2). Results showed that only 12% (264 out of 2172) of the differentially expressed genes are involved in the network, and

589 relationships between genes were formed, suggesting that most of these genes are involved in diverse processes. GO analysis showed that genes associated with energy and stress related processes are enriched in the network (Table S11). This further indicates that root hair development and growth are sensitive to environmental stimuli and are energy-dependent. The conserved root epidermal genes, associated the co-expression analysis of 5409 genes, led to a network composed of 122 nodes (genes) and 306 edges (Figure S3). Unexpectedly, in the module, only one gene was from preys (Figure 4A in red color) and another 14 genes, including EXP7, EXP18, RHS12, RHS13, and RHS19, were from core root epidermal genes. (Figure 4A). Since these core genes were verified to be required for root hair development and growth, therefore it can be suggested that this prey gene plays important roles in root morphogenesis (Figure 4B). These results strongly encourage worth further investigation for those genes with unknown functions associated with the above mentioned networks.

The analysis of RHEs in the differentially expressed genes (5409) resulted in only 194 genes which carry one or two RHEs within the 3000 bp upstream of the start codon (Table S12). In an attempt to find whether such RHE localizes in other positions, we screened RHE in both introns and CDS regions. Subsets of 43 and 89 genes harboring one RHE have been hit, respectively (Tables S13, S14). Further analysis showed that only few genes carry RHE in introns and CDS, but none of them carry RHE within the three different types of positions (Table 3). Similarly, the previous study identified 154 out of 208 “core” epidermal genes in “H” position, namely root hair genes, but only 33 of them carry RHE (Bruex et al., 2012). These results suggest that regulatory elements, other than RHE, are probably involved in the transcriptional regulation of root hair gene expression.

Conclusions

In summary, using the currently popular RNA-seq analysis programs, we here provided genome-wide “digital” information on transcriptional expression of root hair genes. We detected additional 4150 genes that are differentially expressed between RH and NRH. We also identified 14 previously unannotated transcripts, which are also differentially expressed between RH and NRH. The findings in this study well-complement and extend the previous one. Some of the highly up-regulated genes in root hairs, which were not reported in the previous study, such as *RIN4* (of known function) or *AT5G64401* (of unknown function) are worth further study. Gene clustering and the root epidermal-specific co-expression analysis revealed some potentially important genes, such as *AT5G04960*, *AT4G26010*, and *AT5G05500* probably function as putative novel players in root hair morphogenesis.

Materials and Methods

Data Collection and Processing

Transcriptomic data sets were downloaded from a public database (NCBI: SRA045009.1) and analyzed as previously

described (Trapnell et al., 2009, 2010). Microarray data of 2671 ATH1 arrays from the NASCarray database (<http://affymetrix.arabidopsis.info/>) were downloaded and normalized using the RMA function of the Affy package of the Bioconductor software. Three hundred root-related arrays were manually identified as previously described (Lin et al., 2011), and were used as a database for co-expression analysis.

Mapping of RNA-seq Reads and Identification of Differentially Expressed Genes

All analyses were carried out using the Tophat-Cufflinks pipeline (Trapnell et al., 2009, 2010), with the following versions: Tophat v2.0.11, Bowtie2 v2.2.2.0, and Cufflinks v2.2.1. The Arabidopsis TAIR10 genome and gene model annotation file (GFF, TAIR10_GFF3_genes_gff) downloaded from TAIR (www.arabidopsis.org) were used as reference.

To align the RNA-seq reads to the genome, we first generated a Bowtie2 index using TAIR10 genome and then run Tophat with the following options: `-N 2 -read-gap-length 3 -read-edit-dist 3 -read-realign-edit-dist 0 -report-secondary-alignments -coverage-search -microexon-search -library-type fr-unstranded -b2-sensitive`. The resulting aligned reads were then used to create a **RABT** (Reference Annotation Based Transcript) assembly using Cufflinks. First, Cufflinks was run in the discovery mode aimed to identify previously unannotated transcripts. Assemblies both from RH and NRH were then merged into one file using Cuffmerge, using TAIR10_GFF3_genes_gff file as the reference annotation, resulting in a **RABT** assembly, used to quantify transcript abundance. Finally, transcript abundance (RPKM) and identification of differentially expressed genes was performed using Cuffdiff with default parameters ($P < 0.05$ and FDR cutoff of 0.05%) with the options: `-N -u`, corresponding to upper quartile normalization and multi-read-correct. Differential transcript abundance at all genes was calculated as the logarithm base-2 of the expression ratio ($RPKM_{NRH}/RPKM_{RH}$).

Gene Ontology Analysis

GO enrichment analysis using the TopGo “elim” method (Alexa et al., 2006) was based on The Gene Ontology Browsing Utility (GOBU) as previously described (Lin et al., 2006). The elim algorithm iteratively removes the genes mapped to significant terms from higher level GO terms, and thus avoids the increase of unimportant functional categories.

Generation of Co-expression Networks Using the MACCU Toolbox

Gene co-expression networks were constructed on the basis of 300 publicly available root-related microarrays using the MACCU toolbox as previous report (Lin et al., 2011), with a Pearson correlation threshold of equal to or greater than 0.83 based on the GO enrichment analysis. The generated co-expression networks were visualized by Cytoscape (<http://www.cytoscape.org>), and the Cytoscape tool of NetworkAnalyzer was employed to extract connected components (sub-network).

Module Identification of Co-expression Networks

MCODE plugin in Cytoscape software was employed to extract functional modules as previous report (Rivera et al., 2010).

First, a vertex-weighting value was calculated based on the clustering coefficient, C_i [$C_i = 2 * n/K_i * (K_i - 1)$], where K_i represents the node count of the neighborhood of node i ; and n indicates the number of edges among the K_i nodes in the neighborhood. Next, the highest weighted vertex is set as a center point, seed of the region and search node j whose weight ratio (W_j/W_{seed}) was >0.1 . Then, it filters the predicted complexes if the minimum degree of the graph is less than the given threshold and then constructs a module by deleting the searched node from the network. The top modules with a node count >5 were selected in the co-expression networks for GO enrichment analysis.

Acknowledgments

This work was funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB15030103), the Natural Science Foundation of China (31370280, 31470346), the National Science Foundation in Jiangsu Provinces (BK20141470) and Research Fund of State Key Laboratory of Soil and Sustainable Agriculture, Nanjing Institute of Soil Science, Chinese Academy of Science (Y412201446). WL is supported by the Jiangsu Specially-Appointed Professor program. PL is supported by Chinese Academy of Science through its One Hundred Talents Program. We thank Dr. Wen-Dar Lin and Jorge Rodríguez-Celma for their help in using the MACCU software. Dr. Mazen Alazem is most appreciated for English editing of the revision and we are grateful

References

- Alexa, A., Rahnenführer, J., and Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* 22, 1600–1607. doi: 10.1093/bioinformatics/btl140
- Aoki, K., Ogata, Y., and Shibata, D. (2007). Approaches for extracting practical information from gene co-expression networks in plant biology. *Plant Cell Physiol.* 48, 381–390. doi: 10.1093/pcp/pcm013
- Becker, J. D., Takeda, S., Borges, F., Dolan, L., and Feijo, J. A. (2014). Transcriptional profiling of *Arabidopsis* root hairs and pollen defines an apical cell growth signature. *BMC Plant Biol.* 14:197. doi: 10.1186/s12870-014-0197-3
- Bernhardt, C., Lee, M. M., Gonzalez, A., Zhang, F., Lloyd, A., and Schiefelbein, J. (2003). The bHLH genes GLABRA3 (GL3) and ENHANCER OF GLABRA3 (EGL3) specify epidermal cell fate in the *Arabidopsis* root. *Development* 130, 6431–6439. doi: 10.1242/dev.00880
- Bernhardt, C., Zhao, M., Gonzalez, A., Lloyd, A., and Schiefelbein, J. (2005). The bHLH genes GL3 and EGL3 participate in an intercellular regulatory circuit that controls cell patterning in the *Arabidopsis* root epidermis. *Development* 132, 291–298. doi: 10.1242/dev.01565
- Birnbaum, K., Shasha, D. E., Wang, J. Y., Jung, J. W., Lambert, G. M., Galbraith, D. W., et al. (2003). A gene expression map of the *Arabidopsis* root. *Science* 302, 1956–1960. doi: 10.1126/science.1090022
- Brady, S. M., Orlando, D. A., Lee, J. Y., et al. (2007). A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science* 318, 801–806. doi: 10.1126/science.1146265
- Bruex, A., Kainkaryam, R. M., Wiecekowsky, Y., et al. (2012). A gene regulatory network for root epidermis cell differentiation in *Arabidopsis*. *PLoS Genet.* 8:e1002446. doi: 10.1371/journal.pgen.1002446
- Cho, H. T., and Cosgrove, D. J. (2002). Regulation of root hair initiation and expansin gene expression in *Arabidopsis*. *Plant Cell* 14, 3237–3253. doi: 10.1105/tpc.006437

to two reviewers for their invaluable comments and suggestions to substantially improve the manuscript.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2015.00421/abstract>

Figure S1 | Co-expression relationships of the 635 up-regulated genes in root hairs (RH) when compared to non-root hair tissues (NRH), with pearson correlation coefficient cutoff at 0.83. Blue stars indicate the genes required for or associated with root hair development and growth.

Figure S2 | Co-expression relationships of the 2172 up-regulated genes in root hairs (RH) when compared to non-root hair tissues (NRH), with pearson correlation coefficient cutoff at 0.83.

Figure S3 | Hierarchical clustering analysis of changes in transcript abundance of 136 overlapping genes (Table S13 in the Supplementary Material) between 208 “core” root epidermal genes (Bruex et al., 2012) and 5409 differentially expressed genes in this study. Transcript abundance was defined as RPKM (Reads Per Kilobase per Million mapped reads) in the root hairs (RH) and non-root hair tissues (NRH) with two biological repeats. Color key indicates the \log_2 transformed intensity, gray color which not in the color key indicates that the number is missing.

Figure S4 | The “core” root epidermal gene associated co-expression network of the differentially expressed genes between root hairs (RH) and non-root hair tissues (NRH), with pearson correlation coefficient cutoff at 0.83. Genes in green color indicate bait genes from “core” root epidermal gene and genes in red color indicate prey genes identified in the present study.

Figure S5 | Gene Ontology (GO) enrichment analysis of the bait and prey genes involved in the “core” root epidermal gene associated co-expression network.

- Di Cristina, M., Sessa, G., Dolan, L., et al. (1996). The *Arabidopsis* Athb-10 (GLABRA2) is an HD-Zip protein required for regulation of root hair development. *Plant J.* 10, 393–402. doi: 10.1046/j.1365-313X.1996.10030393.x
- Dinneny, J. R., Long, T. A., Wang, J. Y., et al. (2008). Cell identity mediates the response of *Arabidopsis* roots to abiotic stress. *Science* 320, 942–945. doi: 10.1126/science.1153795
- Dolan, L., Linstead, P., Kidner, C., Boudonck, K., Cao, X. F., and Berger, F. (1998). Cell fate in plants. Lessons from the *Arabidopsis* root. *Symp. Soc. Exp. Biol.* 51, 11–17.
- Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U.S.A.* 95, 14863–14868. doi: 10.1073/pnas.95.25.14863
- Galway, M. E., Masucci, J. D., Lloyd, A. M., Walbot, V., Davis, R. W., and Schiefelbein, J. W. (1994). The TTG gene is required to specify epidermal-cell fate and cell patterning in the *Arabidopsis* root. *Dev. Biol.* 166, 740–754. doi: 10.1006/dbio.1994.1352
- Gifford, M. L., Dean, A., Gutierrez, R. A., Coruzzi, G. M., and Birnbaum, K. D. (2008). Cell-specific nitrogen responses mediate developmental plasticity. *Proc. Natl. Acad. Sci. U.S.A.* 105, 803–808. doi: 10.1073/pnas.0709559105
- Grebe, M. (2012). The patterning of epidermal hairs in *Arabidopsis*—updated. *Curr. Opin. Plant Biol.* 15, 31–37. doi: 10.1016/j.pbi.2011.10.010
- Grierson, C., Nielsen, E., Ketelaarc, T., and Schiefelbein, J. (2014). Root hairs. *Arabidopsis Book* 12:e0172. doi: 10.1199/tab.0172
- Hassan, H., Scheres, B., and Bilou, I. (2010). JACKDAW controls epidermal patterning in the *Arabidopsis* root meristem through a non-cell-autonomous mechanism. *Development* 137, 1523–1529. doi: 10.1242/dev.048777
- Hill, K., Porco, S., Lobet, G., et al. (2013). Root systems biology: integrative modeling across scales, from gene regulatory networks to the rhizosphere. *Plant Physiol.* 163, 1487–1503. doi: 10.1104/pp.113.227215

- Ishida, T., Kurata, T., Okada, K., and Wada, T. (2008). A genetic regulatory network in the development of trichomes and root hairs. *Annu. Rev. Plant Biol.* 59, 365–386. doi: 10.1146/annurev.arplant.59.032607.092949
- Jones, M. A., Raymond, M. J., and Smirnov, N. (2006). Analysis of the root-hair morphogenesis transcriptome reveals the molecular identity of six genes with roles in root-hair development in *Arabidopsis*. *Plant J.* 45, 83–100. doi: 10.1111/j.1365-3113X.2005.02609.x
- Kang, Y. H., Kirik, V., Hulskamp, M., et al. (2009). The MYB23 gene provides a positive feedback loop for cell fate specification in the *Arabidopsis* root epidermis. *Plant Cell* 21, 1080–1094. doi: 10.1105/tpc.108.063180
- Kent, W. J. (2002). BLAT—the BLAST-like alignment tool. *Genome Res.* 12, 656–664. doi: 10.1101/gr.229202
- Kirik, V., Simon, M., Huelskamp, M., and Schiefelbein, J. (2004). The ENHANCER OF TRY AND CPC1 gene acts redundantly with TRIPTYCHON and CAPRICE in trichome and root hair cell patterning in *Arabidopsis*. *Dev. Biol.* 268, 506–513. doi: 10.1016/j.ydbio.2003.12.037
- Kiyosue, T., Yamaguchi-Shinozaki, K., and Shinozaki, K. (1994). Characterization of two cDNAs (ERD10 and ERD14) corresponding to genes that respond rapidly to dehydration stress in *Arabidopsis thaliana*. *Plant Cell Physiol.* 35, 225–231.
- Kwak, S. H., and Schiefelbein, J. (2007). The role of the SCRAMBLED receptor-like kinase in patterning the *Arabidopsis* root epidermis. *Dev. Biol.* 302, 118–131. doi: 10.1016/j.ydbio.2006.09.009
- Kwak, S. H., and Schiefelbein, J. (2008). A feedback mechanism controlling SCRAMBLED receptor accumulation and cell-type pattern in *Arabidopsis*. *Curr. Biol.* 18, 1949–1954. doi: 10.1016/j.cub.2008.10.064
- Kwak, S. H., Shen, R., and Schiefelbein, J. (2005). Positional signaling mediated by a receptor-like kinase in *Arabidopsis*. *Science* 307, 1111–1113. doi: 10.1126/science.1105373
- Kwak, S. H., Woo, S., Lee, M. M., and Schiefelbein, J. (2014). Distinct signaling mechanisms in multiple developmental pathways by the SCRAMBLED receptor of *Arabidopsis*. *Plant Physiol.* 166, 976–987. doi: 10.1104/pp.114.247288
- Kwasniewski, M., Nowakowska, U., Szumera, J., Chwialkowska, K., and Szarejko, I. (2013). iRootHair: a comprehensive root hair genomics database. *Plant Physiol.* 161, 28–35. doi: 10.1104/pp.112.206441
- Lan, P., Li, W., Lin, W. D., Santi, S., and Schmidt, W. (2013). Mapping gene activity of *Arabidopsis* root hairs. *Genome Biol.* 14:R67. doi: 10.1186/gb-2013-14-6-r67
- Lee, M. M., and Schiefelbein, J. (1999). WEREWOLF, a MYB-related protein in *Arabidopsis*, is a position-dependent regulator of epidermal cell patterning. *Cell* 99, 473–483. doi: 10.1016/S0092-8674(00)81536-6
- Lin, W.-D., Chen, Y.-C., Ho, J.-M., and Hsiao, C.-D. (2006). GOBU: toward an integration interface for biological objects. *J. Inf. Sci. Eng.* 22, 19.
- Lin, W. D., Liao, Y. Y., Yang, T. J., Pan, C. Y., Buckhout, T. J., and Schmidt, W. (2011). Coexpression-based clustering of *Arabidopsis* root genes predicts functional modules in early phosphate deficiency signaling. *Plant Physiol.* 155, 1383–1402. doi: 10.1104/pp.110.166520
- Liu, C., Li, L. C., Chen, W. Q., Chen, X., Xu, Z. H., and Bai, S. N. (2013). HDA18 affects cell fate in *Arabidopsis* root epidermis via histone acetylation at four kinase genes. *Plant Cell* 25, 257–269. doi: 10.1105/tpc.112.107045
- Liu, J., Elmore, J. M., Fuglsang, A. T., Palmgren, M. G., Staskawicz, B. J., and Coaker, G. (2009). RIN4 functions with plasma membrane H⁺-ATPases to regulate stomatal apertures during pathogen attack. *PLoS Biol.* 7:e1000139. doi: 10.1371/journal.pbio.1000139
- Mackey, D., Holt, B. F. III, Wiig, A., and Dangel, J. L. (2002). RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in *Arabidopsis*. *Cell* 108, 743–754. doi: 10.1016/S0092-8674(02)00661-X
- Masucci, J. D., and Schiefelbein, J. W. (1996). Hormones act downstream of TTG and GL2 to promote root hair outgrowth during epidermis development in the *Arabidopsis* root. *Plant Cell* 8, 1505–1517. doi: 10.1105/tpc.8.9.1505
- Niu, Y., Chai, R., Liu, L., et al. (2014). Magnesium availability regulates the development of root hairs in *Arabidopsis thaliana* (L.) Heynh. *Plant Cell Environ.* 37, 2795–2813. doi: 10.1111/pce.12362
- Pesch, M., Schultheiss, I., Digiuni, S., Uhrig, J. F., and Hulskamp, M. (2013). Mutual control of intracellular localisation of the patterning proteins AtMYC1, GL1 and TRY/CPC in *Arabidopsis*. *Development* 140, 3456–3467. doi: 10.1242/dev.094698
- Rivera, C. G., Vakil, R., and Bader, J. S. (2010). NeMo: Network Module identification in cytoscape. *BMC Bioinformatics* 11(Suppl. 1):S61. doi: 10.1186/1471-2105-11-S1-S61
- Ryan, E., Steer, M., and Dolan, L. (2001). Cell biology and genetics of root hair formation in *Arabidopsis thaliana*. *Protoplasma* 215, 140–149. doi: 10.1007/BF01280310
- Ryu, K. H., Zheng, X., Huang, L., and Schiefelbein, J. (2013). Computational modeling of epidermal cell fate determination systems. *Curr. Opin. Plant Biol.* 16, 5–10. doi: 10.1016/j.pbi.2012.12.003
- Santi, S., and Schmidt, W. (2009). Dissecting iron deficiency-induced proton extrusion in *Arabidopsis* roots. *New Phytol.* 183, 1072–1084. doi: 10.1111/j.1469-8137.2009.02908.x
- Schellmann, S., Schnittger, A., Kirik, V., et al. (2002). TRIPTYCHON and CAPRICE mediate lateral inhibition during trichome and root hair patterning in *Arabidopsis*. *EMBO J.* 21, 5036–5046. doi: 10.1093/emboj/cdf524
- Simon, M., Bruex, A., Kainkaryam, R. M., et al. (2013). Tissue-specific profiling reveals transcriptome alterations in *Arabidopsis* mutants lacking morphological phenotypes. *Plant Cell* 25, 3175–3185. doi: 10.1105/tpc.113.115121
- Tanaka, N., Kato, M., Tomioka, R., et al. (2014). Characteristics of a root hairless line of *Arabidopsis thaliana* under physiological stresses. *J. Exp. Bot.* 65, 1497–1512. doi: 10.1093/jxb/eru014
- Tominaga-Wada, R., and Wada, T. (2014). Regulation of root hair cell differentiation by R3 MYB transcription factors in tomato and *Arabidopsis*. *Front. Plant Sci.* 5:91. doi: 10.3389/fpls.2014.00091
- Trapnell, C., Pachter, L., and Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111. doi: 10.1093/bioinformatics/btp120
- Trapnell, C., Roberts, A., Goff, L., et al. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7, 562–578. doi: 10.1038/nprot.2012.016
- Trapnell, C., Williams, B. A., Pertea, G., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi: 10.1038/nbt.1621
- Usadel, B., Obayashi, T., Mutwil, M., Giorgi, F. M., Bassel, G. W., Tanimoto, M., et al. (2009). Co-expression tools for plant biology: opportunities for hypothesis generation and caveats. *Plant Cell Environ.* 32, 1633–1651. doi: 10.1111/j.1365-3040.2009.02040.x
- Wada, T., Kurata, T., Tominaga, R., et al. (2002). Role of a positive regulator of root hair development, CAPRICE, in *Arabidopsis* root epidermal cell differentiation. *Development* 129, 5409–5419. doi: 10.1242/dev.00111
- Wada, T., Tachibana, T., Shimura, Y., and Okada, K. (1997). Epidermal cell differentiation in *Arabidopsis* determined by a Myb homolog, CPC. *Science* 277, 1113–1116. doi: 10.1126/science.277.5329.1113
- Wilson, M. H., Holman, T. J., Sorensen, I., et al. (2015). Multi-omics analysis identifies genes mediating the extension of cell walls in the *Arabidopsis thaliana* root elongation zone. *Front Cell Dev. Biol.* 3:10. doi: 10.3389/fcell.2015.00010
- Won, S. K., Lee, Y. J., Lee, H. Y., Heo, Y. K., Cho, M., and Cho, H. T. (2009). Cis-element- and transcriptome-based screening of root hair-specific genes and their functional characterization in *Arabidopsis*. *Plant Physiol.* 150, 1459–1473. doi: 10.1104/pp.109.140905
- Xu, C. R., Liu, C., Wang, Y. L., et al. (2005). Histone acetylation affects expression of cellular patterning genes in the *Arabidopsis* root epidermis. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14469–14474. doi: 10.1073/pnas.0503143102

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Li and Lan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Spatial dissection of the *Arabidopsis thaliana* transcriptional response to downy mildew using Fluorescence Activated Cell Sorting

Timothy L. R. Coker^{1,2}, Volkan Cevik^{2†}, Jim L. Beynon² and Miriam L. Gifford^{2*}

¹ Systems Biology Doctoral Training Centre, University of Warwick, Coventry, UK, ² School of Life Sciences, University of Warwick, Coventry, UK

OPEN ACCESS

Edited by:

Sixue Chen,
University of Florida, USA

Reviewed by:

Dilip Shah,
Donald Danforth Plant
Science Center, USA
Mi-Jeong Yoo,
University of Florida, USA

*Correspondence:

Miriam L. Gifford,
School of Life Sciences, University of
Warwick, Gibbet Hill Road, Coventry
CV4 7AL, UK
miriam.gifford@warwick.ac.uk

† Present Address:

Volkan Cevik,
The Sainsbury Laboratory,
Norwich, UK

Specialty section:

This article was submitted to
Plant Systems and Synthetic Biology,
a section of the journal
Frontiers in Plant Science

Received: 16 April 2015

Accepted: 29 June 2015

Published: 10 July 2015

Citation:

Coker TLR, Cevik V, Beynon JL and
Gifford ML (2015) Spatial dissection of
the *Arabidopsis thaliana*
transcriptional response to downy
mildew using Fluorescence Activated
Cell Sorting. *Front. Plant Sci.* 6:527.
doi: 10.3389/fpls.2015.00527

Changes in gene expression form a crucial part of the plant response to infection. In the last decade, whole-leaf expression profiling has played a valuable role in identifying genes and processes that contribute to the interactions between the model plant *Arabidopsis thaliana* and a diverse range of pathogens. However, with some pathogens such as downy mildew caused by the biotrophic oomycete pathogen *Hyaloperonospora arabidopsidis* (*Hpa*), whole-leaf profiling may fail to capture the complete *Arabidopsis* response encompassing responses of non-infected as well as infected cells within the leaf. Highly localized expression changes that occur in infected cells may be diluted by the comparative abundance of non-infected cells. Furthermore, local and systemic *Hpa* responses of a differing nature may become conflated. To address this we applied the technique of Fluorescence Activated Cell Sorting (FACS), typically used for analyzing plant abiotic responses, to the study of plant-pathogen interactions. We isolated haustoriated (*Hpa*-proximal) and non-haustoriated (*Hpa*-distal) cells from infected seedling samples using FACS, and measured global gene expression. When compared with an uninfected control, 278 transcripts were identified as significantly differentially expressed, the vast majority of which were differentially expressed specifically in *Hpa*-proximal cells. By comparing our data to previous, whole organ studies, we discovered many highly locally regulated genes that can be implicated as novel in the *Hpa* response, and that were uncovered for the first time using our sensitive FACS technique.

Keywords: plant-pathogen interactions, oomycete pathogens, biotrophic infection, cell type-specific transcriptomics, Fluorescence Activated Cell Sorting

Introduction

Unlike mammals, plants do not develop specialized immune cells. Instead, they rely on Pattern-Recognition Receptors (PRRs), which detect conserved molecules or motifs associated with foreign micro-organisms (Zipfel, 2014), and cytoplasmic NOD-Like Receptors (NLRs), which detect more specific pathogen-derived effectors that are delivered into the plant cell (Jones and Dangl, 2006). Perception of a pathogen by these receptors triggers a cascade of cellular signaling events, which culminate at the cell nucleus where transcriptional reprogramming occurs (Tsuda and Somssich, 2015).

Transcriptional reprogramming is a crucial part of the immune response, and this makes it a potential target for interference from pathogens. Manipulation of host gene expression may be particularly important for biotrophic pathogens, which must keep their host cells alive while effectively suppressing the immune system and extracting nutrients. A number of pathogenic effectors from *Pseudomonas syringae* and *Hyaloperonospora arabidopsidis* (*Hpa*) have been shown to localize to the host cell nucleus, or to physically interact with transcriptional machinery (Mukhtar et al., 2011; Caillaud et al., 2012, 2013). Several endogenous *Arabidopsis* genes have been shown to be involved in disease susceptibility (Lapin and Van den Ackerveken, 2013; Zeilmaker et al., 2015) and expression of these may be induced by a pathogen to aid infection. Thus, being able to understand the transcriptional response to infection is not only important to understand the mechanisms by which plants resist pathogens, but also those by which pathogens suppress the plant immune system and exploit the endogenous molecular machinery of the plant for their own gain.

The pathosystem of *Arabidopsis* and its downy mildew pathogen *Hpa* has been an invaluable model in plant pathology over the past two decades for a number of reasons (Coates and Beynon, 2010). Firstly, *Hpa* is an oomycete, making it phylogenetically distinct from the many bacterial and fungal pathogens that have received extensive study, but more closely related to the agriculturally important potato blight, *Phytophthora infestans*. Additionally, the remarkable number of *Hpa* isolates, along with the number of differentially susceptible and resistant *Arabidopsis* ecotypes, available for study has made the pathosystem a useful tool for studying gene-for-gene resistance (Holub, 2007). Following this, advancements in genomics have shifted the focus toward large-scale identification of *Hpa*'s RxLR effectors and unraveling their effects on the host (Baxter et al., 2010; Fabro et al., 2011; Mukhtar et al., 2011; Caillaud et al., 2013).

Finally, the pathosystem is perhaps the clearest example of obligate biotrophy in *Arabidopsis*. Upon landing on a leaf surface, an asexual *Hpa* conidiospore germinates and forms an appressorium to penetrate the leaf surface. As early as 1 day post-infection, *Hpa* grows intercellularly as hyphae, before forming lobe-shaped structures called haustoria in almost every cell it contacts during a compatible interaction. These haustoria are invaginations of the plant cell that, while keeping the cell membrane intact, form an intimate interface between host and pathogen that aids nutrient acquisition and the delivery of effectors. Assuming successful infection, *Hpa* completes its life cycle within around 7 days, producing both asexual spores, which are carried by the tree-like conidiophores that emerge from the stomata, and sexual oospores (Coates and Beynon, 2010).

Whereas, progress is being made in identifying the key determinants of pathogenicity in *Hpa* and their effect on the host *Arabidopsis*, this progress is limited in comparison to other pathogens such as *P. syringae*, most notably because *Hpa* cannot be genetically manipulated. Several studies have looked at transcriptional change in response to *Hpa* infection (Huibers et al., 2009; Hok et al., 2011; Wang et al., 2011a; Asai et al., 2014), but it has been suggested that many of the key transcriptional

events, which may occur exclusively in haustoriated cells, are often diluted by the comparative abundance of non-haustoriated cells when taking whole-organ samples (Huibers et al., 2009; Asai et al., 2014). Moreover, very little is known about the localization of *Arabidopsis* responses to *Hpa*, and how events which occur in haustoriated cells may differ from more systemic signaling events on a genome-wide scale. Making this distinction may be crucial in understanding how the haustorial environment influences the behavior of host cells.

In order to identify plant gene expression responses specifically in haustoriated cells, and to compare these to more systemic changes in gene expression during *Hpa* infection, we developed a method of isolating haustoriated cells from seedlings infected with the compatible *Hpa* isolate Noks1. The issue of dilution of highly localized pathogen responses has been previously overcome in the *Arabidopsis*-powdery mildew interaction in one published study, where by isolating infected cells through laser capture microdissection sensitivity of transcriptomic analysis was greatly increased (Chandran et al., 2010). Here, however, we chose to use Fluorescence Activated Cell Sorting (FACS) as it is a rapid way of isolating a large number of cells for gene expression analysis (Karve and Iyer-Pascuzzi, 2015). FACS is a flow cytometry technique that allows sorting of individual cells according to their fluorescence properties (Rogers et al., 2012), and has been a valuable tool for profiling the changing transcriptome of *Arabidopsis* roots during development at high spatial and temporal resolution (Brady et al., 2007). It has also been used extensively to characterize the cell type-specificity of root response to environmental/abiotic factors such as nitrogen content (Gifford et al., 2008) and salinity (Dinneny et al., 2008). FACS has also seen limited application to leaves (Grønlund et al., 2012) and analyzing the shoot apical meristem (Yadav et al., 2009), but has not been used before to study plant-pathogen interactions.

Here we used FACS to isolate haustoriated (*Hpa*-proximal) and non-haustoriated (*Hpa*-distal) cells from *Hpa* Noks1-inoculated *Arabidopsis* seedlings using the *Hpa*-responsive transgene *ProDMR6::GFP* at two time points. We demonstrated that the FACS-isolated cells can be used for transcriptional analysis, and identified 278 transcripts that are differentially expressed between the cell types, relative to uninfected controls or between the two time points. Included in these transcripts were many novel responses which may give us new insight into how infection-site-specific events may influence the outcome of downy mildew infection in *Arabidopsis*.

Materials and Methods

Plant Material and Growth Conditions

A 2.5 kb fragment of the *DMR6* [At5g24530, *Downy Mildew Resistant 6* (van Damme et al., 2008)] promoter was PCR-amplified from *Arabidopsis* (ecotype Col-0) using the primers proDMR-F (AAAAAGCAGGCTTACCGACTCTGTCTGAGTCTGAAGTCCCAAACCATG) and proDMR-R (CAAGAAAGCTGGGTGCCCATTTGATGTCAGAAAATTGAAGAA G), followed by a second amplification with pAttB1 (GGGGACAAGTTTGTACAAAAAAGCAGGCT) and pAttB2

(GGGGACCACTTTGTACAAGAAAGCTGGGT), and cloned into the pDONRZeo plasmid (Invitrogen). The entry clone was then recombined with the binary vector pBGWFS7 (Karimi et al., 2002). The resulting plasmid was then introduced into *Agrobacterium tumefaciens* strain GV3101. *Arabidopsis thaliana* Col-0 plants were transformed using the *Agrobacterium*-mediated floral dipping technique (Clough and Bent, 1998), and successful transformant seeds selected on BASTA. Homozygous T₃ plants with single insertions were used for all experiments. *ProDMR6::GFP* and Col-0 seeds were stratified for a minimum of 24 h before sowing onto soil, and were loosely covered with plastic film to retain moisture for the first 4 days after sowing. Plants were grown in a growth chamber (Weiss Technik, Vejle, Denmark) at 20°C with 10 h of light. The whole experiment was carried out in triplicate.

***Hyaloperonospora arabidopsidis* (Hpa) Propagation and Inoculation**

Hpa isolate Noksl (Rehmany et al., 2005) was maintained on *Arabidopsis* Col-0 by weekly transfer to 7-day-old seedlings. Inoculum was collected from seedlings and sprayed at a concentration of 30,000–60,000 spores ml⁻¹ onto new hosts according to Tomé et al. (2014). Spores were applied to 7-day-old seedlings carrying the *ProDMR6::GFP* transgene, or Col-0 wild type. These plants were then placed in water-tight propagator trays and incubated in a growth chamber (Weiss Technik, Vejle, Denmark) at 18°C with 10 h of light.

Imaging and Microscopy

Images were acquired using a Zeiss LSM 710 confocal microscope, in conjunction with the Zeiss ZEM software.

Protoplast Generation and Fluorescence Activated Cell Sorting (FACS)

Protoplasts were generated from seedling leaves according to Grønlund et al. (2012), but with the following alterations: (i) ProtectRNA and Actinomycin D were not used, (ii) vacuum infiltration was omitted, (iii) petri dishes were rotated on orbital shaker for only 45–60 min, and (iv) only one wash and centrifugation step was performed. FACS was performed according to Grønlund et al. (2012), using a workspace derived from Figure 3 of the publication. Cells were sorted directly into tubes containing 1 ml RLT cell lysis buffer (Qiagen) containing 1% β-mercaptoethanol, then samples stored at –80°C.

RNA Extraction, cDNA Amplification, and Labeling

RNA was extracted using the RNeasy Plant Mini Kit according to manufacturers instructions (Qiagen). DNase treatment was performed on-column using TURBO DNase (Life Technologies), with dose dependent on the approximate number of sorted cells in the sample, as manufacturers instructions: GFP-positive samples, which typically contained ~20,000 cells, were treated with one unit of TURBO DNase and incubated at 37°C for 20 min; all other samples, which contained >100,000 cells, were given a second equal round of DNase I treatment. cDNA was amplified using the Ovation Pico WTA System (NuGen),

then labeled with Cy3 using the One-Color DNA Labeling Kit (NimbleGen) according to manufacturers instructions. RNA integrity was measured using a 2100 Bioanalyzer Picochip (Agilent). cDNA and Cy3-labeled cDNA were quantified using a NanoDrop Spectrophotometer (Thermo Scientific).

Microarray Hybridization and Data Normalization

Labeled cDNA samples were randomized and hybridized for 18 h on a 12x135k expression array custom designed for the TAIR10 *A. thaliana* genome annotation (Design ID OID37507; see GEO GSE58046, NimbleGen), then the arrays were washed, dried and scanned according to manufacturers instructions. The scanned microarray images were imported into DEVA software, and data outputted as raw.xys files. The data were then imported into R (R Development Core Team, 2008). The Robust Multichip Average (RMA) algorithm was used to normalize the data, taking outlier probes into account, and to summarize expression at the transcript level using median polish (Irizarry et al., 2003). All raw and normalized microarray data has been deposited in GEO (GSE67100).

Microarray Data Analysis

Linear Models for Microarray Data (package *limma* in R) was used to fit linear models to pairs of samples (Figure S3), identifying genes that contrasted the most between the experimental pairs (Smyth, 2004). Transcripts were differentially expressed if they showed an absolute log₂ fold-change of ≥0.75 [a threshold previously used by Huibers et al. (2009)] and a Benjamini-Hochberg adjusted $p \leq 0.05$ in at least one comparison. Published data was processed in the same way, except for the data from Huibers et al. (2009), which had been previously normalized. The Cytoscape plugin BiNGO was used to identify gene ontology (GO) terms overrepresented in transcript groups, using the default settings and the “GO full” database, and a significance threshold of Benjamini-Hochberg adjusted $p \leq 0.05$. For grouping, transcripts found to be differentially expressed in any pairwise comparison between sample types at either 5 or 7 days post-inoculation (d.p.i.) were placed in order of their ratio of proximal change to distal change, measured as $\log_2(\text{Expression}_{\text{Proximal}}/\text{Expression}_{\text{Control}})/\log_2(\text{Expression}_{\text{Distal}}/\text{Expression}_{\text{Control}})$ and divided evenly into the final number of groups.

Results

***ProDMR6::GFP* as a Fluorescent Reporter for Host Cells Containing *Hyaloperonospora arabidopsidis* Haustoria**

In order to identify transcriptional events in *A. thaliana* that occur specifically in cells containing *Hpa* haustoria, we developed a method of using FACS to isolate haustoriated cells and non-haustoriated cells from *Hpa*-infected plants. This required a fluorescent reporter that is expressed specifically in haustoriated cells. van Damme et al. (2008) recently characterized the *Arabidopsis* gene Downy Mildew Resistant 6 (*DMR6*), which encodes a 2-oxoglutarate (2OG)-Fe(II) oxygenase and is required for susceptibility to *Hpa* isolate Waco9. By expressing

a GUS reporter under the control of the *DMR6* promoter they demonstrated that *DMR6* expression is induced specifically in haustoriated cells, in both compatible and incompatible interactions with *Hpa* (van Damme et al., 2008). In order to assess *ProDMR6* as a marker for isolating haustoriated cells using FACS, a construct containing 2.5 kb upstream of *DMR6* was fused to the GFP coding sequence and used to transform *Arabidopsis* Col-0 plants.

To investigate *ProDMR6::GFP* expression we screened 10- to 14-day-old T₃ seedlings of four independent transformants using confocal microscopy. GFP expression was observed consistently in all transformants upon inoculation with the compatible *Hpa* isolate Noks1, and all transgenic lines behaved as Col-0 in terms of growth and development. Although van Damme et al. (2008) reported expression of *ProDMR6::GUS* as early as 2 d.p.i., we observed little or no fluorescence at 3 d.p.i. (Figure 1A). Instead, we observed strong fluorescence at 5 (Figure 1B) and 7 d.p.i. (Figure 1C). Fluorescent cells were observed adjacent to each other, suggestive of the pattern of *Hpa* infection (Figure S1), and this was confirmed to correlate with the visibility of conidiophores on the cotyledon surface at 7 d.p.i. (data not shown). We did not observe green fluorescence in Noks1-infected Col-0 seedlings (Figure 1D), or uninoculated *ProDMR6::GFP* seedlings (Figure 1E), at any time point, confirming that the GFP was expressed specifically upon *Hpa* infection in the marker line.

Fluorescence Activated Cell Sorting to Isolate Haustoriated and Non-haustoriated Cells from Infected Tissues

Having isolated an effective and specific marker of *Hpa* haustoriated cells, we designed an experiment allowing us to study the transcriptional response of *Arabidopsis* to *Hpa* Noks1 on a spatial scale (Figure 2). Seven-day-old *ProDMR6::GFP* seedlings were inoculated with *Hpa* isolate Noks1 and cotyledons

sampled at 5 and 7 d.p.i. in three biological replicates. We chose 5 d.p.i., as this was when we could first observe GFP expression under the microscope, and 7 d.p.i., as it represents a point where the *Hpa* life cycle has completed (Coates and Beynon, 2010). Protoplasts were generated from these samples and cells sorted using FACS to obtain two cell populations: GFP-expressing cells, representing the haustoriated cell population and hereon referred to as “*Hpa*-proximal cells,” and non-GFP-expressing cells, representing the non-haustoriated cell population from infected plants, hereon be referred to as “*Hpa*-distal cells.” As a control and baseline for comparison, uninfected *ProDMR6::GFP* seedlings of the same age were also sampled at both time points, protoplasts generated and sorted through FACS.

Protoplasts were generated using a recent protocol for FACS of leaf cells by Grønlund et al. (2012). Immediately prior to FACS, a small subset of the protoplasts derived from infected seedlings express GFP, consistent with the proportion of GFP expressing cells in infected seedling leaves. This GFP expression was detected upon FACS analysis (Figure S2). In contrast, GFP expressing cells were not observed in protoplasts from uninfected seedlings prior to FACS. From the 18 protoplast samples collected (three cell populations × two time points × three biological replicates), RNA was extracted, converted to cDNA, labeled, and hybridized to whole genome oligonucleotide *Arabidopsis* microarrays.

Differential Expression of Genes in *Hpa*-proximal and *Hpa*-distal Cells Gives Insight Local and Systemic Responses to the Pathogen

Microarray gene expression was summarized at the transcript level and normalized using the RMA algorithm (Irizarry et al., 2003) (Table S1). In order to identify transcripts which were differentially expressed (DE) in *Hpa*-proximal cells, and to differentiate these from systemic signaling observed in cells distal to the infection site, we performed pairwise comparisons (Figure S3) across cell populations and time points using Linear Models for Microarray Data (LIMMA) (Smyth, 2004). A total of 278 transcripts were identified as differentially expressed at a cutoff of absolute log₂ fold-change ≥ 0.75 and a Benjamini-Hochberg adjusted *p*-value ≤ 0.05 in at least one pairwise comparison (Table S2).

As a confirmation that the cells isolated by FACS were those that were *Hpa*-associated, among the 278 DE transcripts was *DMR6* (At5g24530), which showed ~seven-fold upregulation in *Hpa*-proximal cells relative to uninfected control cells at 7 d.p.i. (Benjamini-Hochberg adjusted *p* = 0.035), and at 5 d.p.i. (Benjamini-Hochberg adjusted *p* = 0.061). We also observed upregulation of several other genes which have been previously implicated in the *Hpa* response, or as more general regulators of plant-pathogen interactions. These include Impaired Oomycete Susceptibility 1 (*IOS1*, At1g51800), Pathogenesis-Related 4 (*PR4*, At3g04720), Pathogen and Circadian Controlled 1 (*PCC1*, At3g22231), Flg22-induced Receptor-like Kinase 1 (*FLK1*, At2g19190) and *WRKY8* (At5g46350) (Table S2).

Of the 278 total DE transcripts, 81 and 231 transcripts were DE between the three cell types at 5 d.p.i. and 7 d.p.i. respectively, with 35 transcripts being DE over both time points (Figure 3A). 276 transcripts were DE between *Hpa*-proximal

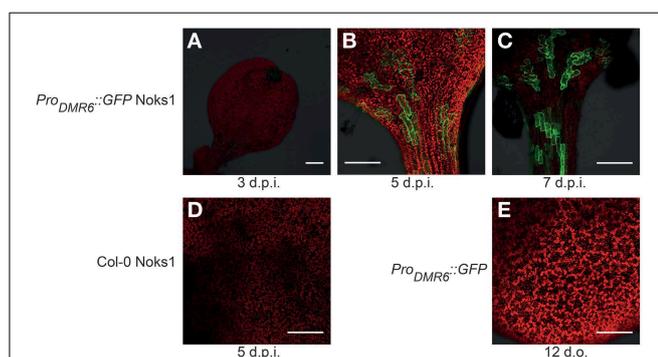
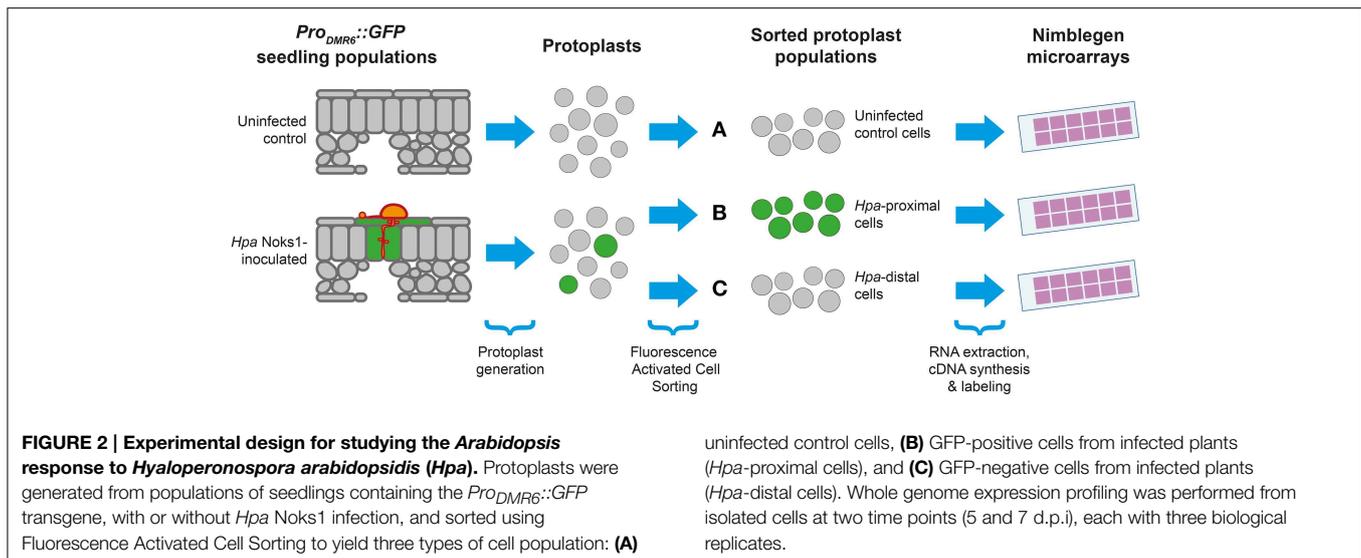


FIGURE 1 | Confocal microscopy images of *Hyaloperonospora arabidopsidis* infection marker *ProDMR6::GFP* expression in *Arabidopsis* seedlings. Scale bar represents 0.25 mm. (A) Absence of GFP expression in *ProDMR6::GFP* seedlings infected with *Hpa* isolate Noks1, 3 d.p.i. (B,C) GFP expression in *ProDMR6::GFP* seedlings infected with Noks1 at (B) 5 d.p.i. and (C) 7 d.p.i. (D) Absence of GFP expression in a Noks1-inoculated Col-0 seedling, 5 d.p.i. (E) Absence of GFP expression in uninfected *ProDMR6::GFP* transgenic seedling; seedling is 12 days old, an equivalent age to a 5 d.p.i. seedlings.



cells and uninfected control cells from the same time point, with 37 transcripts found to be DE between *Hpa*-proximal and *Hpa*-distal cells at the same time point (**Figure 3B**). A single transcript, At2g18660.1 (Plant Natriuretic Peptide A, PNP-A), was found to be DE between GFP-negative (*Hpa*-distal) cells from infected plants and cells from uninfected plants. Together with the detection of previously characterized *Hpa* responsive genes, the observation that the vast majority of transcriptional responses are being identified in the *Hpa*-proximal populations, rather than the *Hpa*-distal populations, from infected plants confirms that *Hpa*-responsive cells can be isolated using FACS.

In order to discover what types of genes are responding locally vs. systemically, i.e., specifically in *Hpa*-proximal cells vs. more generally in both *Hpa*-proximal and *Hpa*-distal cells, the 278 DE transcripts were grouped according to the localization of their response at each of the time points, and these groups were searched for overrepresentation of GO terms using the Cytoscape plugin BiNGO (Benjamini-Hochberg adjusted $p \leq 0.05$, Maere et al., 2005, **Figure 4**, **Table S3**). To take a more granular view of response location we chose to differentiate local and systemic genes based on the ratio of their *Hpa*-proximal response (\log_2 fold-change relative to uninfected control) to their *Hpa*-distal response; for a list of the genes within each group, see **Table 1**, **Table S2**.

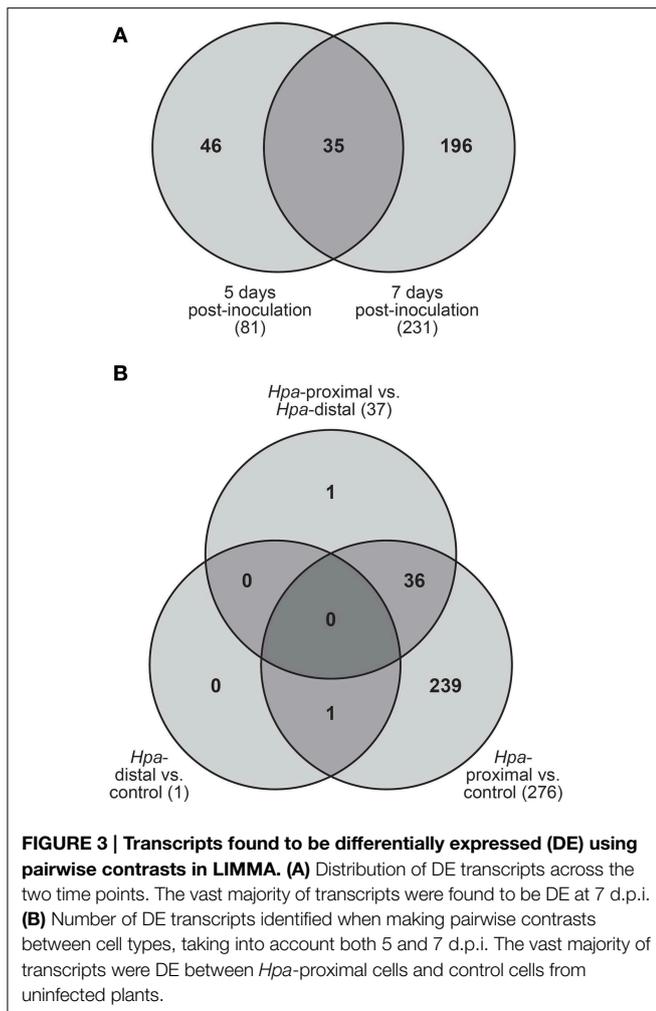
The 81 transcripts DE at 5 d.p.i. were split into three groups (**Figure 4A**). For upregulated genes, we were interested in broadly comparing local and systemic responses, so we split the transcripts found to be upregulated at this time point into two groups—one representing systemic induction (almost equal proximal and distal response), and one representing localized induction (strong proximal response, weak distal response). The systemic induction group showed overrepresentation of pathology-related GO terms such as “response to other organism” and “defense response,” as well as “systemic acquired resistance,” fitting to the systemic expression pattern of the genes

in this group. This suggests that, despite the lack of genes DE in *Hpa*-distal cells relative to the control, this population of cells is capturing systemic signaling in response to *Hpa*. Genes involved in lipid transport and localization were also overrepresented in this group. Individual genes represented in this group include Enhanced Disease Susceptibility to *Erysiphe orontii* (*EDS16*, At1g74710) and AVRPPHB Susceptible 3 (*PBS3*, At5g13320), which have both been implicated in salicylic acid accumulation in plant defense (Wildermuth et al., 2001; Nobuta et al., 2007), and Lysine Histidine Transporter 1 (*LHT1*, At5g40780), which has been shown to influence plant defense in a salicylic acid-mediated manner (Liu et al., 2010). The defense genes Pathogenesis-Related 4 (*PR4*, At3g04720) and Pathogen and Circadian Controlled 1 (*PCC1*, At3g22231) also fell into this group.

In contrast localized induction group did not show overrepresentation of any GO terms, suggesting a diversity of genes within this group. Individual genes represented in this group include the transcription factor *WRKY29* (At4g23550), a terpene synthase (*TPS4*, At1g61120) and a peroxidase superfamily protein (At5g39580). The group also includes cysteine-rich receptor-like protein kinases *ARCK1* (At1g11890) and *CRK26* (At4g38830), and a monodehydroascorbate reductase (*AtMDAR3*, At3g09940) that is crucial for colonization of *Arabidopsis* by the mutualistic fungus *Piriformospora indica* (Vadassery et al., 2009).

Due to the small number of transcripts at this time point, downregulated genes could not effectively be split into “systemic” and “local” responding and were thus considered as one group. This group showed overrepresentation for only one GO term: “cytoskeletal part.” Downregulated genes include Callose Synthase 3 (At5g13000), peroxidase 12 (At1g71695) and a pathogenesis-related thaumatin superfamily protein (At1g73620).

The larger number (231) of transcripts DE at 7 d.p.i. allowed us to split them into more groups (**Figure 4B**). Transcripts



upregulated at this time point were this time split into three gene groups—systemic induction, local induction and infection-site-specific induction, representing increasing localization of their response, such that genes in the infection-site-specific induction group showed a negligible *Hpa*-distal response. As with the systemic induction group at 5 d.p.i., the systemic induction group at 7 d.p.i. showed overrepresentation for the GO terms “lipid transport,” “systemic acquired resistance” and a number of generic defense-related terms such as “defense response.” The GO terms “response to salicylic acid stimulus” and “response to stress” were also additionally overrepresented in this group. Individual genes within this group include Pathogenesis-Related 4 (*PR4*, At3g04720) and 5 (*PR5*, At1g75040), *WRKY59* (At2g21900), *WRKY62* (At5g01900) and *WRKY8* (At5g46350), Accelerated Cell Death 6 (*ACD6*, At4g14400), Plant Natriuretic Peptide A (*PNP-A*, At2g18660) and Late Upregulated in Response to *Hyaloperonospora parasitica* (*LURP1*, At2g14560). Surprisingly, *DMR6* fell into this group, despite being used as our marker for *Hpa*-local cells. This could be due to weaker, more systemic signaling of *DMR6* that was beyond detection using a GFP marker. As this data set is enriched for responses predominantly in *Hpa*-local cells, this too may also be an

indication that even the most systemic responses captured remain fairly localized to the infection site.

The local induction group showed similar GO term enrichment to the systemic induction group at 7 d.p.i., such as the pathology-related terms “response to other organism” and “defense response” and the more generic “response to stress.” A number of receptor-like proteins were present in this group, including Flg22-induced Receptor-like Kinase 1 (*FRK1*, At2g19190), Cysteine-rich Receptor-like Kinase 13 (*CRK13*, At4g23210), Receptor Like Proteins 9 (*AtRLP9*, At1g58190) and 52 (*AtRLP52*, At5g25910) a putative CC-NBS-LRR class disease resistance protein (At1g58400) and a putative TIR-NBS class disease resistance protein (At4g09420). *WRKY47*, (At4g01720), *WRKY72* (At5g15130) and *WRKY38* (At5g22570) were also in this group.

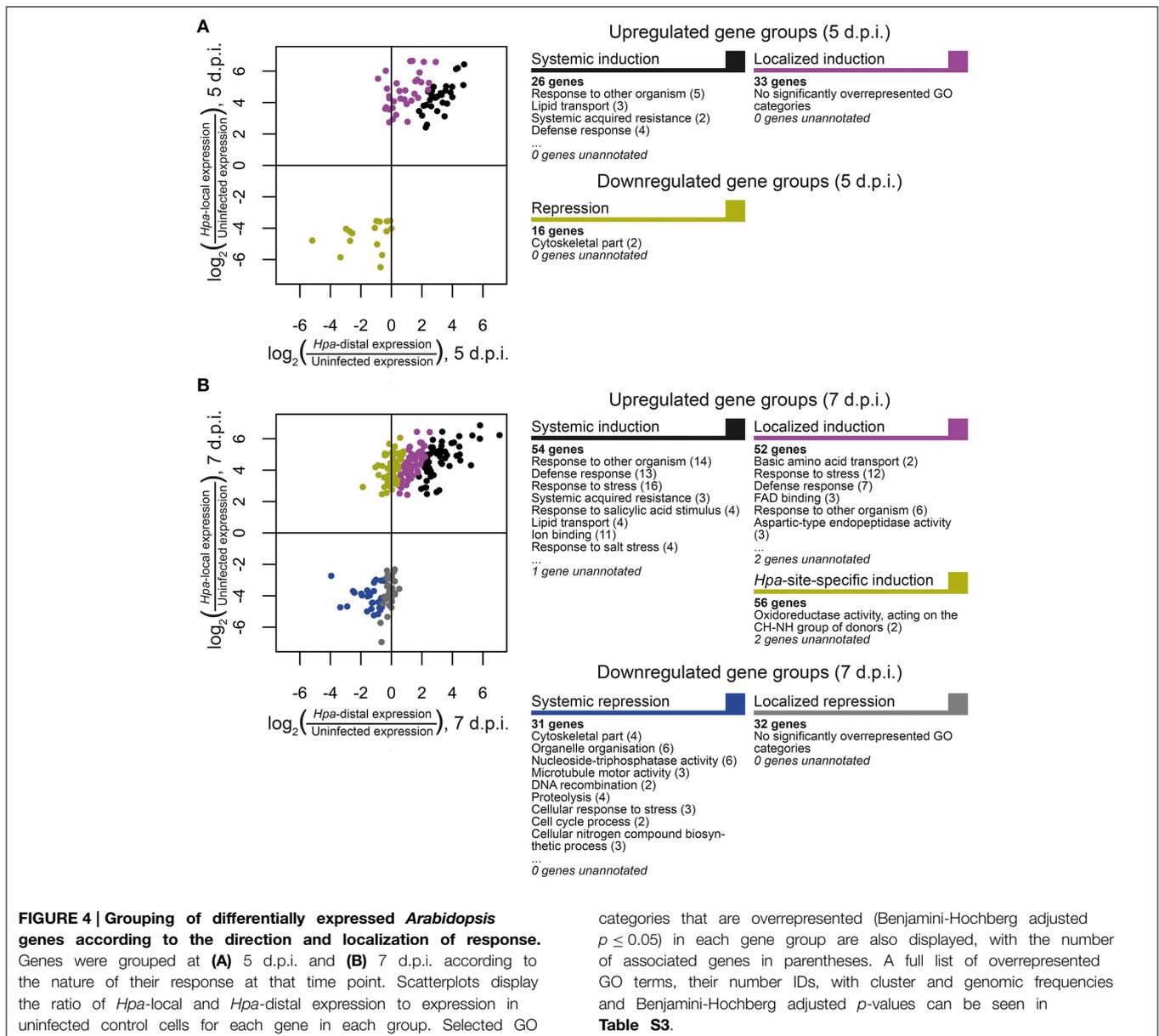
Infection-site-specific induced, representing the most localized genes upregulated at 7 d.p.i., showed overrepresentation of only the GO term “oxidoreductase activity, acting on the CH-NH group of donors.” Genes in this group include the transcription factors *WRKY36* (At1g69810), *NAC3* (At3g29035) and *NAC087* (At5g18270), as well as an RNA-binding Suppressor-of-White-APricot (SWAP) protein (At5g06520).

The larger number of downregulated genes at 7 d.p.i., relative to 5 d.p.i., allowed us to split them into two groups representing systemic and local repression. Genes that showed systemic repression were overrepresented for a number of cellular functions such as “cytoskeletal part,” “organelle organization,” “cell cycle process,” and “nucleoside-triphosphatase activity.” Genes in this group include a histone H1/H5 family member (At1g48620), metacaspase 3 (*MC3*, At5g64240) and *A. thaliana* Kinesins 1 (*ATK1*, At4g21270) and 12B (*ATK12B*, At3g23670).

Finally, there was no overrepresentation of GO terms in the localized repression group. Genes in this group included peroxidase 12 (*PER12*, At1g71695), the receptor protein kinase *ERECTA* (At2g26330), microtubule-associated protein 65-4 (*MAP65-4*, At3g60840) and Starch Synthase 3 (*ATSS3*, At1g11720).

Comparison with Published Data Sets

To ask if the FACS approach identifies novel genes in the *Arabidopsis* response to *Hpa* infection, we compared our list of differentially expressed genes to previously published microarray data from Huibers et al. (2009), Wang et al. (2011a) and Hok et al. (2011). Data from these publications was retrieved from the relevant public databases and processed in a similar manner to the data we present here, i.e., differentially expressed genes identified by making pairwise contrasts in LIMMA. From each published dataset we considered only samples and direct comparisons that were most relevant to our experimental design here. Huibers et al. (2009) used two-color CATMA arrays to profile expression in a compatible *Arabidopsis*-*Hpa* interaction (Landsberg erecta (Ler) and Cala2) and an incompatible interaction (Ler and Waco9), relative to uninfected controls, at 3 d.p.i. Wang et al. (2011a) performed a 6-day timecourse of infection with the incompatible strain Emwa1, in Col-0 and the susceptible mutant *rpp4*. Finally, Hok et al. (2011) measured gene expression in *Arabidopsis* Wassilewskija (WS) seedlings



after mock treatment, and treatment with the compatible isolate Emwa, at an early time point (8 and 24 h post-inoculation) and at a late time point (4 and 6 d.p.i.). For the former two datasets, we considered only the Cala2 interaction and the *rpp4* interaction, respectively, as they represented compatible interactions that result in a similar outcome to the Col-0 and Noks1 interaction, i.e., completion of the *Hpa* lifecycle. For the latter two datasets, which have multiple time points, we considered all time points as to capture as much of the *Hpa* response as possible.

Our 278 differentially expressed transcripts represent 267 different genes—128 of which could be detected in the previously published datasets based on our analysis (Figure 5A). The remaining 139 genes are thus novel *Hpa* responses identified by our FACS-based cell response type specific

approach. However, ~5300 transcripts were previously detected as differentially expressed in one or more of the datasets outlined above, but not differentially expressed in our dataset. A comparison between previous datasets shows that only a small proportion of these are common between datasets (Figure 5B), suggesting that these DE genes arose as differences in experimental design, *Hpa* strain used or otherwise may be false positives.

In order to compare the sensitivity and specificity of our approach to the previously published data, we compared the average fold-change and Benjamini-Hochberg adjusted p -values for all genes for a number of pairwise comparisons across different datasets (Figures 5C–F). We found that our dataset had a larger proportion of genes with significant (≥ 0.75) \log_2

TABLE 1 | Differentially expressed genes grouped according to the direction and localization of their response at 5 and 7 d.p.i.

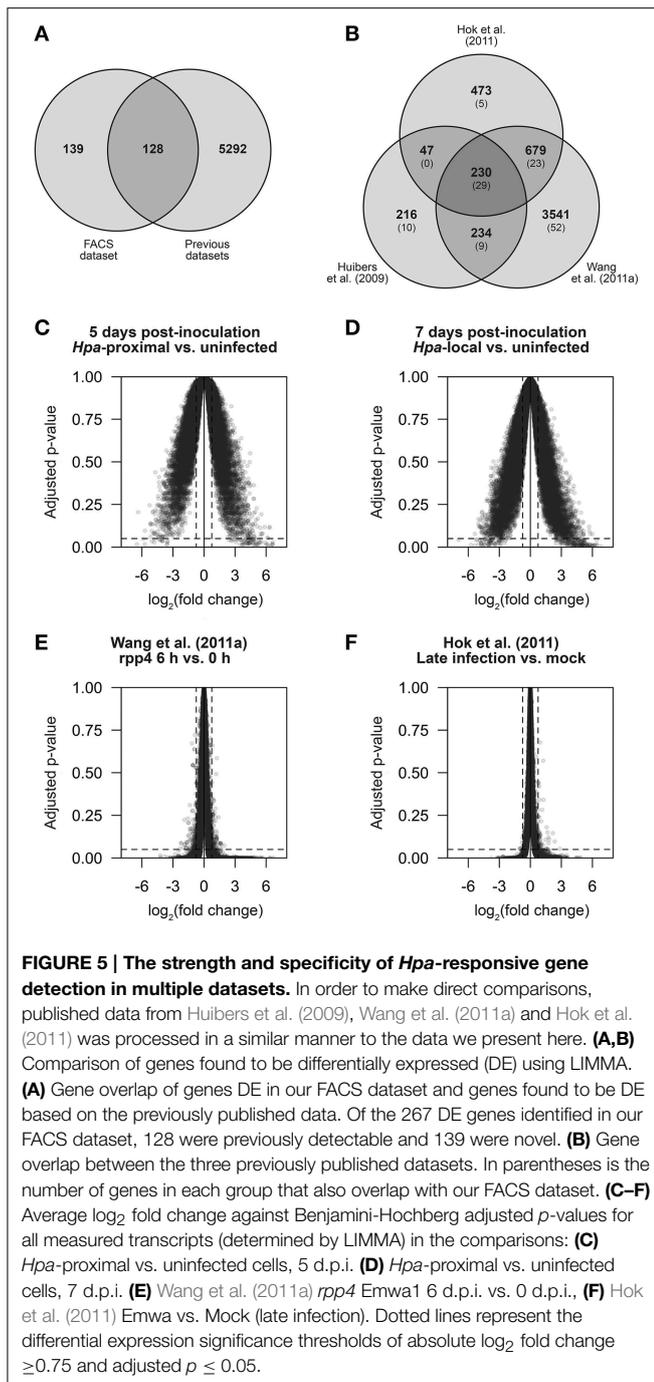
Groups at 5 d.p.i.							
SYSTEMIC INDUCTION (26 GENES)				LOCALIZED INDUCTION (33 GENES)			
ID	Name	ID	Name	ID	Name	ID	Name
AT1G02850	BGLU11	AT3G49210		AT1G06770	DRIP1	AT3G09940	ATMDAR3,
AT1G14880	PCR1	AT3G49620	DIN11	AT1G25390		AT3G11340	
AT1G35710		AT3G53600		AT1G27020		AT3G14225	GLIP4
AT1G49050		AT4G11890		AT1G30730		AT3G22600	
AT1G73805		AT4G12480	pEARLI 1	AT1G35260	MPL165	AT3G25655	IDL1
AT1G74710	EDS16, ATICS1	AT4G12500		AT1G44130		AT3G60120	BGLU27
AT2G25510		AT4G20000		AT1G51800		AT4G11890	
AT2G27660		AT4G21850	MSRB9	AT1G53470	MSL4	AT4G23550	WRKY29
AT3G04720	PR4, HEL	AT4G23150	CRK7	AT1G61120	TPS04, GES	AT4G25950	VATG3
AT3G22231	PCC1	AT5G03350		AT1G69930	GSTU11	AT4G38830	CRK26
AT3G22235		AT5G13320	ATGH3.12	AT1G74010		AT5G12340	
AT3G25610		AT5G40780		AT2G19500	CKX2, ATCKX2	AT5G22540	
AT3G29034		AT5G44568		AT2G20805		AT5G37490	
REPRESSION (16 GENES)				AT2G28110	FRA8, IRX7	AT5G38900	
ID	Name	ID	Name	AT2G31990		AT5G39580	
AT1G05910		AT3G28460		AT2G36810		AT5G48657	
AT1G71050	HIPP20	AT3G52770	ZPR3	AT3G02240	RGF7		
AT1G71695		AT4G00400	GPAT8,				
AT1G73620		AT5G02770					
AT2G22400		AT5G13000	gsl12				
AT2G32860	BGLU33	AT5G17410					
AT3G18160	PEX3-1	AT5G40640					
AT3G19960	ATM1	AT5G58240	FHIT				
Groups at 7 d.p.i.							
SYSTEMIC INDUCTION (54 GENES)				LOCALIZED INDUCTION (52 GENES)			
ID	Name	ID	Name	ID	Name	ID	Name
AT1G01680	PUB54	AT3G29130		AT1G02850	BGLU11	AT3G21710	
AT1G02520	PGP11	AT3G47480		AT1G05260	RCI3, RCI3A	AT3G46616	
AT1G02850	BGLU11	AT3G61280		AT1G26420		AT3G52710	
AT1G14880	PCR1	AT4G01350		AT1G30720		AT3G61390	
AT1G19610	LCR78, PDF1.4	AT4G12480	pEARLI 1	AT1G30730		AT4G01720	WRKY47
AT1G33960	AIG1	AT4G12490		AT1G34460	CYCB1;5, CYC3	AT4G09420	
AT1G35710		AT4G12500		AT1G34670	MYB93	AT4G16563	
AT1G55790		AT4G14400	ACD6	AT1G44130		AT4G18540	
AT1G66280	BGLU22	AT4G16260		AT1G53830	PME2	AT4G21120	AAT1, CAT1
AT1G73805		AT4G20000		AT1G56550	RXGT1	AT4G23210	CRK13
AT1G74710	EDS16, ICS1, SID2	AT4G20110	VSR7, VSR3;1, BP80-3;1	AT1G58190	RLP9	AT4G38830	CRK26
AT1G75040	PR5, PR-5	AT4G23150	CRK7	AT1G58400		AT5G02230	
AT2G14560	LURP1	AT4G39830		AT1G64583		AT5G15130	WRKY72
AT2G18660	PNP-A	AT5G01900	WRKY62	AT1G65090		AT5G18780	
AT2G21900	WRKY59, ATWRKY59	AT5G10760		AT1G66920		AT5G21280	
AT2G44380		AT5G11210	GLR2.5	AT1G69930	GSTU11	AT5G22570	WRKY38
AT2G44890	CYP704A1	AT5G13320	PBS3, GDG1, WIN3, GH3.12	AT1G71910		AT5G25260	
AT2G45510	CYP704A2	AT5G22540		AT1G77380	AAP3	AT5G25910	RLP52

(Continued)

TABLE 1 | Continued

SYSTEMIC INDUCTION (54 GENES)				LOCALIZED INDUCTION (52 GENES)			
ID	Name	ID	Name	ID	Name	ID	Name
AT3G04720	PR4, HEL	AT5G24530	DMR6	AT2G19190	FRK1	AT5G37415	AGL105
AT3G09940	MDHAR, MDAR3,	AT5G37490		AT2G27180		AT5G37540	
AT3G11340		AT5G37600	GLN1;1, GSR 1	AT2G28110	FRA8, IRX7	AT5G38540	
AT3G18250		AT5G38550		AT2G30550		AT5G48290	
AT3G21080		AT5G38900		AT2G35980	YLS9, NHL10	AT5G50200	WR3, NRT3.1
AT3G22235		AT5G39580		AT2G45220		AT5G57450	XRCC3
AT3G22600		AT5G44585		AT2G47550		AT5G59930	
AT3G26210	CYP71B23	AT5G44920		AT3G09940	MDHAR MDAR3, MDAR2	AT5G61640	PMSR1, ATMSRA1
AT3G29034		AT5G46350	WRKY8				
Hpa-SITE-SPECIFIC INDUCTION (56 GENES)							
ID	Name	ID	Name	ID	Name	ID	Name
AT1G01150		AT1G76370		AT3G50190		AT4G37710	
AT1G05880	ARI12	AT2G09840		AT3G54730		AT5G06520	
AT1G13480		AT2G19500	CKX2	AT3G55150	EXO70H1	AT5G07610	
AT1G15640		AT2G21550		AT3G55700		AT5G11400	
AT1G17020	SRG1	AT2G30395	OFFP17	AT3G60120	BGLU27	AT5G18270	ANAC087
AT1G21360	GLTP2	AT2G35770	scpl28	AT3G61827		AT5G19270	
AT1G29600		AT2G38365		AT3G62640		AT5G20330	BETAG4
AT1G51915		AT2G43730		AT4G01750	RGXT2	AT5G24080	
AT1G53980		AT3G01420	ALPHA-DOX1	AT4G03950		AT5G28190	
AT1G60095		AT3G06260	GATL4	AT4G04775		AT5G28235	
AT1G61750		AT3G14225	GLIP4	AT4G14630	GLP9	AT5G39560	
AT1G63245	CLE14	AT3G15340	PPI2	AT4G15417	RTL1	AT5G42120	
AT1G68630		AT3G26470		AT4G19950		AT5G61160	AACT1
AT1G69810	WRKY36	AT3G29035	ANAC059, NAC3	AT4G20470		AT5G63225	
SYSTEMIC REPRESSION (31 GENES)				LOCALIZED REPRESSION (32 GENES)			
ID	Name	ID	Name	ID	Name	ID	Name
AT1G10930	ATSGS1, RECQ4A	AT3G04460	PEX12, APM4	AT1G07320	RPL4	AT2G36490	DML1, ROS1
AT1G12244		AT3G04850		AT1G11720	ATSS3, SS3	AT3G05730	
AT1G16350		AT3G23670	PAKRP1L, KINESIN-12B	AT1G12845		AT3G15353	MT3
AT1G21440		AT3G28460		AT1G13380		AT3G19450	ATCAD4,
AT1G48620	HON5	AT4G11990		AT1G14690	MAP65-7	AT3G45850	
AT1G48650		AT4G21270	ATK1, KATAP	AT1G27385		AT3G54190	
AT1G54820		AT4G22930	PYR4, DHOASE	AT1G35780		AT3G60840	MAP65-4
AT1G58060		AT4G30610	BRS1, SCPL24	AT1G48600	ATPMEAMT	AT4G37080	
AT1G63630		AT4G34210	ASK11, SK11	AT1G53560		AT5G01015	
AT1G66510		AT4G37110		AT1G70370	PG2	AT5G15310	ATMYB16, ATMIXTA,
AT2G21380		AT5G17220	GST26, TT19, GSTF12	AT1G71695		AT5G15740	
AT2G22330	CYP79B3	AT5G17410		AT1G79200		AT5G20630	ATGER3,
AT2G26680		AT5G46390		AT1G79280	NUA, AtTPR	AT5G39790	
AT2G40760		AT5G51350		AT2G26330	ER, QRP1	AT5G48600	ATCAP-C
AT2G45440	DHDPS2	AT5G64240	MC3	AT2G30540		AT5G50740	
AT3G02900					AT2G32880		AT5G57130

At 5 and 7 d.p.i., transcripts were classified as either "upregulated" or "downregulated," then split into groups according to the ratio of a Hpa-local response (measured as the log₂ fold-change between Hpa-local cells and uninfected cells at that time point) and a Hpa-distal response (the fold-change between Hpa-distal and uninfected cells). See **Figure 4** for a graphical representation of their expression pattern, **Table S2** for expression values, and **Table S3** for GO term analysis details.



fold changes relative to an uninfected control than in the previously published datasets, and those identified as DE showed DE of a higher magnitude, highlighting that by specifically analyzing *Hpa*-proximal cells, we observe greater sensitivity in expression changes during infection (compare the width of plots in **Figures 5C,D** to **Figures 5E,F**). Conversely, the published datasets had a larger proportion of genes within the significance threshold of adjusted $p \leq 0.05$, but with almost-zero fold-changes (**Figures 5E,F**). This suggests that, relative to the published datasets, although our data shows higher sensitivity,

in this instance noise may be a limiting factor in *Hpa*-responsive gene detection.

Discussion

Here we present the novel use of FACS to isolate *A. thaliana* cells infected by the downy mildew pathogen *Hpa*. To our knowledge, this is the first use of FACS to specifically isolate plant cells responding to infection, although this has previously been achieved in animal systems (Richman et al., 2002; Thöne et al., 2007).

We demonstrate that cells isolated by FACS of *Hpa*-infected seedlings can be used for transcriptomic analysis of the local vs. systemic response to *Hpa* infection. Consistent with expectations that the majority of transcriptional events would occur at the infection site, all differentially expressed genes were either significantly upregulated or downregulated in the *Hpa*-proximal cell population, over time, or relative to an uninfected control or *Hpa*-distal cells from infection plants at the same time point. In contrast, only a single transcript showed significant differential expression between *Hpa*-distal cells and uninfected control cells. The identity of this transcript as Plant Natriuretic Peptide A (*PNP-A*, At2g18660) is assuring as PNP-A has been previously described as a secreted signal working systemically during both abiotic and biotic stress (Wang et al., 2011b). Ideally we would have identified further genes to be significantly differentially expressed in the *Hpa*-distal population, representing systemic signaling. However, as the *Hpa*-distal cell population was simply a collection of cells not expressing the haustoriated cell marker *ProDMR6::GFP*, we might expect this population to be heterogeneous, containing cells at varying proximity to the pathogen, many of which may not be responding to the pathogen at all. To address this potential dilution of systemic responses, we considered that many of the genes differentially expressed in *Hpa*-proximal cells may also be responding more systemically, and grouped these into the “Systemic Induction” and “Systemic Repression” groups in **Figure 4**. Several of the genes and GO terms associated with these groups are consistent with what is already known about defense signaling in *Arabidopsis*, such as the role of salicylic acid and salicylic acid-responsive gene expression in systemic acquired resistance (Durrant and Dong, 2004). However, no firm conclusions can currently be made from the analysis in **Figure 4** and further experiments are needed to validate the localization of these responses, and to unravel their significance in the *Hpa*-*Arabidopsis* interaction.

The use of FACS to study cells specifically at the site of infection has potential to increase the sensitivity of transcriptomic or other high-throughput analyses, such as proteomics. We have shown that in general, the magnitude of up- or down-regulation of genes is greater in our FACS-isolated *Hpa*-proximal cells than in previous whole-leaf datasets, relative to uninfected controls (**Figures 5C–F**). We have also identified a number of genes that are differentially expressed in *Hpa*-proximal cells not previously detected in microarray studies (**Figure 5A**). However, we have also failed to detect many genes previously associated with *Hpa* infection. While many of these could potentially be attributed to differences

in experimental design or the *Hpa* isolate used, it seems that noise is largely a contributing factor. Greater optimization of the FACS protocol will hopefully help to overcome this in the future.

A crucial development in the use of FACS for studying local vs. systemic signaling during *Arabidopsis* infection will be the development of new cell markers. A key challenge, particularly for the *Hpa* pathosystem, is that the pathogen and the proteins that it delivers into host cells cannot currently be fluorescently labeled through genetic manipulation. As such, isolation of *Hpa*-contacting cells relies entirely on pathogen-responsive *Arabidopsis* promoters, which may not be induced immediately and are likely to show changes in expression over the course of infection. This seems to be an issue with the *DMR6* promoter, from which we could not detect GFP expression until 5 d.p.i. This prevented us from studying earlier stages of infection, which is unfortunate as it is at these stages that the use of FACS will be most informative, as the limited spread of the pathogen precludes the use of whole tissue microarrays. An additional caveat more relevant to this dataset, is that, at the later time points (e.g., 5 and 7 d.p.i.), recently haustoriated cells may not fluoresce, and may instead be interpreted as *Hpa*-distal cells. Characterizations of new, early-induced haustoriated cell markers, as well as an in-depth study of their expression patterns will be crucial in developing a refined FACS approach. Furthermore, to avoid dilution of the systemic response, one could use a second fluorophore to mark cells within a certain range of the pathogen. This could potentially be complex as signals and responses spread over space. In addition to developing new methods to study pathogen signaling at a cell-specific resolution, we must in turn develop theoretical methods to understand the data being generated, and perhaps take into account some of the assumptions and limitations of the FACS approach. As these methods develop, we can better understand the events that occur specifically at the *Arabidopsis*-*Hpa* interface, and how these might influence more widespread signaling in the plant.

Author Contributions

TC, MG, and JB contributed to design of the experiment and interpretation of the data. VC generated the *ProDMR6::GFP* construct and plant lines, and TC and VC performed microscopy of these plants. Protoplast generation, FACS and microarray analysis was performed by TC. All authors wrote the manuscript.

Acknowledgments

We thank D. Tomé and J. Steinbrenner for maintenance of *Hpa* cultures and inoculation of experimental plants. We also thank D. Patel and J. Hulsmans for technical support with FACS. This work was supported by a BBSRC New Investigator

grant BB/H109502/1 to MG, a BBSRC grant BB/F001347 to JB and the EPSRC/BBSRC funded Warwick Systems Biology Doctoral Training Centre to TLRC. All microarray data have been deposited in GEO (Series GSE58046), released upon publication.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2015.00527>

Figure S1 | Confocal microscopy images of *Hyaloperonospora arabidopsidis* (*Hpa*) infection marker *ProDMR6::GFP* expression in an *Arabidopsis* cotyledon, 7 d.p.i. with compatible *Hpa* isolate Noks1. (A,B)

Expression of the marker follows the pattern of pathogen spread across the cotyledon. (C–E) Cells expressing the marker appear to contain haustoria.

Figure S2 | Fluorescence expression profiles for cells analyzed and sorted with FACS. (A,B) Dot plots of output from the 580/30 nm vs. 530/40 nm bandpass filters on the BD Influx, using a workspace derived from Grönlund et al. (2012). (A) Protoplasts generated from uninfected, 14-day-old *ProDMR6::GFP* seedlings, where cells were collected exclusively from the low 580/low 530 (GFP-negative) fate. (B) Protoplasts generated from *ProDMR6::GFP* inoculated with *Hpa* isolate Noks1, at 7 d.p.i., where cells were collected from both the low 580/low 530 (GFP-negative) and low 580/high 530 (GFP-positive, ~0.5%) gates. The high 580/low 530 gate represents cell debris.

Figure S3 | Pair-wise comparisons used to identify differentially expressed genes. Expression at 7 d.p.i. and 5 d.p.i. was compared for each cell type. Genes significantly differentially expressed in either *Hpa*-distal or *Hpa*-proximal cells, but not uninfected control cells, over time were considered as differentially expressed. To investigate differential expression between cell types, six comparisons were performed: *Hpa*-distal cells vs. uninfected control cells, *Hpa*-proximal cells vs. uninfected control cells, and *Hpa*-proximal cells vs. *Hpa*-distal cells, independently for each time point.

Figure S4 | Normalized \log_2 expression data for uninfected control, *Hpa*-proximal, and *Hpa*-distal cells at 5 and 7 d.p.i. Gene IDs, names, and descriptions are given based on the TAIR10 genome annotation. Column names are given in the format "Cell type, time point, replicate," where "Control5a" is control cells at 5 d.p.i., replicate 1.

Figure S5 | Transcripts differentially expressed between uninfected control, *Hpa*-proximal, and *Hpa*-distal cells at 5 and 7 d.p.i. Gene IDs, names and descriptions are given based on the TAIR10 genome annotation. Mean \log_2 expression values are given in columns D–I—the names of these columns are given in the format "Cell type, time point," where "Control5" is control cells at 5 d.p.i. Columns J–M are fold-changes for distal and proximal cells at 5 d.p.i. and 7 d.p.i., relative to uninfected controls at the respective time points. Columns N ("Group@5dpi") and O ("Group@7dpi") give the names of the groups which that transcript belonged to (see Figure 4), or state "Not DE" if not differentially expressed at that time point. Column P ("Newly detected?") states "Yes" or "No" as to whether the gene is newly detected in our dataset, based on the analysis in Figure 5A.

Figure S6 | Overrepresentation of Gene Ontology (GO) terms in *Hpa*-responsive gene groups. *Hpa*-responsive genes were grouped according to the localization of their response at 5 and 7 d.p.i. (see Table S2, Figure 4). All GO terms found to be overrepresented (Benjamini-Hochberg adjusted $p \leq 0.05$) in the groups are included, other than those that were only represented by a single gene in a group, which were excluded. The percentage frequency of each GO term in the cluster and in the *Arabidopsis thaliana* genome (of the 27,594 GO-annotated genes), as well as a list of genes annotated with the GO term are included.

References

- Asai, S., Rallapalli, G., Piquerez, S. J. M., Caillaud, M. C., Furzer, O. J., Ishaque, N., et al. (2014). Expression profiling during Arabidopsis/downy mildew interaction reveals a highly-expressed effector that attenuates responses to salicylic acid. *PLoS Pathog.* 10:e1004443. doi: 10.1371/journal.ppat.1004443
- Baxter, L., Tripathy, S., Ishaque, N., Boot, N., Cabral, A., Kemen, E., et al. (2010). Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. *Science* 330, 1549–1551. doi: 10.1126/science.1195203
- Brady, S. M., Orlando, D. A., Lee, J. Y., Wang, J. Y., Koch, J., Dinneny, J. R., et al. (2007). A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science* 318, 801–806. doi: 10.1126/science.1146265
- Caillaud, M. C., Asai, S., Rallapalli, G., Piquerez, S., Fabro, G., and Jones, J. D. G. (2013). A downy mildew effector attenuates salicylic acid-triggered immunity in Arabidopsis by interacting with the host mediator complex. *PLoS Biol.* 11:e1001732. doi: 10.1371/journal.pbio.1001732
- Caillaud, M. C., Piquerez, S. J., Fabro, G., Steinbrenner, J., Ishaque, N., Beynon, J., et al. (2012). Subcellular localization of the Hpa RxLR effector repertoire identifies a tonoplast-associated protein HaRxL17 that confers enhanced plant susceptibility. *Plant J.* 69, 252–265. doi: 10.1111/j.1365-313X.2011.04787.x
- Chandran, D., Inada, N., Hather, G., Kleindt, C. K., and Wildermuth, M. C. (2010). Laser microdissection of Arabidopsis cells at the powdery mildew infection site reveals site-specific processes and regulators. *Proc. Natl. Acad. Sci. U.S.A.* 107, 460–465. doi: 10.1073/pnas.0912492107
- Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743. doi: 10.1046/j.1365-313x.1998.00343.x
- Coates, M. E., and Beynon, J. L. (2010). *Hyaloperonospora arabidopsidis* as a pathogen model. *Annu. Rev. Phytopathol.* 48, 329–345. doi: 10.1146/annurev-phyto-080508-094422
- Dinneny, J. R., Long, T. A., Wang, J. Y., Mace, D., Pointer, S., Barran, C., et al. (2008). Cell identity mediates the response of Arabidopsis roots to salinity and iron stress. *Science* 320, 942–945. doi: 10.1126/science.1153795
- Durrant, W. E., and Dong, X. (2004). Systemic acquired resistance. *Annu. Rev. Phytopathol.* 42, 185–209. doi: 10.1146/annurev-phyto.42.040803.140421
- Fabro, G., Steinbrenner, J., Coates, M., Ishaque, N., Baxter, L., Studholme, D. J., et al. (2011). Multiple candidate effectors from the oomycete pathogen *Hyaloperonospora arabidopsidis* suppress host plant immunity. *PLoS Pathog.* 7:e1002348. doi: 10.1371/journal.ppat.1002348
- Gifford, M. L., Dean, A., Gutierrez, R. A., Coruzzi, G. M., and Birnbaum, K. D. (2008). Cell-specific nitrogen responses mediate developmental plasticity. *Proc. Natl. Acad. Sci. U.S.A.* 105, 803–808. doi: 10.1073/pnas.0709559105
- Grönlund, J. T., Eyres, A., Kumar, S., Buchanan-Wollaston, V., and Gifford, M. L. (2012). Cell specific analysis of Arabidopsis leaves using fluorescence activated cell sorting. *J. Vis. Exp.* 4:4214. doi: 10.3791/4214
- Hok, S., Danchin, E. G. J., Allasia, V., Panabieres, F., Attard, A., and Keller, H. (2011). An Arabidopsis (malectin-like) leucine-rich repeat receptor-like kinase contributes to downy mildew disease. *Plant Cell Environ.* 34, 1944–1957. doi: 10.1111/j.1365-3040.2011.02390.x
- Holub, E. B. (2007). Natural variation in innate immunity of a pioneer species. *Curr. Opin. Plant Biol.* 10, 415–424. doi: 10.1016/j.pbi.2007.05.003
- Huibers, R. P., de Jong, M., Dekter, R. W., and Van den Ackerveken, G. (2009). Disease-specific expression of host genes during downy mildew infection of Arabidopsis. *Mol. Plant Microbe Interact.* 22, 1104–1115. doi: 10.1094/MPMI-22-9-1104
- Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., et al. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4, 249–264. doi: 10.1093/biostatistics/4.2.249
- Jones, J. D. G., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi: 10.1038/nature05286
- Karimi, M., Inze, D., and Depicker, A. (2002). GATEWAY vectors for Agrobacterium-mediated plant transformation. *Trends Plant Sci.* 7, 193–195. doi: 10.1016/S1360-1385(02)02251-3
- Karve, R., and Iyer-Pascuzzi, A. S. (2015). Digging deeper: high-resolution genome-scale data yields new insights into root biology. *Curr. Opin. Plant Biol.* 24, 24–30. doi: 10.1016/j.pbi.2015.01.007
- Lapin, D., and Van den Ackerveken, G. (2013). Susceptibility to plant disease: more than a failure of host immunity. *Trends Plant Sci.* 18, 546–554. doi: 10.1016/j.tplants.2013.05.005
- Liu, G., Ji, Y., Bhuiyan, N. H., Pilot, G., Selvaraj, G., Zou, J., et al. (2010). Amino acid homeostasis modulates salicylic acid-associated defense responses in Arabidopsis. *Plant Cell* 22, 3845–3863. doi: 10.1105/tpc.110.079392
- Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: a cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* 21, 3448–3449. doi: 10.1093/bioinformatics/bti551
- Mukhtar, M. S., Carvunis, A. R., Dreze, M., Epple, P., Steinbrenner, J., Moore, J., et al. (2011). Independently evolved virulence effectors converge onto hubs in a plant immune system network. *Science* 333, 596–601. doi: 10.1126/science.1203659
- Nobuta, K., Okrent, R. A., Stoutemyer, M., Rodibaugh, N., Kempema, L., Wildermuth, M. C., et al. (2007). The GH3 acyl adenylase family member PBS3 regulates salicylic acid-dependent defense responses in Arabidopsis. *Plant Physiol.* 144, 1144–1156. doi: 10.1104/pp.107.097691
- R Development Core Team. (2008). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: URL <http://www.R-project.org>
- Rehmany, A. P., Gordon, A., Rose, L. E., Allen, R. L., Armstrong, M., Whisson, S. C., et al. (2005). Differential recognition of highly divergent downy mildew avirulence gene alleles by RPP1 resistance genes from two Arabidopsis lines. *Plant Cell* 17, 1839–1850. doi: 10.1105/tpc.105.031807
- Richman, L., Meylan, P. R., Munoz, M., Pinaud, S., and Mirkovitch, J. (2002). An adenovirus-based fluorescent report vector to identify and isolate HIV-infected cells. *J. Virol. Methods* 99, 9–21. doi: 10.1016/S0166-0934(01)00375-5
- Rogers, E. D., Jackson, T., Moussaieff, A., Sharoni, A., and Benfey, P. N. (2012). Cell type-specific transcriptional profiling: implications for metabolite profiling. *Plant J.* 70, 5–17. doi: 10.1111/j.1365-313X.2012.04888.x
- Smyth, G. K. (2004). Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* 3, Article 3. doi: 10.2202/1544-6115.1027
- Thöne, F., Schwanhauser, B., Becker, D., Ballmaier, M., and Bumann, D. (2007). FACS-isolation of Salmonella-infected cells with defined bacterial load from mouse spleen. *J. Microbiol. Methods* 71, 220–224. doi: 10.1016/j.mimet.2007.08.016
- Tomé, D. F., Steinbrenner, J., and Beynon, J. L. (2014). A growth quantification assay for *Hyaloperonospora arabidopsidis* isolates in *Arabidopsis thaliana*. *Methods Mol. Biol.* 1127, 145–158. doi: 10.1007/978-1-62703-986-4_12
- Tsuda, K., and Somssich, I. E. (2015). Transcriptional networks in plant immunity. *New Phytol.* 206, 932–947. doi: 10.1111/nph.13286
- Vadassery, J., Tripathi, S., Prasad, R., Varma, A., and Oelmüller, R. (2009). Monodehydroascorbate reductase 2 and dehydroascorbate reductase 5 are crucial for a mutualistic interaction between *Piriformospora indica* and Arabidopsis. *J. Plant Physiol.* 166, 1263–1274. doi: 10.1016/j.jplph.2008.12.016
- van Damme, M., Huibers, R. P., Elberse, J., and Van den Ackerveken, G. (2008). Arabidopsis DMR6 encodes a putative 2OG-Fe(II) oxygenase that is defense-associated but required for susceptibility to downy mildew. *Plant J.* 54, 785–793. doi: 10.1111/j.1365-313X.2008.03427.x
- Wang, W., Barnaby, J. Y., Tada, Y., Li, H., Tor, M., Caldelari, D., et al. (2011a). Timing of plant immune responses by a central circadian regulator. *Nature* 470, 110–115. doi: 10.1038/nature09766
- Wang, Y. H., Gehring, C., and Irving, H. R. (2011b). Plant natriuretic peptides are apoplastic and paracrine stress responses molecules. *Plant Cell Physiol.* 52, 837–850. doi: 10.1093/pcp/pcr036
- Wildermuth, M. C., Dewdney, J., Wu, G., and Ausubel, F. M. (2001). Isochorismate synthase is required to synthesize salicylic acid for plant defence. *Nature* 414, 562–565. doi: 10.1038/35107108

- Yadav, R. K., Girke, T., Pasala, S., Xie, M., and Reddy, G. V. (2009). Gene expression map of the Arabidopsis shoot apical meristem stem cell niche. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4941–4946. doi: 10.1073/pnas.0900843106
- Zeilmaker, T., Ludwig, N. R., Elberse, J., Seidl, M. F., Berke, L., Van Doorn, A., et al. (2015). DOWNY MILDEW RESISTANT 6 and DMR6-LIKE OXYGENASE 1 are partially redundant but distinct suppressors of immunity in Arabidopsis. *Plant J.* 81, 210–222. doi: 10.1111/tpj.12719
- Zipfel, C. (2014). Plant pattern-recognition receptors. *Trends Immunol.* 35, 345–351. doi: 10.1016/j.it.2014.05.004

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Coker, Cevik, Beynon and Gifford. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Identification of a core set of rhizobial infection genes using data from single cell-types

Da-Song Chen¹, Cheng-Wu Liu², Sonali Roy², Donna Cousins², Nicola Stacey² and Jeremy D. Murray^{2*}

¹ State Key Laboratory of Agricultural Microbiology, College of Life Science and Technology, Huazhong Agricultural University, Wuhan, China, ² John Innes Centre, Department of Cell and Developmental Biology, Norfolk, UK

OPEN ACCESS

Edited by:

Marc Libault,
University of Oklahoma, USA

Reviewed by:

Ulrike Mathesius,
Australian National University,
Australia

Md Shakhawat Hossain,
University of Missouri, USA

*Correspondence:

Jeremy D. Murray,
John Innes Centre, Department
of Cell and Developmental Biology,
Norwich Research Park, Norwich,
Norfolk NR4 7UH, UK
jeremy.murray@jic.ac.uk

Specialty section:

This article was submitted to
Plant Systems and Synthetic Biology,
a section of the journal
Frontiers in Plant Science

Received: 12 May 2015

Accepted: 13 July 2015

Published: 28 July 2015

Citation:

Chen D-S, Liu C-W, Roy S,
Cousins D, Stacey N and Murray JD
(2015) Identification of a core set
of rhizobial infection genes using data
from single cell-types.
Front. Plant Sci. 6:575.
doi: 10.3389/fpls.2015.00575

Genome-wide expression studies on nodulation have varied in their scale from entire root systems to dissected nodules or root sections containing nodule primordia (NP). More recently efforts have focused on developing methods for isolation of root hairs from infected plants and the application of laser-capture microdissection technology to nodules. Here we analyze two published data sets to identify a core set of infection genes that are expressed in the nodule and in root hairs during infection. Among the genes identified were those encoding phenylpropanoid biosynthesis enzymes including *Chalcone-O-Methyltransferase* which is required for the production of the potent Nod gene inducer 4',4-dihydroxy-2-methoxychalcone. A promoter-GUS analysis in transgenic hairy roots for two genes encoding *Chalcone-O-Methyltransferase* isoforms revealed their expression in rhizobially infected root hairs and the nodule infection zone but not in the nitrogen fixation zone. We also describe a group of *Rhizobially Induced Peroxidases* whose expression overlaps with the production of superoxide in rhizobially infected root hairs and in nodules and roots. Finally, we identify a cohort of co-regulated transcription factors as candidate regulators of these processes.

Keywords: infection threads, methoxychalcone, medicarpin, CCAAT-box, infection zone, nod genes, Nod factors, nodulation

Introduction

Most legumes are able to interact with soil bacteria called rhizobia to form special root structures called nodules within which the bacteria are able to fix atmospheric nitrogen, a process which is fueled by host photosynthate. In *Medicago truncatula*, as with many legumes, the rhizobia infects the host root by colonizing special intracellular tubular invaginations called infection threads that form first in root hairs and then later in mature nodules. This process begins with rhizobial attachment near the tip of growing root hairs which then curl and entrap a rhizobial microcolony in an infection pocket, also called an infection focus. The infection thread then extends as a tubular structure from this pocket through the length of the root hair, becoming colonized by rhizobia as it grows (Fournier et al., 2008). The process of infection thread formation is then repeated in the underlying cell layers allowing the rhizobia access to the inner root layers which have divided to form the NP. Root hair infection involves the upregulation of 100s of genes that regulate several different processes including host-symbiont signaling and diverse developmental processes including cell growth and engagement of the cell cycle (Libault et al., 2010; Breakspear et al., 2014). While maturing the nodule forms several developmental zones: an apical

meristem (Zone I), an infection zone containing cells that form infection threads (ZII), a nitrogen fixing zone comprised of giant cells filled with endocytosed nitrogen-fixing rhizobia (ZIII), an interzone (IZ), and a senescent zone, where no nitrogen fixation takes place (ZIV). The process of infection, from the initial entry of the rhizobia into the infection threads up until their release into symbiosomes, requires constant communication between the host and symbiont. At the heart of the dialog is the production of bacterial signaling compounds called Nod factors. Nod factors are produced in response to plant flavonoids, phenylpropanoid compounds which are secreted by the roots into the rhizosphere (Peters et al., 1986; Redmond et al., 1986; Maxwell et al., 1989; Kape et al., 1992; Phillips et al., 1994; Zuanazzi et al., 1998; Subramanian et al., 2007). The secreted flavonoids induce the expression of rhizobial nod genes required for the production of Nod factors through activation of the transcription factor NodD (Mulligan and Long, 1985; Rossen et al., 1985). The flavonoids produced by a given host are often specific to certain symbionts, only activating Nod genes in nodulation-competent rhizobia (Györgypal et al., 1988). In *M. truncatula*, several flavonoids capable of inducing Nod gene in its symbiont *Sinorhizobium meliloti* have been identified and one of the most potent is 4,4'-dihydroxy 2'-methoxychalcone (Kapulnik et al., 1987; Maxwell et al., 1989; Orgambide et al., 1994). The enzyme Chalcone O-Methyltransferase (ChOMT) in the closely related *M. sativa* is required for production of this compound from isoliquiritigenin (4,2',4'-trihydroxychalcone; Maxwell et al., 1992). Our recent study has shown that the *M. truncatula* ortholog, *ChOMT1*, and three other close homologs (*ChOMT2*, *ChOMT3*, and *ChOMT4*) were induced during infection of root hairs (Breakspear et al., 2014). Transcripts for two of these isoforms were also detected in the infection zone (Zone II) of mature nodules (Roux et al., 2014), but a spatio-temporal analysis of the expression of these genes during early infection is lacking.

One of the first physiological events occurring during nodulation is the generation of ROS (Cook et al., 1995; Ramu et al., 2002), which is coincident with enhanced flavonoid biosynthesis (van Brussel et al., 1990; Mathesius et al., 1998; Hassan and Mathesius, 2012). Earlier studies on the production and breakdown of ROS by rhizobia suggest that ROS levels must be maintained between certain limits for the symbiosis to be successful (Santos et al., 2000; Jamet et al., 2003, 2007). The expression of *Rhizobial Induced Peroxidase 1 (RIP1)* was shown to increase in response to rhizobial infection or Nod factors and was found to be correlated with the production of ROS during rhizobial infection (Cook et al., 1995; Ramu et al., 2002). Recently, transcriptomic analysis of root hairs revealed that *RIP1* belongs to a family of 10 *RIPs* that are similarly strongly induced upon inoculation with rhizobia or Nod factors, further suggesting an important role for the modulation of ROS during nodulation (Breakspear et al., 2014).

In this study we compared published data from studies using single-cell types, i.e., microarray data for root hairs collected from seedlings inoculated with rhizobia (Breakspear et al., 2014) and RNAseq data for laser-capture microdissected nodules (Roux et al., 2014) to identify a core set of genes associated with infection. We then investigated the expression of two

ChOMT genes and the generation of superoxides during the different stages of nodulation in infected root hair and nodule cells. Finally, we identified transcription factors induced during rhizobial infection of root hairs and preferentially expressed in the nodule infection zone as potential regulators of ROS and phenylpropanoid production.

Materials and Methods

Analysis of Transcriptomics Data

The first data set used for our analysis was from Roux et al. (2014). This study generated RNAseq data by laser-capture microdissection of *M. truncatula* cv. Jemalong A17 nodules harvested 15 days post inoculation with *S. meliloti* 2011. This was designated as data set A. A second data set was from a microarray analysis of root hairs of harvested from *M. truncatula* cv. Jemalong A17 seedlings 1, 3, and 5 days post inoculation with *S. meliloti* 1021 or treated with Nod factor (24 h post treatment; Breakspear et al., 2014), designated data set B. Roux et al. (2014) assigned all plant and bacterial genes that were differentially expressed between zones into 13 hierarchical clusters based on their expression across the sampled nodule zones. To compare the two data sets we identified all genes in dataset B that could be assigned to the above-mentioned clusters and compared their frequencies to that of all genes assigned to clusters in dataset A using a chi-squared test with the online GraphPad QuickCalc software <http://graphpad.com/quickcalcs/chisquared1.cfm>.

Plant Growth Conditions

Composite plants and seedlings were grown in controlled environment chambers with 16 h day length with a light intensity of 90–130 $\mu\text{mol m}^{-2} \text{s}^{-1}$ and constant temperature of 20°C. For rhizobial inoculation a 24-h culture of *S. meliloti* 1021 rhizobia was spun down and resuspended in buffered nodulation medium to an optical density of 0.02 at 600 nm, and 3 mL of the culture was used to inoculate each of the growing *M. truncatula* plants.

Promoter-GUS Analysis

For the promoter-GUS analyses, the promoter regions of *ChOMT2* (Medtr3g021440; from –23 to –1803 bp) and *ChOMT3* (Medtr7g011900; from –23 to –1885 bp) were amplified by PCR using Phusion High Fidelity DNA polymerase (NEB). The fragments were then cloned into pDONR207 and after sequence confirmation were transferred to the destination vector pKGWFS7 upstream of the GUS open reading frame using the GATEWAY cloning system (Life Technologies) as per the manufacturer's recommended protocol. The vector was then transformed into *M. truncatula* (A17) using *Agrobacterium rhizogenes*-mediated hairy root transformation as described previously (Breakspear et al., 2014). The composite plants with transgenic roots were then transferred to a soil mixture of Terragreen (Oil-Dri UK) and silver sand 1:1 mixture, inoculated with *S. meliloti* 1021, and watered as needed with distilled water. Nodulated roots were then stained for GUS activity at 1, 2, and 3 weeks post inoculation, as previously described (Breakspear et al., 2014).

Nitroblue tetrazolium (NBT) Staining

For nitroblue tetrazolium (NBT) staining, *M. truncatula* seedlings were grown in Terragreen (Oil-Dri UK) and silver sand mixture (1:1) for 7 days and then inoculated with *S. meliloti* 1021. At 14 days post inoculation the nodulated roots were cut off and stained in 0.1% NBT water solution (Promega UK) for 40 min in the dark at room temperature. Imbedding and sectioning of plant tissues was carried out as previously described (Guan et al., 2013).

Results

Infection threads form in both root hairs and in cells of nodule ZII. To identify genes common to infection thread formation in these two cell types we compared genes that were induced either by rhizobia or by purified Nod factors in root hairs with the genes expressed in different nodule zones (Roux et al., 2014). To do this we used the RNAseq analysis described by Roux et al. (2014) which assigned genes to 13 clusters based on their expression across tissues sampled from each nodule zone using laser-capture microdissection. The advantage of this approach is that genes that are only expressed in the nodule but not in the root hair or vice versa are not considered. For example, the large family genes encoding for nodule-specific cysteine-rich (NCR) peptides, which are highly expressed in the nodule but are not expressed in root hairs, are thereby excluded. A total of 768 genes were identified that met the clustering criteria and were induced by at least one treatment in root hairs. When these genes were assigned to their clusters a disproportionate number of the rhizobially induced genes in the WT and *sickle* (*skl*) mutant belonged to clusters 2–5, which are genes that are primarily expressed in the

nodule meristem (ZI) and distal ZII with some expression in proximal ZII (Table 1). Nod factor-induced genes displayed an overlapping pattern, with clusters 1–4 being over-represented, cluster one containing genes with the majority of their expression in the meristem (ZI) with some expression in distal ZII (Table 1). For each treatment and the combined treatments the observed frequencies in the different clusters were significantly different than the expected frequencies (Chi-squared tests, all p -values < 0.0001). This analysis suggests that a core set of genes underlies infection processes in both tissue types.

We then considered the genes that were induced by rhizobial infection and were found in the over-represented clusters 2–5 (Supplementary Table S1). Amongst these genes were several members of the isoflavonoid biosynthesis pathway analyzed in Breakspear et al. (2014) including *Chalcone Reductase* (*CHR*) which is required for the production of isoliquiritigenin and *ChOMT* which converts isoliquiritigenin to the even more potent nod gene inducer methoxychalcone. Four *ChOMT* genes are induced in root hairs after rhizobial infection (Breakspear et al., 2014), but their expression patterns during nodulation have not been investigated in detail. We investigated the expression of two of these genes, *ChOMT2* and *ChOMT3*, during the early stages of infection and nodule formation using promoter-GUS analysis in *A. rhizogenes* transformed roots. We found that both genes were expressed in root hairs undergoing infection (Figures 1A–D) and throughout the NP (Figures 1E,F). As the nodule matured the expression of both genes was restricted to the apex (Figures 1G,H), and in root tips (data not shown). The expression of both genes was typically observed at numerous infection sites along the root (Figures 1A,B,D,I). Hand sectioning of the nodules revealed expression of *ChOMT2* in the

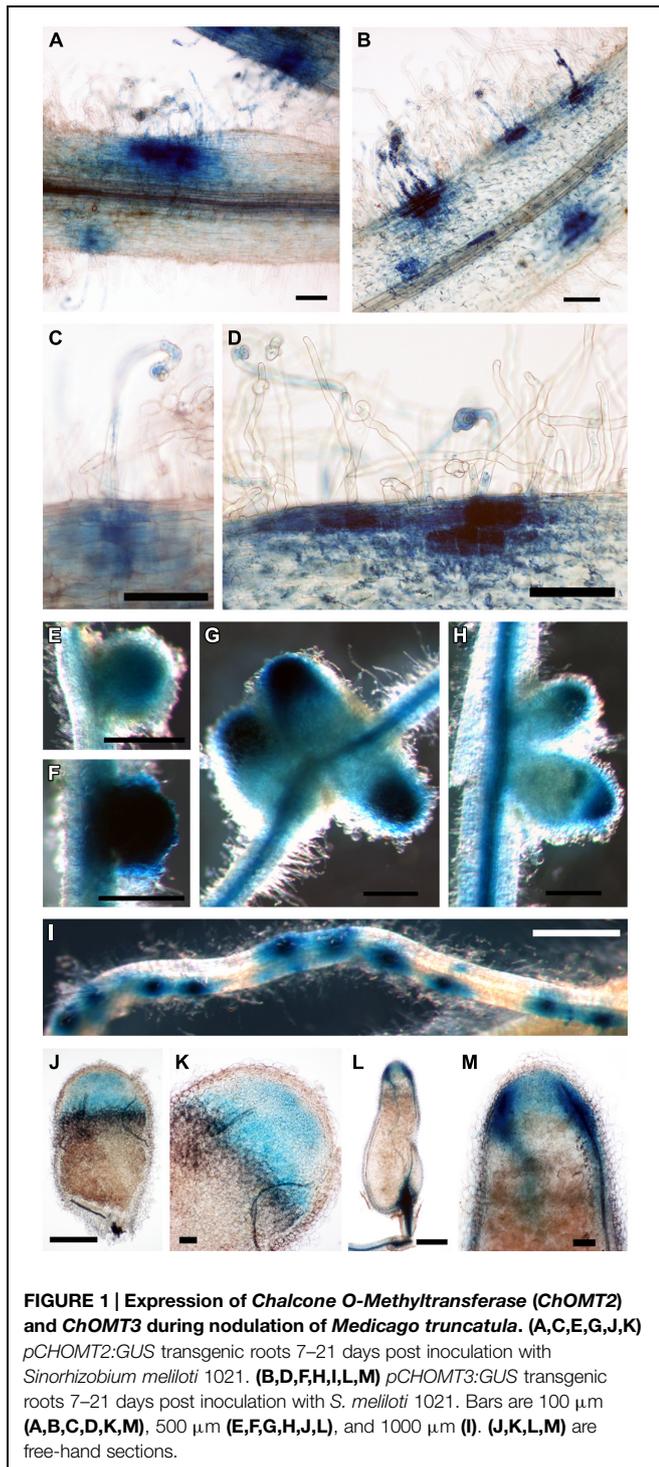
TABLE 1 | Comparison of the observed number of genes induced by rhizobial inoculation in root hairs belonging to different clusters of nodule expression (Roux et al., 2014).

Cluster ¹	RHIZ + WT		RHIZ + <i>skl</i>		Nod factors + WT		All treatments	
	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp
1	13	16.5	25	31.9	14	11.9	34	38.9
2	75	31.0	108	59.8	57	22.4	131	72.9
3	41	17.1	69	32.9	31	12.3	79	40.1
4	78	33.0	143	63.5	34	23.8	168	77.4
5	55	52.6	142	101.4	28	38.0	166	123.6
6	6	24.1	12	46.4	14	17.4	24	56.6
7	4	9.9	8	19.1	7	7.2	12	23.3
8	5	8.2	6	15.8	4	5.9	9	19.2
9	3	25.5	4	49.1	5	18.4	9	59.8
10	3	10.2	5	19.7	0	7.4	6	24.0
11	17	37.8	39	72.9	16	27.3	47	88.8
12	13	21.9	23	42.1	12	15.8	30	51.4
13	14	39.1	46	75.4	14	28.2	53	91.9
#genes ²	327		630		236		768	
% ³	76.1%		73.3%		57.6%		70.8%	

¹clusters from Roux et al. (2014).

²number of genes induced by rhizobia or Nod factor (Breakspear et al., 2014).

³percentage of genes in overrepresented clusters (shaded cells).



nodule meristem, the infection zone, with some expression in the IZ, closely matching the published LCM data (Figures 1J,K; Roux et al., 2014). *ChOMT3* expression was also found in the nodule meristem and but was absent from the IZ and proximal infection zone (Figures 1L,M), again closely reflecting the data of Roux et al. (2014). Expression of both genes was absent in the nitrogen fixation zone. *ChOMT3* expression was also strong in

the apical part of the nodule vasculature, including the nodule vascular meristems (Figures 1L,M).

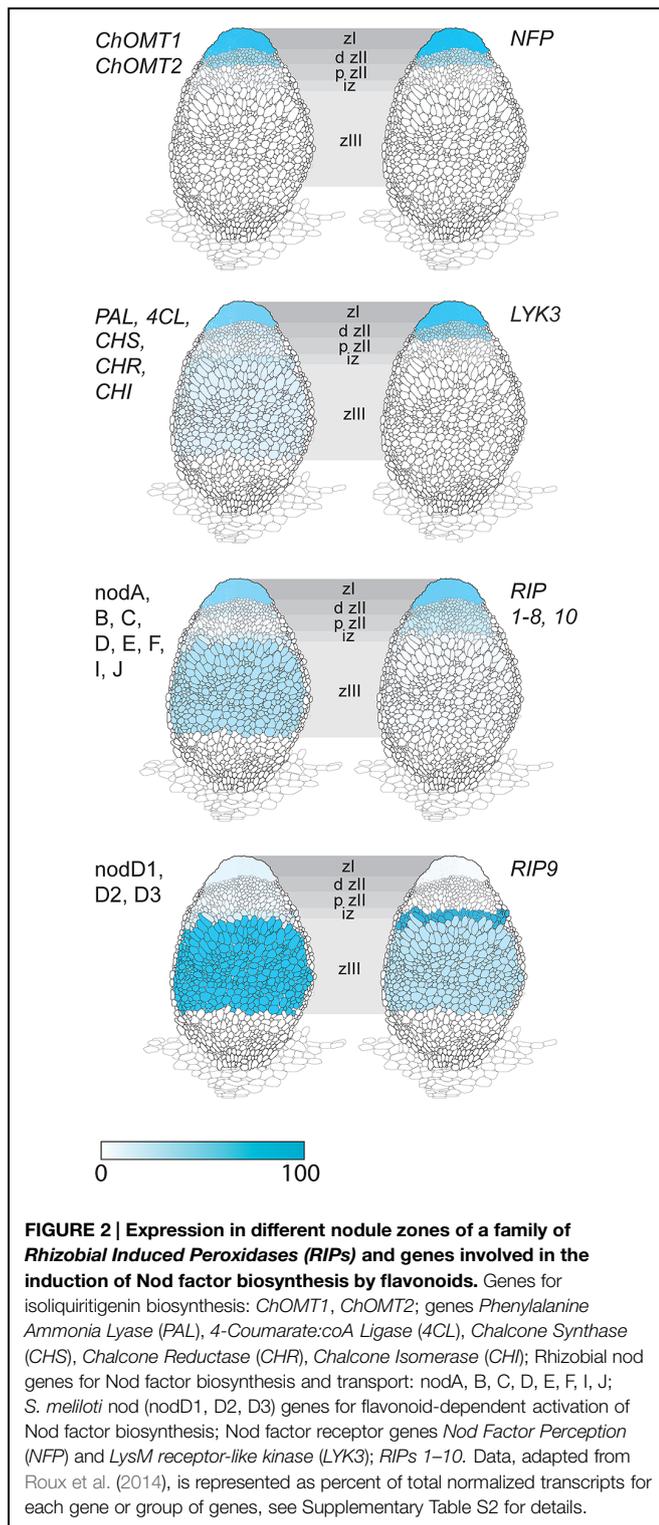
The expression of *ChOMT2* and *ChOMT3* corresponded well with the other genes encoding enzymes for isoliquiritigenin production which had an average of 65% of their total expression in ZI and ZII (Figure 2, Supplementary Table S1). However, some transcripts of these genes (~10% of the total) were also detected in each of the remaining zones (IZ, ZIII). The expression of the rhizobial nod genes required for the biosynthesis and secretion of Nod factors which are known to be induced by flavonoids also had strong expression in ZI and proximal ZII (on average 55% of the total reads), however, as noted by Roux et al. (2014), they also showed expression (~25% of total reads) in the nitrogen fixation zone (ZIII; Figure 2). Based on our results it seems unlikely that methoxychalcone is responsible for the induction of nod gene expression in ZIII. Instead this expression could be induced by isoliquiritigenin for which the required genes are expressed in ZIII (Figure 2). In contrast to the Nod factor biosynthesis and transport genes, the three *nodD* genes were mostly expressed in ZIII, with only one (*nodD3*) having some expression in the nodule apex (ZI; Figure 2).

Rhizobial Induced Peroxidases

Another gene class identified in the overlap between the inoculated root hair and nodule data sets was type III peroxidases, which are apoplastic/cell wall localized enzymes have a role in cell wall remodeling (Francoz et al., 2015). The genes encoding this family of enzymes have been shown to be inducible by rhizobia and Nod factors in root hairs and designated as *RIP1-10* (Breakspear et al., 2014). Most of the previously identified *RIPs* are also present in ZI and ZII (Figure 2, data from Roux et al., 2014). Unexpectedly one of them, *RIP9*, was more highly expressed than all the others combined (Supplementary Table S3) and had a dramatically different expression domain, being >90% expressed in the IZ and ZIII (Figure 2). Since secreted peroxidases are capable of generating reactive oxygen species, we monitored superoxide production at the different stages of nodulation using NBT staining (Doke, 1983). The stain was seen in microcolonies and in infection threads at the early stages of nodulation (Figures 3A,B). The nascent NP were also strongly stained (Figure 3C) and as nodules matured, the staining was mainly found in the nodule apex (Figure 3D). Nodule sectioning revealed staining in a patchwork of individual cells located mostly in ZII and sometimes in ZIII (Figures 3E,F). Closer inspection revealed that the staining was present in cells containing infection threads in ZII (Figure 3G) and also in ZIII (Figure 3H). There was also staining in a thin layer of non-infected cells near the nodule apex (Figure 3F). In addition, staining was also strong in root tips of both mature primary roots and emerging lateral roots (Figure 3G).

Transcription Factors

The transcription factors controlling the expression of genes for flavonoid biosynthesis and ROS production in the nodule are currently not known. A large number of transcription factors (Table 2) were shared between the genes in nodule clusters 2–5 and genes induced in root hairs of infected plants (Table 3).



These included the well-studied *ERN1*, *ERN2*, and *NSP2* which are required for early infection events and nodule development (Oldroyd and Long, 2003; Kaló et al., 2005; Heckmann et al., 2006; Middleton et al., 2007; Cerri et al., 2012). In addition, four genes encoding CCAAT-box transcription factor subunits

(*NF-YA1*, *NF-YA2*, *NF-YC2*, *NF-YB7*) were in this group. *NF-YA1* and *NF-YC2* are important for both rhizobial infection and nodule development (Combiere et al., 2006, 2008; Zanetti et al., 2010; Soyano et al., 2013; Laporte et al., 2014), while *NF-YA2* is a close homolog of *NF-YA1*. This co-regulation across root hair and nodule-zones suggests that the encoded subunits may act as part of heterocomplex along with *NF-YB7* to regulate infection. Notably the *nf-yal1* mutant forms nodules with either reduced or absent meristems that fail to release the bacteria from the infection thread into symbiosomes (Xiao et al., 2014). A summary of all data (Breakspear et al., 2014; Roux et al., 2014) for the entire family of *M. truncatula* CCAAT-box transcription factors is included in Supplementary Table S4.

The roles of the remaining transcription factors in nodulation have not been defined. Of note in this group are the nodule-specific GRF-Zinc finger gene *N20*, and the *M. truncatula* ortholog of *Arabidopsis thaliana* Zinc Finger Protein 6 (*ZFP6*). *ZFP6* has been assigned a role in integration of gibberellic acid and cytokinin signaling (Zhou et al., 2013).

Discussion

Root Hair Infection-Genes are Expressed in Nodule Zone II

Current molecular genetic studies of nodulation rely heavily on gene expression data to provide insight into symbiotic processes. Candidate genes identified in this manner can then be followed up using reverse genetic studies such as the *Tnt1* or *LORE1* transposon mutant collections available in *M. truncatula* or *Lotus japonicus*, respectively (Tadege et al., 2008; Cheng et al., 2011; Fukai et al., 2012; Urbański et al., 2012). Recently laser-capture has been used to target the specific regions of indeterminate nodules (Roux et al., 2014), including the infection zone which contain the cells becoming infected by rhizobia. A subsequent study, which transcriptionally profiled root hairs from seedlings undergoing infection (Breakspear et al., 2014), provides a unique opportunity to compare gene expression responses during infection of these two cell types. A disproportionate number of genes induced in root hairs of infected plants were expressed in the nodule apex. The majority of genes induced in root hairs in response to purified Nod factors were also expressed in the nodule apex (71% belonging to clusters 2–5), mirroring the analysis by Roux et al. (2014) who reported that 63.4% of genes induced by Nod factors in intact root samples in an earlier study (Czaja et al., 2012) belonged to clusters 1–4. This overlap suggests that a core set of genes required to support rhizobial infection in both epidermal and cortical (nodule) tissues are directly induced by Nod factor signaling.

ChOMT Genes are Expressed Specifically in Infected Roots Hairs and in the Nodule Apex While Other Flavonoid Biosynthesis Genes are Expressed Throughout the Nodule

Expression of the *ChOMT* genes required for the production of the nod-gene inducing flavonoid methoxychalcone was high in infected root hairs and within the nodule was restricted to the

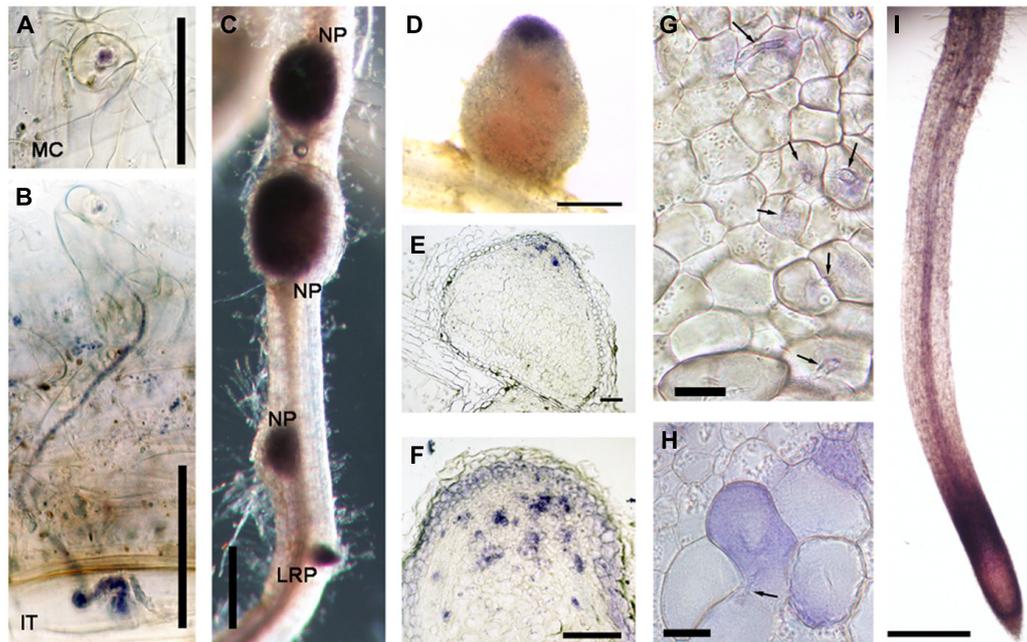


FIGURE 3 | Nitroblue tetrazolium (NBT) staining of *M. truncatula* roots nodulated by *S. meliloti* 1021. (A) A microcolony within an infection focus. (B) An infection thread in a root hair. (C) Nodule primordia (NP) and a lateral root primordium (LRP). (D) An intact

nodule (E,F,G,H) nodule sections, 10 μ m thick (I) the primary root tip. In (G) and (H) arrows indicate infection threads. Plants harvested 14 dpi with *S. meliloti* 1021. Bars = 50 μ m (A,B), 100 μ m (E,F), and 500 μ m (C,D,I), 20 μ m (G,H).

TABLE 2 | Transcription factors associated with rhizobial infection.

Gene	Gene model/ Position ³	Class	Cluster	⁴ F1%	FIID%	FIIP%	IZ%	ZIII%
ZFP6 (At)	Chr7 31519725-31520145	zinc finger	2	73.8	25.7	0.0	0.5	0.0
MYB3R-1 (At)	Medtr3g110028	MYB	2	71.2	24.5	4.1	0.0	0.2
NSP2	Medtr3g072710	GRAS	2	57.5	31.7	9.9	0.7	0.1
WRKY40 (At)	Medtr2g105060	WRKY	2	47.4	26.1	18.6	7.9	0.0
ERN1	Medtr7g085810	AP2	2	55.8	36.1	7.9	0.1	0.0
BTF3 (At)	Medtr3g020660	BTF3b-like	4	41.7	26.9	18.7	7.2	5.5
	Medtr7g105790	ovate	4	46.6	25.4	22.0	5.1	0.9
WRKY23 (At)	Medtr7g109800	WRKY	4	46.5	36.0	8.2	3.1	6.2
ERF12 (At)	Medtr2g014300	ERF	4	45.7	34.7	13.3	3.2	3.1
ERF9 (At)	Medtr4g078710	ERF	4	38.6	36.6	13.6	5.4	5.7
MYB6 (At)	Medtr3g097450	MYB	4	30.4	36.0	13.8	13.7	6.2
NF-YA1 ¹	Medtr1g056530	CCAAT-box	3	26.3	33.6	27.2	8.5	4.4
CRF4 (At)	Medtr3g090760	ERF	3	26.0	37.8	25.4	7.1	3.7
	Medtr6g092540	MYB-like	3	23.9	48.2	17.0	5.7	5.1
NF-YC2	Medtr7g113680	CCAAT-box	3	21.2	34.6	12.3	18.8	13.1
	Medtr8g101650	MYB-like	3	22.1	45.5	10.9	8.9	12.6
NF-YA2	Medtr7g106450	CCAAT-box	3	17.3	45.6	31.5	4.7	0.9
ERN2	Medtr6g029180	AP2	3	9.8	65.9	10.6	13.1	0.7
RRA3 ²	Medtr3g088630	RR	3	12.0	55.5	30.8	1.5	0.1
N20	Medtr7g086040	GRF-ZF	3	10.2	63.0	26.6	0.3	0.0
NF-YB7	Medtr8g091720	CCAAT-box	3	2.9	41.3	51.2	4.6	0.0

Data from Roux et al. (2014).

Cells shading indicates relative strength of expression for a given gene across nodule zones.

¹Previously HAP2.1.

²Previously MtARR8, see notes on family nomenclature (Liu et al., 2015).

³Chromosome position Mtv4.0.

⁴Percentage of normalized RNAseq reads across samples (Roux et al., 2014).

TABLE 3 | Transcription factors associated with rhizobial infection.

Gene name	Gene model/Position	probeset	Fold change vs. control				
			1 dpi ¹	3 dpi	5 dpi	5 dpi <i>skI</i>	24 h NF ²
ZFP6 (At)	Chr7 31519725-31520145	Mtr.32254.1.S1_at	2.0	1.5	6.4*	9.0*	1.6
MYB3R-1 (At)	Medtr3g110028	Mtr.36181.1.S1_at	1.0	1.1	2.1	2.6	2.1*
NSP2	Medtr3g072710	Mtr.44789.1.S1_at	1.6*	0.9	1.7	3.0*	0.7*
WRKY40 (At)	Medtr2g105060	Mtr.10896.1.S1_s_at	0.5	0.7	0.8	0.8	2.6*
ERN1	Medtr7g085810	Mtr.7556.1.S1_at	1.5*	1.2	3.9*	8.9*	2.5*
BTF3 (At)	Medtr3g020660	Mtr.12254.1.S1_x_at	1.2	1.4	1.5*	2.1*	1.2
	Medtr7g105790	Mtr.25635.1.S1_at	3.4	3.3	4.3	22.9*	1.1
WRKY23 (At)	Medtr7g109800	Mtr.25503.1.S1_at	1.3	1.6*	1.5	2.4*	0.9
ERF12 (At)	Medtr2g014300	Mtr.985.1.S1_at	8.7*	15.8*	13.7*	19.4*	5.4*
ERF9 (At)	Medtr4g078710	Mtr.26158.1.S1_at	3.3*	3.2*	4.7*	23.5*	2.9*
MYB6 (At)	Medtr3g097450	Mtr.1797.1.S1_at	1.2	1.2	2.1	3.3	2.4*
NF-YA1	Medtr1g056530	Mtr.43750.1.S1_at	36.9*	157.1*	114.5*	549.6*	59.0*
CRF4 (At)	Medtr3g090760	Mtr.41691.1.S1_at	2.6	2.8	5.3*	8.5*	1.5
	Medtr6g092540	Mtr.25945.1.S1_at	1.9*	2.8	5.6*	9.6*	10.5*
NF-YC2	Medtr7g113680	Mtr.48660.1.S1_at	1.7*	2.1	2.0*	4.6*	2.0*
	Medtr8g101650	Mtr.9513.1.S1_at	13.1	11.2*	9.8*	10.3*	11.6*
NF-YA2	Medtr7g106450	Mtr.1584.1.S1_at	1.7*	2.8	2.9*	5.6*	3.6*
ERN2	Medtr6g029180	Mtr.43947.1.S1_at	1.2	0.8	1.8	2.6*	1.9*
RRA3	Medtr3g088630	Mtr.31738.1.S1_at	2.3	7.4	12.1*	42.6*	1.0
N20	Medtr7g086040	Mtr.14503.1.S1_at	1.1	3.9	9.3*	15.0*	1.6
NF-YB7	Medtr8g091720	Mtr.4282.1.S1_at	10.2*	3.2	3.0	2.0	5.1*

Data from Breakspear et al. (2014).

¹dpi, days post inoculation with *Sinorhizobium meliloti* 1021.

²NF, 24 h post treatment with 10 nM Nod factors.

*Significantly different compared to control $p < 0.05$, also shaded light red (see Breakspear et al., 2014 for details).

infection zone, matching the pattern of expression of rhizobial Nod factor synthesis/export genes and the Nod factor receptors *Nod Factor Perception (NFP)* and *LysM receptor-like kinase (LYK3)*, (Figure 2; Roux et al., 2014). This is consistent with the detection of flavonoids in root hairs undergoing infection (Hassan and Mathesius, 2012) and is consistent with evidence showing that Nod factor production by the rhizobia is required at all stages of infection thread development (Marie et al., 1992; Den Herder et al., 2007). While *ChOMT* expression was restricted to ZII, transcripts for genes required to make the methoxychalcone precursor isoliquiritigenin as well as rhizobial nod genes were also detected in the N-fixation zone (ZIII; Figure 2; Roux et al., 2014). The authors further showed using promoter-GUS analysis that the expression of the nod genes coincided with the relatively infrequent infection threads observed in ZIII. The significance of these sporadic infections which can also be observed in mature determinate nodules of *L. japonicus* (unpublished results) remains to be determined. Our analysis suggests that the nod gene inducer in the N-fixing zone of the nodule is unlikely to be methoxychalcone since *ChOMT* expression is tightly confined to the nodule apex. On the other hand isoliquiritigenin seems to fit the role as several genes required for its synthesis, including *CHR* and *Chalcone Isomerase*, have moderate levels of expression in ZIII. Another function recently highlighted for flavonoids is their role as antioxidants, having roles in stomatal closure and drought

response (Nakabayashi et al., 2014; Watkins et al., 2014). It seems possible that these compounds could serve to help buffer against damage from ROS generated in cells undergoing infection. Two *Vestitone Reductase* genes needed for the production of medicarpin, one of which is expressed at infection sites in the epidermis (Breakspear et al., 2014), were also found to be expressed in ZI and the IZ (Supplementary Table S2). A potential role for medicarpin is selection against incompatible rhizobia and other opportunistic bacteria as previously discussed (Breakspear et al., 2014).

The expression of the *nodD* genes did not correspond with the Nod factor signaling domain that was clearly delineated by the expression of the Nod factor receptors, rhizobial nod factor biosynthesis genes, and the *ChOMTs*, with only one gene, *nodD3*, having about 30% of its expression in this domain and *nodD1* and *nodD2* being mainly expressed in ZIII. The lack of correspondence between *nodD* expression and the Nod factor signaling domain may reflect the fact that the *nodD* gene, at least in *Rhizobium leguminosarum*, is constitutively expressed and is not induced by flavonoids (Rossen et al., 1985). Furthermore, the NodD protein is present at relatively low levels in the rhizobia, suggesting that low constitutive expression is sufficient for its activity (Schlaman et al., 1991). While *nodD* expression is not induced by flavonoids, activation of nod genes by some NodD proteins is flavonoid-dependent; methoxychalcone strongly induces nod

gene expression in rhizobia containing extra copies of *nodD1* or *nodD2* in *S. meliloti* (Hartwig et al., 1990). In contrast, *NodD3* induction of *nod* gene expression is flavonoid-independent and requires the transcriptional regulator *syrM* (Mulligan and Long, 1989). Accordingly, in nodules *syrM* expression closely matches that of *nodD3* Roux et al. (2014; Supplementary Table S2). A study which examined all three single *nodD* mutants showed a small reduction in nodulation in the *nodD3* mutant only, but strongly reduced nodulation for all combinations of double mutants, suggesting all three isoforms are important in the interaction with *M. truncatula* (Smith and Long, 1998). Similar results were obtained with *M. sativa* (Honma and Ausubel, 1987). The available data suggest that *nodD1* and *nodD2* activation of *nod* gene expression in the infection zone occurs mainly through the action of methoxychalcone, but expression of these genes in ZIII may be mediated by another flavonoid, such as isoliquiritigenin, whilst *NodD3* operates independently of flavonoids and is sufficient to sustain nodulation.

Rhizobial Induced Peroxidases have Complementary Patterns of Expression in the Nodule

Also common to epidermal and cortical infection was the strong expression of a subset of *RIPs*. Type III peroxidases have been implicated in generation of apoplastic reactive oxygen species (Martinez et al., 1998; Bindschedler et al., 2006), with one important role being to promote cell wall hardening (Passardi et al., 2004). A corresponding role for these peroxidases has been proposed in the rigidification of the infection thread cell wall and matrix (Wisniewski et al., 2000). Expression of nine of the ten *RIPs* was strictly limited to Zones I and II of the nodule while another family member, *RIP9*, showed a complementary pattern, being very highly expressed (~7 times higher than all other *RIPs* combined) in the IZ and ZIII. The expression of *RIP9* therefore coincides with the low pO_2 in the proximal zones of the nodule; Roux et al. (2014) report that the leghemoglobin genes required for microaerobic conditions are strongly and abruptly upregulated in the IZ and remain highly expressed in ZIII. Indeed, low oxygen availability in these zones might explain the need for a highly expressed, separately regulated peroxidase isoform. While the link between type III peroxidases with rhizobial infection is well established (Cook et al., 1995; Ramu et al., 2002; this study) functional analysis is confounded by the presence of multiple family members with similar expression patterns. *RIP9* may therefore present an opportunity to use genetics to help study the role of these enzymes in nodulation.

Based on co-expression and co-regulation we identified candidate transcription factors involved in rhizobial infection. The group included *ERN1*, *ERN2*, *NSP2*, *NF-YA1*, and *NF-YC2*, all of which have been functionally implicated in rhizobial infection (Oldroyd and Long, 2003; Kaló et al., 2005; Comber et al., 2006, 2008; Heckmann et al., 2006; Middleton et al., 2007; Zanetti et al., 2010; Cerri et al., 2012; Soyano et al., 2013; Laporte et al., 2014). Two other genes encoding CCAAT-box subunits were also identified in the analysis, *NF-YA1*, a

close homolog of *NF-YA1*, and *NF-YB7*. As these transcriptional regulators act in heterocomplexes having one of each A, B, and C subunits, it is tempting to speculate that these act together to control infection, with *NF-YA1* and *NF-YA2* acting interchangeably. Among the unstudied members of this group is a GRF zinc finger protein N20, which has expression that is highly nodulation-specific but with only weak aa sequence homology to other legumes. Also of interest is *ZPF6* which encodes a C2H2 transcription factor that is required for trichome development, and has been proposed as an integrator of GA and cytokinin signaling in this process (Zhou et al., 2013). Also found were the transcription factors cytokinin response regulator *RRA3*, an ERF (Medtr3g090760) with homology to *Arabidopsis Cytokinin Response Factor 4* (Table 2), and two genes encoding the GA biosynthetic enzymes, Ent-Kaurenoic Acid Oxidase 1 (KAO1) and Gibberellin 3-Oxidase 1 (GA3OX1; Supplementary Table S1), which were part of a larger set of GA and cytokinin related genes identified as upregulated after rhizobial inoculation (Breakspear et al., 2014; Liu et al., 2015). This indicates that the regulation of these hormones, along with auxin, is important during infection in both root hairs and within the nodule.

Challenges and Future Directions

This study illustrates the power of single-cell type approaches to study the key mechanisms that underlie a biological process. Through the study of gene expression associated with infection thread formation in two different tissues, tissue-specific and background effects can be eliminated and core processes exposed. Improvements can of course still be made; Roux et al. (2014) estimated a 10% rate of contamination of ZII transcripts in the ZI sample, which argues the need for sampling smaller areas of tissue. Furthermore, it can be difficult to discern different cell types, for instance between meristematic cells and cells of the distal infection zone (Limpens et al., 2013). Similarly, while root hair isolation methods used by Breakspear et al. (2014) or Libault et al. (2010) are technically less challenging than laser-capture microdissection, only a small percentage of root hairs in the sample are undergoing infection. Ultimately, a single cell-based approach offers the greatest resolution, having the potential to sub-classify cells from the same tissue. Such an approach would avoid the averaging out of highly localized phenomena such as the production of superoxide reported in this study. Another limitation of the described approach is evident from the analysis of the transcriptional regulators. It is clear that infection thread development is comprised of numerous processes occurring in parallel sometimes within the same cells which cannot easily be separated even using single-cell approaches. This can be addressed in three ways. The first is to use a developmental time series, as employed in Breakspear et al. (2014), which allowed partial resolution of events occurring before and during infection. The second is the use of relevant sensory or chemical inputs that can perturb individual components of the system. For instance a subset of the genes may be inducible by treatment with ROS, allowing them to be partitioned away from the larger set of co-regulated genes. In this respect the

ever-growing gene expression atlases available for medicago and other plants present a useful resource (Benedito et al., 2008; Hruz et al., 2008). The third and most powerful approach is the use of mutants. Careful comparison of specific mutants defective in one or a few processes will provide a clearer picture of the transcriptional network underlying rhizobial infection.

Our work compares two different types of data sets and identifies a core set of infection genes common to infection in both root hairs and nodules with specific attention to transcription factors that can serve as a starting point for future studies. In addition we show for the first time the expression pattern of two genes encoding *ChOMT* isoforms, an enzyme that plays a key role in the symbiosis that so far has only been studied at the biochemical level. We confirmed expression of these genes in the nodule infection zone, and have extended this knowledge by showing that these genes are specifically induced in infected root hairs, and that one gene is expressed in the nodule vascular

bundle. We show using NBT staining that superoxide is being produced specifically in cells undergoing infection in root hairs and the nodule further demonstrating the tight link between ROS production and the infection process.

Acknowledgments

This work was supported by the Biotechnology and Biological Sciences Research Council Grants BB/G023832/1 and BB/L010305/1 and the John Innes Foundation.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2015.00575>

References

- Benedito, V. A., Torres-Jerez, I., Murray, J. D., Andriankaja, A., Allen, S., Kakar, K., et al. (2008). A gene expression atlas of the model legume *Medicago truncatula*. *Plant J.* 55, 504–513. doi: 10.1111/j.1365-313X.2008.03519.x
- Bindschedler, L. V., Dewdney, J., Blee, K. A., Stone, J. M., Asai, T., Plotnikov, J., et al. (2006). Peroxidase-dependent apoplastic oxidative burst in *Arabidopsis* required for pathogen resistance. *Plant J.* 47, 851–863. doi: 10.1111/j.1365-313X.2006.02837.x
- Breakspear, A., Liu, C., Roy, S., Stacey, N., Rogers, C., Trick, M., et al. (2014). The root hair 'infectome' of *Medicago truncatula* uncovers changes in cell cycle genes and reveals a requirement for auxin signalling in rhizobial infection. *Plant Cell* 26, 4680–4701. doi: 10.1105/tpc.114.133496
- Cerri, M. R., Frances, L., Laloum, T., Auriac, M. C., Niebel, A., Oldroyd, G. E., et al. (2012). *Medicago truncatula* ERN transcription factors: regulatory interplay with NSP1/NSP2 GRAS factors and expression dynamics throughout rhizobial infection. *Plant Physiol.* 160, 2155–2172. doi: 10.1104/pp.112.203190
- Cheng, X., Wen, J., Tadege, M., Ratet, P., and Mysore, K. S. (2011). Reverse genetics in *Medicago truncatula* using Tnt1 insertion mutants. *Methods Mol. Biol.* 678, 179–190. doi: 10.1007/978-1-60761-682-5_13
- Comber, J. P., de Billy, F., Gamas, P., Niebel, A., and Rivas, S. (2008). Trans-regulation of the expression of the transcription factor MtHAP2-1 by a uORF controls root nodule development. *Genes Dev.* 22, 1549–1559. doi: 10.1101/gad.461808
- Comber, J. P., Frugier, F., de Billy, F., Boualem, A., El-Yahyaoui, F., Moreau, S., et al. (2006). MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in *Medicago truncatula*. *Genes Dev.* 20, 3084–3088. doi: 10.1101/gad.402806
- Cook, D., Dreyer, D., Bonnet, D., Howell, M., Nony, E., and VandenBosch, K. (1995). Transient induction of a peroxidase gene in *Medicago truncatula* precedes infection by *Rhizobium meliloti*. *Plant Cell* 7, 43–55. doi: 10.1105/tpc.7.1.43
- Czaja, L. F., Hogeckamp, C., Lamm, P., Maillet, F., Martinez, E. A., Samain, E., et al. (2012). Transcriptional responses toward diffusible signals from symbiotic microbes reveal MtNFP- and MtDMI3-dependent reprogramming of host gene expression by arbuscular mycorrhizal fungal lipochito oligosaccharides. *Plant Physiol.* 159, 1671–1685. doi: 10.1104/pp.112.195990
- Den Herder, J., Vanhee, C., De Rycke, R., Corich, V., Holsters, M., and Goormachtig, S. (2007). Nod factor perception during infection thread growth fine-tunes nodulation. *Mol. Plant Microbe Interact.* 20, 129–137. doi: 10.1094/MPMI-20-2-0129
- Doke, N. (1983). Involvement of superoxide anion generation in the hypersensitive response of potato tuber tissues to infection with an incompatible race of *Phytophthora infestans* and to the hyphal wall components. *Physiol. Plant Pathol.* 23, 345–357. doi: 10.1016/0048-4059(83)90019-X
- Fournier, J., Timmers, A. C., Sieberer, B. J., Jauneau, A., Chabaud, M., and Barker, D. G. (2008). Mechanism of infection thread elongation in root hairs of *Medicago truncatula* and dynamic interplay with associated rhizobial colonization. *Plant Physiol.* 148, 1985–1995. doi: 10.1104/pp.108.125674
- Francoz, E., Ranocha, P., Nguyen-Kim, H., Jamet, E., Burlat, V., and Dunand, C. (2015). Roles of cell wall peroxidases in plant development. *Phytochemistry* 112, 15–21. doi: 10.1016/j.phytochem.2014.07.020
- Fukai, E., Soyano, T., Umehara, Y., Nakayama, S., Hirakawa, H., Tabata, S., et al. (2012). Establishment of a *Lotus japonicus* gene tagging population using the exon-targeting endogenous retrotransposon LORE1. *Plant J.* 69, 720–730. doi: 10.1111/j.1365-313X.2011.04826.x
- Guan, D., Stacey, N., Liu, C., Wen, J., Mysore, K. S., Torres-Jerez, I., et al. (2013). Rhizobial infection is associated with the development of peripheral vasculature in nodules of *Medicago truncatula*. *Plant Physiol.* 162, 107–115. doi: 10.1104/pp.113.215111
- Györgypal, Z., Iyer, N., and Kondorosi, A. (1988). Three regulatory nodD alleles of diverged flavonoid-specificity are involved in host-dependent nodulation by *Rhizobium meliloti*. *Mol. Gen. Genet.* 21, 85–92. doi: 10.1007/BF00322448
- Hartwig, U. A., Maxwell, C. A., Joseph, C. M., and Phillips, D. A. (1990). Effects of alfalfa nod gene-inducing flavonoids on nodABC transcription in *Rhizobium meliloti* strains containing different nodD genes. *J. Bacteriol.* 172, 2769–2773.
- Hassan, S., and Mathesius, U. (2012). The role of flavonoids in root-rhizosphere signalling: opportunities and challenges for improving plant-microbe interaction. *J. Exp. Bot.* 63, 3429–3444. doi: 10.1093/jxb/err430
- Heckmann, A. B., Lombardo, F., Miwa, H., Perry, J. A., Bunnewell, S., Parniske, M., et al. (2006). *Lotus japonicus* nodulation requires two GRAS domain regulators, one of which is functionally conserved in a non-legume. *Plant Physiol.* 142, 1739–1750. doi: 10.1104/pp.106.089508
- Honma, M. A., and Ausubel, F. M. (1987). *Rhizobium meliloti* has three functional copies of the nodD symbiotic regulatory gene. *Proc. Natl. Acad. Sci. U.S.A.* 84, 8558–8562. doi: 10.1073/pnas.84.23.8558
- Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., et al. (2008). Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv. Bioinformatics* 2008, 420747. doi: 10.1155/2008/420747
- Jamet, A., Mandon, K., Puppo, A., and Herouart, D. (2007). H₂O₂ is required for optimal establishment of the *Medicago sativa*/*Sinorhizobium meliloti* and symbiosis. *J. Bacteriol.* 189, 8741–8745. doi: 10.1128/JB.01130-07
- Jamet, A., Sigaud, S., Van de Sype, G., Puppo, A., and Herouart, D. (2003). Expression of the bacterial catalase genes during *Sinorhizobium*

- meliloti*–*Medicago sativa* symbiosis and their crucial role during the infection process. *Mol. Plant Microbe Interact.* 16, 217–225. doi: 10.1094/MPMI.2003.16.3.217
- Kaló, P., Gleason, C., Edwards, A., Marsh, J., Mitra, R. M., Hirsch, S., et al. (2005). Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators. *Science* 308, 1786–1789. doi: 10.1126/science.1110951
- Kape, R., Parniske, M., Brandt, S., and Werner, D. (1992). Isoliquiritigenin, a strong nod gene- and glyceollin resistance-inducing flavonoid from soybean root exudate. *Appl. Environ. Microbiol.* 58, 1705–1710.
- Kapulnik, Y., Joseph, C. M., and Phillips, D. A. (1987). Flavone limitations to root nodulation and symbiotic nitrogen fixation in alfalfa. *Plant Physiol.* 84, 1193–1196. doi: 10.1104/pp.84.4.1193
- Laporte, P., Lepage, A., Fournier, J., Catrice, O., Moreau, S., Jardinaud, M. F., et al. (2014). The CCAAT box-binding transcription factor NF-YA1 controls rhizobial infection. *J. Exp. Bot.* 65, 481–495. doi: 10.1093/jxb/ert392
- Libault, M., Farmer, A., Brechenmacher, L., Drnevich, J., Langley, R. J., Bilgin, D. D., et al. (2010). Complete transcriptome of the soybean root hair cell, a single-cell model, and its alteration in response to *Bradyrhizobium japonicum* infection. *Plant Physiol.* 152, 541–552. doi: 10.1104/pp.109.148379
- Limpens, E., Moling, S., Hooiveld, G., Pereira, P. A., Bisseling, T., Becker, J. D., et al. (2013). Cell- and tissue-specific transcriptome analyses of *Medicago truncatula* root nodules. *PLoS ONE* 8:e64377. doi: 10.1371/journal.pone.0064377
- Liu, C., Breakspear, A., Roy, S., and Murray, J. D. (2015). Cytokinin responses counterpoint auxin signalling during rhizobial infection. *Plant Signal. Behav.* 10:e1019982. doi: 10.1080/15592324.2015.1019982
- Marie, C., Barny, M. A., and Downie, J. A. (1992). *Rhizobium leguminosarum* has two glucosamine synthases, GlmS and NodM, required for nodulation and development of nitrogen-fixing nodules. *Mol. Microbiol.* 6, 843–851. doi: 10.1111/j.1365-2958.1992.tb01535.x
- Martinez, C., Montillet, J. L., Bresson, E., Agnel, J. P., Daï, G. H., Daniel, J. F., et al. (1998). Apoplastic peroxidase generates superoxide anions in cells of cotton cotyledons undergoing the hypersensitive reaction to *Xanthomonas campestris* pv. *malvacearum* race 18. *Mol. Plant Microbe Interact.* 11, 1038–1047. doi: 10.1094/MPMI.1998.11.11.1038
- Mathesius, U., Bayliss, C., Weinman, J. J., Schlaman, H. R. M., Spaink, H. P., Rolfe, B. G., et al. (1998). Flavonoids synthesized in cortical cells during nodule initiation are early developmental markers in white clover. *Mol. Plant Microbe Interact.* 11, 1223–1232. doi: 10.1094/MPMI.1998.11.12.1223
- Maxwell, C. A., Edwards, R., and Dixon, R. A. (1992). Identification, purification, and characterization of S-adenosyl-L-methionine: isoliquiritigenin 2'-O-methyltransferase from alfalfa (*Medicago sativa* L.). *Arch. Biochem. Biophys.* 293, 158–166. doi: 10.1016/0003-9861(92)90379-B
- Maxwell, C. A., Hartwig, U. A., Joseph, C. M., and Phillips, D. A. (1989). A chalcone and two related flavonoids released from alfalfa roots induce nod genes of *Rhizobium meliloti*. *Plant Physiol.* 91, 842–847. doi: 10.1104/pp.91.3.842
- Middleton, P. H., Jakab, J., Penmetsa, R. V., Starker, C. G., Doll, J., Kaló, P., et al. (2007). An ERF transcription factor in *Medicago truncatula* that is essential for Nod factor signal transduction. *Plant Cell* 19, 1221–1234. doi: 10.1105/tpc.106.048264
- Mulligan, J. T., and Long, S. R. (1985). Induction of *Rhizobium meliloti* nodC expression by plant exudate requires nodD. *Proc. Natl. Acad. Sci. U.S.A.* 82, 6609–6613. doi: 10.1073/pnas.82.19.6609
- Mulligan, J. T., and Long, S. R. (1989). A family of activator genes regulates expression of *Rhizobium meliloti* nodulation genes. *Genetics* 122, 7–18.
- Nakabayashi, R., Yonekura-Sakakibara, K., Urano, K., Suzuki, M., Yamada, Y., Nishizawa, T., et al. (2014). Enhancement of oxidative and drought tolerance in *Arabidopsis* by overaccumulation of antioxidant flavonoids. *Plant J.* 77, 367–379. doi: 10.1111/tpj.12388
- Oldroyd, G. E. D., and Long, S. R. (2003). Identification and characterization of nodulation-signaling pathway 2, a gene of *Medicago truncatula* involved in Nod factor signaling. *Plant Physiol.* 131, 1027–1032. doi: 10.1104/pp.102.010710
- Orgambide, G. G., Philip-Hollingsworth, S., Hollingsworth, R. I., and Dazzo, F. B. (1994). Flavone-enhanced accumulation and symbiosis-related biological activity of a diglycosyl diacylglycerol membrane glycolipid from *Rhizobium leguminosarum* biovar trifolii. *J. Bacteriol.* 176, 4338–4347.
- Passardi, F., Penel, C., and Dunand, C. (2004). Performing the paradoxical: how plant peroxidases modify the cell wall. *Trends Plant Sci.* 9, 534–540. doi: 10.1016/j.tplants.2004.09.002
- Peters, N. K., Frost, J. W., and Long, S. R. (1986). A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes. *Science* 233, 977–980. doi: 10.1126/science.3738520
- Phillips, D. A., Dakora, F. D., Sande, E., Joseph, C. M., and Zon, J. (1994). Synthesis, release and transmission of alfalfa signals to rhizobial symbionts. *Plant Soil* 161, 69–80. doi: 10.1007/BF02183086
- Ramu, S. K., Peng, H. M., and Cook, D. R. (2002). Nod factor induction of reactive oxygen species production is correlated with expression of the early nodulin gene rip1 in *Medicago truncatula*. *Mol. Plant Microbe Interact.* 15, 522–528. doi: 10.1094/MPMI.2002.15.6.522
- Redmond, J. W., Batley, M., Djordjevic, M. A., Innes, R. W., Kuempel, P. L., and Rolfe, B. G. (1986). Flavones induce expression of nodulation genes in *Rhizobium*. *Nature* 323, 632–635. doi: 10.1038/323632a0
- Rossen, L., Shearman, C. A., Johnston, A. W., and Downie, J. A. (1985). The nodD gene of *Rhizobium leguminosarum* is autoregulatory and in the presence of plant exudate induces the nodA,B,C genes. *EMBO J.* 4, 3369–3373.
- Roux, B., Rodde, N., Jardinaud, M. F., Timmers, T., Sauviac, L., Cottret, L., et al. (2014). An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. *Plant J.* 77, 817–837. doi: 10.1111/tpj.12442
- Santos, R., Hérouart, D., Puppo, A., and Touati, D. (2000). Critical protective role of bacterial superoxide dismutase in *Rhizobium*–legume symbiosis. *Mol. Microbiol.* 38, 750–759. doi: 10.1046/j.1365-2958.2000.02178.x
- Schlaman, H. R., Horvath, B., Vijgenboom, E., Okker, R. J., and Lugtenberg, B. J. (1991). Suppression of nodulation gene expression in bacteroids of *Rhizobium leguminosarum* Biovar viciae. *J. Bacteriol.* 173, 4277–4287.
- Smith, L. S., and Long, S. R. (1998). Requirements for syrM and nodD genes in the nodulation of *M. truncatula* by *Rhizobium meliloti* 1021. *Mol. Plant Microbe Interact.* 11, 937–940. doi: 10.1094/MPMI.1998.11.9.937
- Soyano, T., Kouchi, H., Hirota, A., and Hayashi, M. (2013). NODULE INCEPTION directly targets NF-Y subunit genes to regulate essential processes of root nodule development in *Lotus japonicus*. *PLoS Genet.* 9:e1003352. doi: 10.1371/journal.pgen.1003352
- Subramanian, S., Stacey, G., and Yu, O. (2007). Distinct, crucial roles of flavonoids during legume nodulation. *Trends Plant Sci.* 12, 282–285. doi: 10.1016/j.tplants.2007.06.006
- Tadege, M., Wen, J., He, J., Tu, H., Kwak, Y., Eschstruth, A., et al. (2008). Large-scale insertional mutagenesis using the Tnt1 retrotransposon in the model legume *Medicago truncatula*. *Plant J.* 54, 335–347. doi: 10.1111/j.1365-313X.2008.03418.x
- Urbański, D. F., Malolepszy, A., Stougaard, J., and Andersen, S. U. (2012). Genome-wide LORE1 retrotransposon mutagenesis and high-throughput insertion detection in *Lotus japonicus*. *Plant J.* 69, 731–741. doi: 10.1111/j.1365-313X.2011.04827.x
- van Brussel, A. A., Recourt, K., Pees, E., Spaink, H. P., Tak, T., Wijffelman, C. A., et al. (1990). A biovar-specific signal of *Rhizobium leguminosarum* bv. *viciae* induces increased nodulation gene-inducing activity in root exudate of *Vicia sativa* subsp. *nigra*. *J. Bacteriol.* 172, 5394–5401.
- Watkins, J. M., Hechler, P. J., and Muday, G. K. (2014). Ethylene-induced flavonol accumulation in guard cells suppresses reactive oxygen species and moderates stomatal aperture. *Plant Physiol.* 164, 1707–1717. doi: 10.1104/pp.113.233528
- Wisniewski, J. P., Rathbun, E. A., Knox, J. P., and Brewin, N. J. (2000). Involvement of diamine oxidase and peroxidase in insolubilization of the extracellular matrix: implications for pea nodule initiation by *Rhizobium leguminosarum*. *Mol. Plant Microbe Interact.* 13, 413–420. doi: 10.1094/MPMI.2000.13.4.413
- Xiao, T. T., Schilderink, S., Moling, S., Deinum, E. E., Kondorosi, E., Franssen, H., et al. (2014). Fate map of *Medicago truncatula* root nodules. *Development* 141, 3517–3528. doi: 10.1242/dev.110775
- Zanetti, M. E., Blanco, F. A., Beker, M. P., Battaglia, M., and Aguilar, O. M. (2010). A C subunit of the plant nuclear factor NF-Y required for rhizobial infection

- and nodule development affects partner selection in the common bean-*Rhizobium etli* symbiosis. *Plant Cell* 22, 4142–4157. doi: 10.1105/tpc.110.079137
- Zhou, Z., Sun, L., Zhao, Y., An, L., Yan, A., Meng, X., et al. (2013). Zinc Finger Protein 6 (ZFP6) regulates trichome initiation by integrating gibberellin and cytokinin signaling in *Arabidopsis thaliana*. *New Phytol.* 198, 699–708. doi: 10.1111/nph.12211
- Zuanazzi, J. A. S., Clergeot, P. H., Quirion, J.-C., Husson, H.-P., Kondorosi, A., and Ratet, P. (1998). Production of *Sinorhizobium meliloti* nod gene activator and repressor flavonoids from *Medicago sativa* roots. *Mol. Plant Microbe Interact.* 11, 784–794. doi: 10.1094/MPMI.1998.11.8.784

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Chen, Liu, Roy, Cousins, Stacey and Murray. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Spatially resolved *in vivo* plant metabolomics by laser ablation-based mass spectrometry imaging (MSI) techniques: LDI-MSI and LAESI

Benjamin Bartels and Aleš Svatoš*

Research Group Mass Spectrometry/Proteomics, Max Planck Institute for Chemical Ecology, Jena, Germany

OPEN ACCESS

Edited by:

Marc Libault,
University of Oklahoma, USA

Reviewed by:

Sixue Chen,
University of Florida, USA
Zhibo Yang,
University of Oklahoma, USA

*Correspondence:

Aleš Svatoš,
Research Group Mass
Spectrometry/Proteomics, Max
Planck Institute for Chemical Ecology,
Max-Planck-Gesellschaft,
Hans-Knöll-Straße 8, Jena D-07745,
Germany
svatos@ice.mpg.de

Specialty section:

This article was submitted to
Plant Systems and Synthetic Biology,
a section of the journal
Frontiers in Plant Science

Received: 06 May 2015

Accepted: 15 June 2015

Published: 10 July 2015

Citation:

Bartels B and Svatoš A (2015)
Spatially resolved *in vivo* plant
metabolomics by laser ablation-based
mass spectrometry imaging (MSI)
techniques: LDI-MSI and LAESI.
Front. Plant Sci. 6:471.
doi: 10.3389/fpls.2015.00471

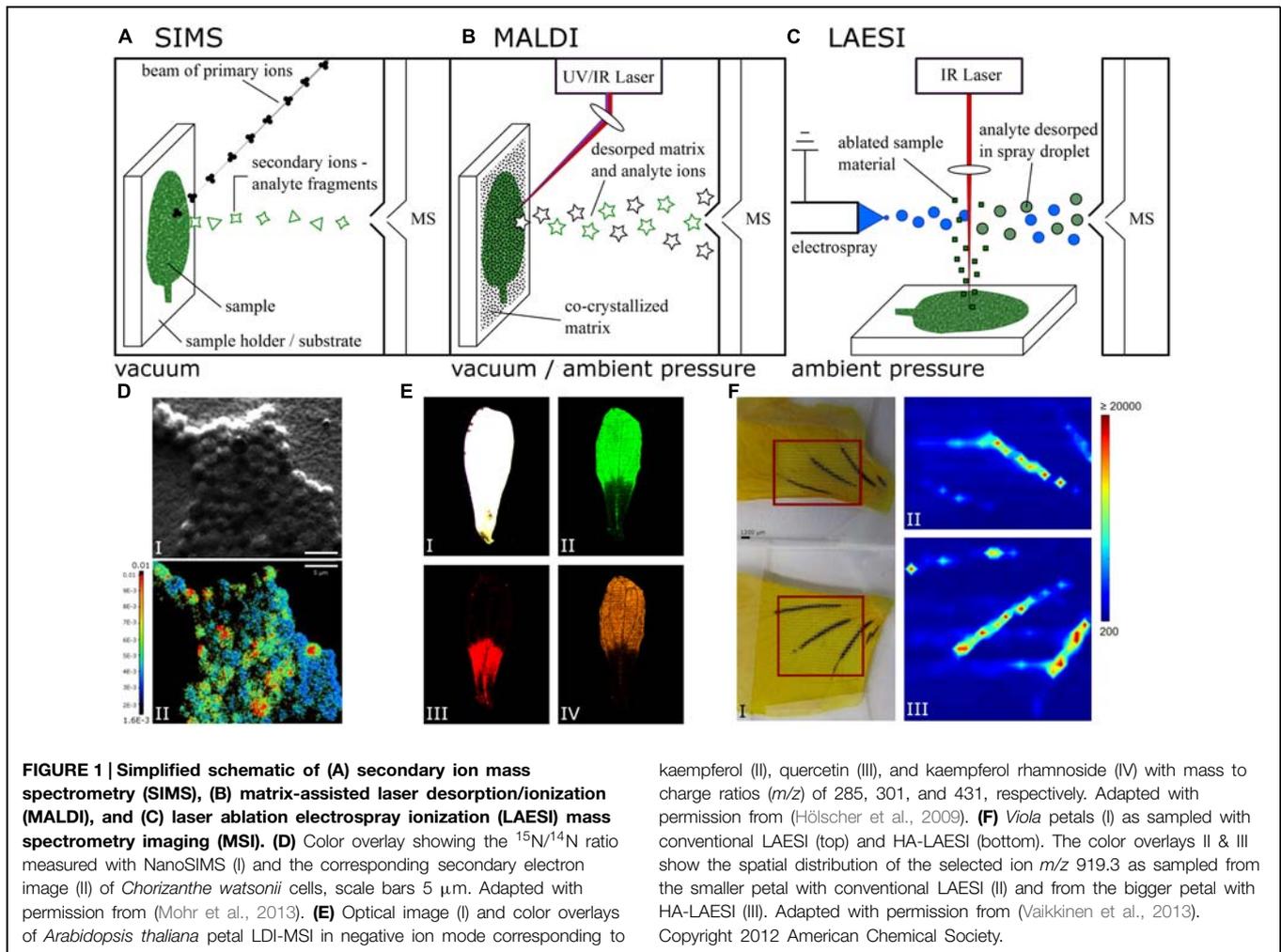
This short review aims to summarize the current developments and applications of mass spectrometry-based methods for *in situ* profiling and imaging of plants with minimal or no sample pre-treatment or manipulation. Infrared-laser ablation electrospray ionization and UV-laser desorption/ionization methods are reviewed. The underlying mechanisms of the ionization techniques—namely, laser ablation of biological samples and electrospray ionization—as well as variations of the LAESI ion source for specific targets of interest are described.

Keywords: ambient, ionization, mass spectrometry, laser ablation, electrospray

Introduction

Sample preparation is an important step that precedes acquisition of many kinds of data. However, often sample preparation is associated with artificially altering the biological or biochemical status of the system under study. In order to minimize this effect, we would like to have little to no sample preparation. If we can perform analysis directly *in vivo*, our data might fully represent the actual system. The usual workflow relies on sample dissection, solvent or thermal extraction and subsequent analysis using chromatographic methods connected to a detector with the needed selectivity. Minimal sample preparation facilitates the analytic process, by allowing people with minimal experience in analytical chemistry to perform the necessary steps without highly involved training. The sheer number of emerging ionization techniques involving minimal, ambient pressure sample preparation demonstrates the current interest, but, sadly, an alphabet soup of abbreviations has been created. Recent reviews (Bhardwaj and Hanley, 2014; El-Baba et al., 2014; Venter et al., 2014) summarize established techniques for most of the possible applications to date, providing an excellent guide for beginners to the field. These techniques are especially interesting for the life sciences (Alberici et al., 2010; Shrivastava and Setou, 2012), due to the delicate nature of biological samples. Biological mass spectrometry imaging (MSI) is profoundly profiting from these developments.

In addition to being the least intrusive approach, spatial resolution is an important feature for any imaging technique. Secondary ion mass spectrometry (SIMS) is the ionization technique for mass spectrometry (MS) that offers highest spatial resolution down to reported values of below one micron (Svatoš, 2010). Because it uses an ion beam to create secondary ions from the sample (Figure 1A), SIMS is not considered a soft ionization technique. Molecules tend to fragment upon ionization, and the utilization of SIMS is intrinsically linked to extensive sample preparation. SIMS



has successfully been used on biological samples for imaging (McMahon et al., 1995). In 2013, SIMS was successfully used to investigate the dynamics of nitrogen gas fixation of cyanobacteria at the level of a single cell (Mohr et al., 2013; **Figure 1D**). MSI of intact biomolecules, however, struggles to reach the level of a bacterial cell. In contrast, recent advances report single-cell resolution on eukaryotes with matrix-assisted ionization techniques, involving extensive sample preparation prior to analysis (Boggio et al., 2011). In early 2015, single-cell imaging was done within a tissue (Li et al., 2015b) utilizing laser ablation electrospray ionization (LAESI), which requires considerably less sample preparation.

A prominent ionization technique used in MSI of large biomolecule imaging is matrix-assisted laser desorption/ionization (MALDI; Caprioli et al., 1997; Bjarnholt et al., 2014; El-Baba et al., 2014). MALDI instrumentation for MSI is commercially available with a spatial resolution of 10 μm (FLEX series, Bruker, Bremen, Germany). MALDI requires the samples to be pre-processed extensively by dissolution in and co-crystallization together with a matrix. Originally restricted to vacuum application (Feigl et al., 1983; Karas et al., 1985, 1987), MALDI has since been adapted to work under atmospheric

pressure (Laiko et al., 2000; Li et al., 2007). Desorption and ionization of co-crystallized samples with matrix is facilitated by an ultraviolet (UV) laser and recently has also been used in conjunction with infrared (IR) lasers. The matrix molecules absorb most of the energy deposited to the sample by the laser and transfer the energy to the sample analytes more gently than via direct irradiation (Caprioli et al., 1997; Karas and Kruger, 2003), as depicted in **Figure 1B**. With MALDI, scientists can ionize very big molecules, e.g., proteins, non-destructively, which is one of the reasons why MALDI is used in protein MSI analysis. The method requires reliable matrix deposition and high ion yield (Karas and Kruger, 2003; El-Baba et al., 2014). To image plant cells – some as large as 50 μm – the spatial resolution of commercial instruments is sufficient. Laser desorption ionization (LDI) works similarly to MALDI but does not require an externally applied matrix. Because samples are not pre-treated with a matrix, spatial resolution is not compromised by matrix crystals, which could be larger than the studied cells.

Electrospray ionization (ESI) was originally designed to ionize long polymer chains (Dole et al., 1968) and has subsequently evolved (Yamashita and Fenn, 1984; Whitehouse et al., 1985)

to a commonly used ion source in mass spectroscopy. ESI has become very popular (Bhardwaj and Hanley, 2014), for example, in combination with liquid chromatography (Whitehouse et al., 1985) and been used for MSI as well, especially in the form of desorption electrospray ionization (DESI; Bjarnholt et al., 2014) and the closely related nano-DESI (Lanekoff et al., 2012). These techniques have been shown to achieve 50 and 20 μm spatial resolution, respectively (Campbell et al., 2012; Lanekoff et al., 2012). Instead of extracting analytes prior to analysis, both techniques extract analytes *in situ* prior to ionization directly from the sample surface (Venter et al., 2014). Control over the amount of sample surfaces wetted becomes imperative to avoid cross contamination and maintain spatial resolution.

In 2007, LAESI was introduced (Nemes and Vertes, 2007). The basic principle of LAESI combines LDI and ESI: ablation with a laser, and ionization via ESI, as shown in **Figure 1C**. However, LAESI uses an IR laser and relies on water present in the sample as a makeshift matrix (Apitz and Vogel, 2005; Nemes et al., 2012), a condition that most samples in life sciences fulfill. This way the deposition of an external matrix is not required, sample handling is simplified and the need to manipulate the samples prior to analysis is reduced. In a LAESI source, IR-laser light of 2940 nm wavelength is used to irradiate samples. At this wavelength, water has a major peak in its absorption spectrum and thus acts as a chromophore absorbing the deposited energy (Hale and Querry, 1973; Downing and Williams, 1975). Essential work describing the physics of ablating biological tissue with a laser was done recently (Vogel and Venugopalan, 2003b). The event of sample ablation can be split into at least two different phases based on the tensile strength of the sample (Vogel and Venugopalan, 2003a; Apitz and Vogel, 2005). Initially, irradiated sample material is heated and vaporization of molecules from the surface takes place (Vogel and Venugopalan, 2003a). When the energy deposition of the laser is larger than the energy consumption of the vaporization process, the water content of the sample is further heated and driven into a superheated state, leading to phase explosion upon relaxation to a stable state (Vogel and Venugopalan, 2003a; Apitz and Vogel, 2005; Chen et al., 2006). This results in material expulsion as well as tissue rupture and is primarily responsible for ablation efficiency

(Apitz and Vogel, 2005). The resulting ablation plume consists mostly of neutral matter in the form of nanoparticles, droplets, and large particulates. Experimental data suggest droplets from the electrospray plume intercept and fuse with the ablation plume nanoparticles, extracting analytes in the process (Nemes and Vertes, 2007). At this point, post-ionization by ESI takes over. A review of the research done on most of the aspects governing ESI (Kebarle and Verkerk, 2009) provides an excellent introduction to the field. Once ions have been generated from the sample, mass analyzers provide the means of detection.

The following section provides examples of instrumentation to illustrate the capabilities of the LAESI technique. LAESI displays promising potential for application in animal and plant metabolomics (Stolee et al., 2012; Stopka et al., 2014) and MSI of living plant tissue (Nemes and Vertes, 2007; Li et al., 2015b). For more information on different types of MSI methods, refer to **Table 1**.

Application of LAESI

The first realization of a LAESI ion source, as described by Nemes and Vertes (2007), consisted of a custom-built electrospray system, an Er:YAG laser tuned to a wavelength of 2940 μm , and a time-of-flight (TOF) mass spectrometer. One of the proof-of-concept experiments carried out was metabolic profiling of *Tagetes patula* seedlings *in vivo*. Several tentative assignments of metabolites from roots, leaves and stems were made. For that, accurate mass measurements, isotope patterns and metabolomic databases of model organisms such as *Arabidopsis thaliana* were considered. Cautious use of these databases was justified under the presumption that plants share certain metabolomics features (Nemes and Vertes, 2007; Nemes et al., 2008). Although LAESI is classified as a destructive method, seedlings subjected to the single-shot laser ablation were reported to survive the 350 μm wide ablation craters in roots, leaves, and stems.

Nemes et al. (2008) used a combination of LAESI and TOF mass analyzer techniques to show the usability of LAESI for MSI of plant tissues. Leaves of *Aphelandra squarrosa* with variegation patterns were subjected to two-dimensional imaging with a

TABLE 1 | Ionization techniques used for mass spectrometry imaging (MSI) of biological samples.

Ionization technique	Typical spot size/spatial resolution	Requirements/sample preparation	Reference
Secondary ion mass spectrometry (SIMS)	~100 nm, subcellular resolution possible	Sample must be stable enough in vacuum environment	McMahon et al. (1995), Colliver et al. (1997), Mohr et al. (2013)
Matrix-assisted laser desorption/ionization (MALDI)	~10 μm with commercially available instruments	Matrix molecules need to be co-crystallized with sample	Karas and Kruger (2003), El-Baba et al. (2014)
Laser desorption/ionization (LDI)	~5 μm with commercially available instruments	UV-absorbing analytes increase desorption/ionization	Hölscher et al. (2009), Kroiss et al. (2010), Hoelscher et al. (2014)
Matrix-assisted laser desorption electrospray ionization (MALDESI)	Spot size is 250–300 μm , spatial resolution of 45 μm with oversampling reported	Similar to MALDI but higher ion yield achievable through post ionization step	Sampson et al. (2006), Robichaud et al. (2014)
Desorption electrospray ionization (DESI)	50–20 μm spatial resolution, depending on source instrumentation	No particular sample preparation needed but sensitive to surface wetting	Campbell et al. (2012), Lanekoff et al. (2012)
Laser ablation electrospray ionization (LAESI)	350–15 μm spot size, depending on source instrumentation	Water in sample, e.g., in the form of cytosol	Nemes and Vertes (2007), Shrestha and Vertes (2009)

spatial resolution of 400 μm and depth profiling with a resolution of 50 μm . The actual spot size of the laser was reported as 350 μm , but a bigger step size was chosen to limit cross-talk in the acquisition of mass spectra. Nemes et al. (2008) were able to show that localization of the secondary metabolites kaempferol and luteolin, as well as certain derivatives with sugar moieties, coincides with the variegation pattern. The spatial distribution was then combined with the information gathered from depth profiling to visualize the spatial distribution of secondary plant metabolites in three dimensions. Depth profiling was realized by consecutive irradiation of the same spot (Nemes et al., 2008, 2009).

The work of Nemes et al. (2008, 2009) showed the feasibility of a LAESI ion source for analyzing and imaging metabolites in plant samples. Shrestha and Vertes (2009) improved upon the LAESI concept by using an etched, GeO_2 -based glass fiber to focus and deliver the laser to the sample. This made it possible to decrease the diameter of the ablation marks to slightly larger than $2R$, with R being the radius of the glass fiber tip's curvature, reported as roughly 15 μm in size and as forming ablation craters of ca. 30 μm . The metabolome of single epithelial cells from *Allium cepa* and *Narcissus pseudonarcissus* bulbs was analyzed and compared across species, but also compared to relative species within a particular sample tissue. Interestingly, the same cell type, *A. cepa* bulb epithelial cells and their *N. pseudonarcissus* equivalent, showed different contents of metabolites, with oligosaccharides and alkaloid, respectively, abundant (Shrestha and Vertes, 2009). By looking at epithelia from different layers of the same bulb, differently aged *A. cepa* cells were compared. The content of arginine was reported to decrease with increasing cell age, while the alliin gradient was oriented the other way around. Cells in an *A. cepa* bulb are older when located in the outer layers. Shrestha et al. (2011) also determined the influence of ablating event on single cells within a tissue on the surrounding cells and found no major disturbance compared to similar cells in undisturbed areas of the sampled tissue.

The same experimental set-up was also used to find biomarkers in the oil glands of *Citrus aurantium* leaves. For the initial mass spectra from achlorophyllous cells of *C. aurantium*, leaf oil glands and epidermal cells from distant parts of the same leaf were first measured and then compared. Different terpenes and terpenoids were found in the oil gland cells, which are absent in the epidermal cells and which contained flavonoids compounds not present in the gland cells (Shrestha et al., 2011).

The step to subcellular resolution was taken by Stolee et al. (2012). The LAESI set-up described previously (Shrestha and Vertes, 2009) was improved upon by adding a micro-dissection needle made out of tungsten. Prior to sample irradiation by the IR laser, the needle with a tip diameter of approximately 1 μm was used to cut open and peel back the cell wall of *A. cepa* epithelial cells. Metabolites such as hexose and alliin were reportedly found with higher abundance in cytosolic areas of a cell, whereas the amino acids arginine and glutamine were found more commonly in the area of the cell nucleus (Stolee et al., 2012). However, the improvement made by ablating the sample precisely goes hand in hand with the small sample volume from which ions

can be generated. This limitation obviously reduces sensitivity of the method and poses a general problem of spatially confined ionization techniques.

Depending on the properties of the electrospray solution used, imaging substances with strongly diverging polarities may be difficult to ionize simultaneously. A LAESI source was modified to address this problem (Vaikkinen et al., 2013). By adding a nebulizer chip blowing heated nitrogen gas toward the MS orifice, a more efficient ionization of both polar and non-polar compounds was expected (Careri et al., 1999; Boscaro et al., 2002). Compared to an unmodified LAESI ion source, heat-assisted LAESI (HA-LAESI) has shown to better ionize compounds with low polarity, as demonstrated on *Persea americana* mesocarp (Vaikkinen et al., 2013). A high abundance of signals assigned to triglycerides was observed in the MS spectrum measured with HA-LAESI. These particular peaks were less pronounced when using LAESI. To demonstrate imaging capabilities, Vaikkinen et al. (2013) used *Viola* flower petals and visualized the distribution of glycosides known to be present in *Viola* (Saito et al., 1983) as shown in **Figure 1F**. To further improve on ionizing low and non-polar compounds, a krypton discharge lamp for photo-ionization was added to the LAESI set-up to ionize anisole molecules with UV light that in turn ionize analytes in subsequent reactions taking place in the gas phase. The electrospray was exchanged for a nebulizer chip with an anisole and heated nitrogen gas flow (Vaikkinen et al., 2014), very similar to HA-LAESI. The technique was called laser ablation atmospheric pressure photoionization (LAAPPI). MSI was performed on *Salvia officinalis* leaves, and tentative assignment of multiple terpene and terpenoid compounds could be made (Vaikkinen et al., 2014). Because the IR light was focused using a lens instead of an etched glass fiber (Shrestha and Vertes, 2009) as described by Nemes et al. (2008), spatial resolution was reported as 400 μm .

Until recently, MSI was performed by measuring a sample step-wise using a predefined raster. Resolution of the mapping thus depended on the smallest possible step preventing pixel cross-talk. Li et al. (2015b) reported a procedure for LAESI-MSI, integrating light microscopy to assess and identify single cells within a sample tissue. An imaging raster consisting of cells defining that particular sample tissue was then created and used for systematic cell-by-cell imaging. Feasibility and proof-of-concept experiments on *A. cepa* bulb and *Lilium longiflorum* were performed using the precision of LAESI with an etched, GeO_2 -based glass fiber (Shrestha and Vertes, 2009). The capacity for separating isobaric and structurally isomeric ions in LAESI-MSI experiments was demonstrated by Li et al. (2015a) on *Pelargonium peltatum* leaves and mouse brain tissue.

Trying to make LAESI more compatible with complementary methods such as light microscopy, Compton et al. (2015) tried to spatially separate laser ablation from ESI. After ablation, the produced plume was carried into transfer tubing with nitrogen gas, and analytes were ionized with ESI after emerging from the 60 cm long tubing. Parts of *Viola* and *Acer* sp. were analyzed using remote-LAESI as proof-of-principle experiments. Signal strength was reported to be 27% of the intensity detected using conventional LAESI (Compton et al., 2015).

Laser ablation electrospray ionization was recently used as one of the methods to confirm the quantitative MSI of surface-occurring glucosinolate on *A. thaliana* leaf surfaces (Shroff et al., 2015). Data obtained from LAESI and liquid extraction surface analysis (LESA; Kertesz and Van Berkel, 2010) unambiguously supported the data obtained using a 9-aminacridine matrix sublimed on the leaves and imaged using vacuum MALDI-MSI.

In addition, LAESI has been applied to human- and animal-derived samples. The applicability of LAESI to blood and serum samples for medical purposes as well as antihistamine quantification directly from human urine samples has been shown (Nemes and Vertes, 2007). Since then, metabolomic and lipidomic analysis of the electric organ of *Torpedo californica* (Sripadi et al., 2009), rat and mouse brain (Nemes et al., 2010; Shrestha et al., 2010), fish gills (Shrestha et al., 2013), and other samples (Parsiegla et al., 2012; Shrestha et al., 2014) has been reported. A LAESI system, DP-1000 LAESI, is now available commercially from Protea Bioscience (Morgantown, WV, USA). The spatial resolution of the system is ca. 200 μm and can be attached to diverse mass spectrometers. Early data on MSI of pesticides, mycotoxins, and plant metabolites from lemon or rose leaves have recently been published (Nielen and van Beek, 2014) using this source.

Application of LDI-MSI in Planta

Laser desorption ionization can be applied *in planta*, as many important secondary metabolites contain conjugated double-bond systems like aromatic/heteroaromatic rings and show strong UV adsorption at 337 or 355 nm; both levels are emitted by the most common UV lasers. Plant pigments and compounds of the polyketide family readily absorb UV light and serve to desorb/ionize themselves. Elimination of MALDI matrices makes MSI in cellular resolution possible; see, for example, hypercins in glandular pigment cells of *Hypericum perforatum* or quercetin glucosides in *A. thaliana* petals or sepals as demonstrated by Hölscher et al. (2009) and shown in **Figure 1E**. A vacuum MALDI system Ultraflex (Bruker) with smart beam technology provided 10 μm spatial resolutions. Hypercins were shown to co-localize with dark pigment glands. A recent advance in developing systems with even higher spatial resolution as well as mass accuracy was commercialized in the AP-SMALDI imagine10 (TransMIT, Giessen, Germany) source attached to a Q-Exactive system with orbital mass analyzer (Thermo Scientific, San Jose, CA, USA). Laser spot sizes smaller than 5 μm are possible, and LDI measurements can be performed at ambient conditions thus preventing plant sample desiccation and deformation. This

References

- Alberici, R. M., Simas, R. C., Sanvido, G. B., Romao, W., Lalli, P. M., Benassi, M., et al. (2010). Ambient mass spectrometry: bringing MS into the “real world”. *Anal. Bioanal. Chem.* 398, 265–294. doi: 10.1007/s00216-010-3808-3
- Apitz, I., and Vogel, A. (2005). Material ejection in nanosecond Er : YAG laser ablation of water, liver, and skin. *Appl. Phys. A Mater. Sci. Process.* 81, 329–338. doi: 10.1007/s00339-005-3213-5

method is not limited to plants as was documented by MSIs of nematodes ingesting plant toxins from infected banana roots (Hoelscher et al., 2014) or on various MSI of antibiotics produced by actinomycetes on beewolf cocoons (Kroiss et al., 2010). LDI coupled with a plasma torch, also known as laser ablation inductively coupled plasma MS (LA-ICP-MS), is used for imaging distribution of metals *in planta* (Becker et al., 2010) or to localize proteins labeled with antibodies containing a metal-reporter ion (Bendall et al., 2011). This method shows extreme sensitivity, and as desorbed tissue debris undergoes post-ionization in a plasma torch, the technique is also quantitative.

Conclusion

Although plant tissues have been employed to characterize LAESI since the introduction of the technique in 2007, its application in plant metabolomics and MSI is still limited to proof-of-concept experiments, for example, with onion (*A. cepa*) bulbs. This limited use may be a result of the apparent dominance of MALDI applications in imaging with high spatial resolution and the initial barrier of acquiring a LAESI source, since instrumentation with high spatial resolution is not yet commercially available. Even custom-built realizations do not reach the benchmark resolutions reported for MALDI. Advantages such as the absence of an external matrix and the potential for direct correlation with microscopically gathered data through the means of software evaluation may, however, promote the use of LAESI over time. Interdisciplinary work, in particular, which is usually characterized by a wide variety of methods and thus depends on data correlation, might profit from these ionization techniques. As the literature reviewed here shows, the performance of the LAESI ion source is sufficient for utilization in larger studies of plant metabolomes, especially in MSI of target metabolites, and for answering current biological questions. The same can be said about LDI. It is less intrusive than MALDI, because it does not require an externally applied matrix. Additionally, the spatial resolution is not compromised by the matrix crystals, which could be larger than the studied cells. Typically, using diverse orthogonal methods can be fruitful and is of help in reducing experimental bias.

Acknowledgments

We thank Emily Wheeler for editorial assistance and the Max Planck Society for a stipend to BB and for financial support.

- Becker, J. S., Zoriy, M., Matusch, A., Wu, B., Salber, D., and Palm, C. (2010). Bioimaging of metals by laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS). *Mass Spectrom. Rev.* 29, 156–175.
- Bendall, S. C., Simonds, E. F., Qiu, P., Amir, E. A. D., Krutzik, P. O., Finck, R., et al. (2011). Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* 332, 687–696. doi: 10.1126/science.1198704

- Bhardwaj, C., and Hanley, L. (2014). Ion sources for mass spectrometric identification and imaging of molecular species. *Nat. Prod. Rep.* 31, 756–767. doi: 10.1039/c3np70094a
- Bjarnholt, N., Li, B., D'Alvise, J., and Janfelt, C. (2014). Mass spectrometry imaging of plant metabolites - principles and possibilities. *Nat. Prod. Rep.* 31, 818–837. doi: 10.1039/C3NP70100J
- Boggio, K. J., Obasuyi, E., Sugino, K., Nelson, S. B., Agar, N. Y., and Agar, J. N. (2011). Recent advances in single-cell MALDI mass spectrometry imaging and potential clinical impact. *Expert Rev. Proteomics* 8, 591–604. doi: 10.1586/epr.11.53
- Boscaro, F., Pieraccini, G., La Marca, G., Bartolucci, G., Luceri, C., Luceri, F., et al. (2002). Rapid quantitation of globotriaosylceramide in human plasma and urine: a potential application for monitoring enzyme replacement therapy in Anderson-Fabry disease. *Rapid Commun. Mass Spectrom.* 16, 1507–1514. doi: 10.1002/rcm.728
- Campbell, D. I., Ferreira, C. R., Eberlin, L. S., and Cooks, R. G. (2012). Improved spatial resolution in the imaging of biological tissue using desorption electrospray ionization. *Anal. Bioanal. Chem.* 404, 389–398. doi: 10.1007/s00216-012-6173-6
- Caprioli, R. M., Farmer, T. B., and Gile, J. (1997). Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal. Chem.* 69, 4751–4760. doi: 10.1021/ac970888i
- Careri, M., Elviri, L., and Mangia, A. (1999). Liquid chromatography-electrospray mass spectrometry of beta-carotene and xanthophylls - validation of the analytical method. *J. Chromatogr. A* 854, 233–244. doi: 10.1016/S0021-9673(99)00541-5
- Chen, Z. Y., Bogaerts, A., and Vertes, A. (2006). Phase explosion in atmospheric pressure infrared laser ablation from water-rich targets. *Appl. Phys. Lett.* 89, 3. doi: 10.1063/1.2243961
- Colliver, T. L., Brummel, C. L., Pacholski, M. L., Swanek, F. D., Ewing, A. G., and Winograd, N. (1997). Atomic and molecular imaging at the single-cell level with TOF-SIMS. *Anal. Chem.* 69, 2225–2231. doi: 10.1021/ac9701748
- Compton, L. R., Reschke, B., Friend, J., Powell, M., and Vertes, A. (2015). Remote laser ablation electrospray ionization mass spectrometry for non-proximate analysis of biological tissues. *Rapid Commun. Mass Spectrom.* 29, 67–73. doi: 10.1002/rcm.7077
- Dole, M., Mack, L. L., and Hines, R. L. (1968). Molecular beams of macroions. *J. Chem. Phys.* 49, 2240. doi: 10.1063/1.1670391
- Downing, H. D., and Williams, D. (1975). Optical-constants of water in infrared. *J. Geophys. Res.* 80, 1656–1661. doi: 10.1029/JC080i012p01656
- El-Baba, T. J., Lutowski, C. A., Wang, B. X., Inutan, E. D., and Trimpin, S. (2014). Toward high spatial resolution sampling and characterization of biological tissue surfaces using mass spectrometry. *Anal. Bioanal. Chem.* 406, 4053–4061. doi: 10.1007/s00216-014-7778-8
- Feigl, P., Schueler, B., and Hillenkamp, F. (1983). LAMMA-1000, a new instrument for bulk microprobe mass analysis by pulsed laser irradiation. *Int. J. Mass Spectrom. Ion Process.* 47, 15–18. doi: 10.1016/0020-7381(83)87125-3
- Hale, G. M., and Querry, M. R. (1973). Optical-constants of water in 200-nm to 200- μ m wavelength region. *Appl. Opt.* 12, 555–563. doi: 10.1364/AO.12.000555
- Hoelscher, D., Dhakshinamoorthy, S., Alexandrov, T., Becker, M., Bretschneider, T., Buerkert, A., et al. (2014). Phenalenone-type phytoalexins mediate resistance of banana plants (*Musa* spp.) to the burrowing nematode *Radopholus similis*. *Proc. Natl. Acad. Sci. U.S.A.* 111, 105–110. doi: 10.1073/pnas.1314168110
- Hölscher, D., Shroff, R., Knop, K., Gottschaldt, M., Creelius, A., Schneider, B., et al. (2009). Matrix-free UV-laser desorption/ionization (LDI) mass spectrometric imaging at the single-cell level: distribution of secondary metabolites of *Arabidopsis thaliana* and *Hypericum* species. *Plant J.* 60, 907–918. doi: 10.1111/j.1365-3113X.2009.04012.x
- Karas, M., Bachmann, D., Bahr, U., and Hillenkamp, F. (1987). Matrix-assisted ultraviolet-laser desorption of nonvolatile compounds. *Int. J. Mass Spectrom. Ion Process.* 78, 53–68. doi: 10.1016/0168-1176(87)87041-6
- Karas, M., Bachmann, D., and Hillenkamp, F. (1985). Influence of the wavelength in high-irradiance ultraviolet-laser desorption mass-spectrometry of organic-molecules. *Anal. Chem.* 57, 2935–2939. doi: 10.1021/ac00291a042
- Karas, M., and Kruger, R. (2003). Ion formation in MALDI: the cluster ionization mechanism. *Chem. Rev.* 103, 427–439. doi: 10.1021/cr010376a
- Kebarle, P., and Verkerk, U. H. (2009). Electrospray: from ions in solution to ions in the gas phase, what we know now. *Mass Spectrom. Rev.* 28, 898–917. doi: 10.1002/mas.20247
- Kertesz, V., and Van Berkel, G. J. (2010). Fully automated liquid extraction-based surface sampling and ionization using a chip-based robotic nanoelectrospray platform. *J. Mass Spectrom.* 45, 252–260. doi: 10.1002/jms.1709
- Kroiss, J., Kaltenpoth, M., Schneider, B., Schwinger, M. G., Hertweck, C., Maddula, R. K., et al. (2010). Symbiotic streptomycetes provide antibiotic combination prophylaxis for wasp offspring. *Nat. Chem. Biol.* 6, 261–263. doi: 10.1038/nchembio.331
- Laiko, V. V., Baldwin, M. A., and Burlingame, A. L. (2000). Atmospheric pressure matrix assisted laser desorption/ionization mass spectrometry. *Anal. Chem.* 72, 652–657. doi: 10.1021/ac990998k
- Lanekoff, I., Heath, B. S., Liyu, A., Thomas, M., Carson, J. P., and Laskin, J. (2012). Automated platform for high-resolution tissue imaging using nanospray desorption electrospray ionization mass spectrometry. *Anal. Chem.* 84, 8351–8356. doi: 10.1021/ac301909a
- Li, H., Smith, B. K., Mark, L., Nemes, P., Nazarian, J., and Vertes, A. (2015a). Ambient molecular imaging by laser ablation electrospray ionization mass spectrometry with ion mobility separation. *Int. J. Mass Spectrom.* 377, 681–689. doi: 10.1016/j.ijms.2014.06.025
- Li, H., Smith, B. K., Shrestha, B., Mark, L., and Vertes, A. (2015b). “Automated cell-by-cell tissue imaging and single-cell analysis for targeted morphologies by laser-ablation electrospray ionization mass spectrometry,” in *Mass Spectrometry Imaging of Small Molecules*, ed. L. He (New York, NY: Humana Press), 117–127.
- Li, Y., Shrestha, B., and Vertes, A. (2007). Atmospheric pressure molecular imaging by infrared MALDI mass spectrometry. *Anal. Chem.* 79, 523–532. doi: 10.1021/ac061577n
- McMahon, J. M., Dookeran, N. N., and Todd, P. J. (1995). Organic ion imaging beyond the limit of static secondary-ion mass-spectrometry. *J. Am. Soc. Mass Spectrom.* 6, 1047–1058. doi: 10.1016/1044-0305(95)00526-9
- Mohr, W., Vagner, T., Kuypers, M. M., Ackermann, M., and Laroche, J. (2013). Resolution of conflicting signals at the single-cell level in the regulation of cyanobacterial photosynthesis and nitrogen fixation. *PLoS ONE* 8:e66606. doi: 10.1371/journal.pone.0066606
- Nemes, P., Barton, A. A., Li, Y., and Vertes, A. (2008). Ambient molecular imaging and depth profiling of live tissue by infrared laser ablation electrospray ionization mass spectrometry. *Anal. Chem.* 80, 4575–4582. doi: 10.1021/ac8004082
- Nemes, P., Barton, A. A., and Vertes, A. (2009). Three-dimensional imaging of metabolites in tissues under ambient conditions by laser ablation electrospray ionization mass spectrometry. *Anal. Chem.* 81, 6668–6675. doi: 10.1021/ac900745e
- Nemes, P., Huang, H. H., and Vertes, A. (2012). Internal energy deposition and ion fragmentation in atmospheric-pressure mid-infrared laser ablation electrospray ionization. *Phys. Chem. Chem. Phys.* 14, 2501–2507. doi: 10.1039/c2cp23411d
- Nemes, P., and Vertes, A. (2007). Laser ablation electrospray ionization for atmospheric pressure, in vivo, and imaging mass spectrometry. *Anal. Chem.* 79, 8098–8106. doi: 10.1021/ac071181r
- Nemes, P., Woods, A. S., and Vertes, A. (2010). Simultaneous imaging of small metabolites and lipids in rat brain tissues at atmospheric pressure by laser ablation electrospray ionization mass spectrometry. *Anal. Chem.* 82, 982–988. doi: 10.1021/ac902245p
- Nielen, M. W. F., and van Beek, T. A. (2014). Macroscopic and microscopic spatially-resolved analysis of food contaminants and constituents using laser-ablation electrospray ionization mass spectrometry imaging. *Anal. Bioanal. Chem.* 406, 6805–6815. doi: 10.1007/s00216-014-7948-8
- Parsiegla, G., Shrestha, B., Carriere, F., and Vertes, A. (2012). Direct analysis of phycohemisomal antenna proteins and metabolites in small cyanobacterial populations by laser ablation electrospray ionization mass spectrometry. *Anal. Chem.* 84, 34–38. doi: 10.1021/ac202831w
- Robichaud, G., Barry, J. A., and Muddiman, D. C. (2014). IR-MALDESI mass spectrometry imaging of biological tissue sections using ice as a matrix. *J. Am. Soc. Mass Spectrom.* 25, 319–328. doi: 10.1007/s13361-013-0787-6
- Saito, N., Timberlake, C. F., Tucknott, O. G., and Lewis, I. A. S. (1983). Fast atom bombardment mass-spectrometry of the anthocyanins Violanin and Platyconin. *Phytochemistry* 22, 1007–1009. doi: 10.1016/0031-9422(83)85043-2

- Sampson, J. S., Hawkrigde, A. M., and Muddiman, D. C. (2006). Generation and detection of multiply-charged peptides and proteins by matrix-assisted laser desorption electrospray ionization (MALDESI) Fourier transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass Spectrom.* 17, 1712–1716. doi: 10.1016/j.jasms.2006.08.003
- Shrestha, B., Javonillo, R., Burns, J. R., Pirger, Z., and Vertes, A. (2013). Comparative local analysis of metabolites, lipids and proteins in intact fish tissues by LAESI mass spectrometry. *Analyst* 138, 3444–3449. doi: 10.1039/c3an00631j
- Shrestha, B., Nemes, P., Nazarian, J., Hathout, Y., Hoffman, E. P., and Vertes, A. (2010). Direct analysis of lipids and small metabolites in mouse brain tissue by AP IR-MALDI and reactive LAESI mass spectrometry. *Analyst* 135, 751–758. doi: 10.1039/b922854c
- Shrestha, B., Patt, J. M., and Vertes, A. (2011). In situ cell-by-cell imaging and analysis of small cell populations by mass spectrometry. *Anal. Chem.* 83, 2947–2955. doi: 10.1021/ac102958x
- Shrestha, B., Sripadi, P., Reschke, B. R., Henderson, H. D., Powell, M. J., Moody, S. A., et al. (2014). Subcellular metabolite and lipid analysis of *Xenopus laevis* eggs by LAESI mass spectrometry. *PLoS ONE* 9:e115173. doi: 10.1371/journal.pone.0115173
- Shrestha, B., and Vertes, A. (2009). In situ metabolic profiling of single cells by laser ablation electrospray ionization mass spectrometry. *Anal. Chem.* 81, 8265–8271. doi: 10.1021/ac901525g
- Shrivastava, K., and Setou, M. (2012). “Imaging mass spectrometry: sample preparation, instrumentation, and applications,” in *Advances in Imaging and Electron Physics*, Vol. 171, ed. P. W. Hawkes (Waltham: Academic Press), 145–193. doi: 10.1016/b978-0-12-394297-5.00004-0
- Shroff, R., Schramm, K., Jeschke, V., Nemes, P., Vertes, A., Gershenzon, J., et al. (2015). Quantification of plant surface metabolites by matrix-assisted laser desorption-ionization mass spectrometry imaging: glucosinolates on *Arabidopsis thaliana* leaves. *Plant J.* 81, 961–972. doi: 10.1111/tpj.12760
- Sripadi, P., Nazarian, J., Hathout, Y., Hoffman, E. P., and Vertes, A. (2009). In vitro analysis of metabolites from the untreated tissue of *Torpedo californica* electric organ by mid-infrared laser ablation electrospray ionization mass spectrometry. *Metabolomics* 5, 263–276. doi: 10.1007/s11306-008-0147-x
- Stolee, J. A., Shrestha, B., Mengistu, G., and Vertes, A. (2012). Observation of subcellular metabolite gradients in single cells by laser ablation electrospray ionization mass spectrometry. *Angew. Chem. Int. Edit.* 51, 10386–10389. doi: 10.1002/anie.201205436
- Stopka, S. A., Shrestha, B., Marechal, E., Falconet, D., and Vertes, A. (2014). Metabolic transformation of microalgae due to light acclimation and genetic modifications followed by laser ablation electrospray ionization mass spectrometry with ion mobility separation. *Analyst* 139, 5945–5953. doi: 10.1039/C4AN01368A
- Svatoš, A. (2010). Mass spectrometric imaging of small molecules. *Trends Biotechnol.* 28, 425–434. doi: 10.1016/j.tibtech.2010.05.005
- Vaikkinen, A., Shrestha, B., Koivisto, J., Kostianen, R., Vertes, A., and Kauppila, T. J. (2014). Laser ablation atmospheric pressure photoionization mass spectrometry imaging of phytochemicals from sage leaves. *Rapid Commun. Mass Spectrom.* 28, 2490–2496. doi: 10.1002/rcm.7043
- Vaikkinen, A., Shrestha, B., Nazarian, J., Kostianen, R., Vertes, A., and Kauppila, T. J. (2013). Simultaneous detection of nonpolar and polar compounds by heat-assisted laser ablation electrospray ionization mass spectrometry. *Anal. Chem.* 85, 177–184. doi: 10.1021/ac302432h
- Venter, A. R., Douglass, K. A., Shelley, J. T., Hasman, G., and Honarvar, E. (2014). Mechanisms of real-time, proximal sample processing during ambient ionization mass spectrometry. *Anal. Chem.* 86, 233–249. doi: 10.1021/ac4038569
- Vogel, A., and Venugopalan, V. (2003a). “Kinetics of phase transitions in pulsed IR laser ablation of biological tissues,” in *Proceedings of the Conference on Laser-Tissue Interaction XIV 4961*, San Jose, 66–74. doi: 10.1117/12.519895
- Vogel, A., and Venugopalan, V. (2003b). Mechanisms of pulsed laser ablation of biological tissues. *Chem. Rev.* 103, 577–644. doi: 10.1021/cr010379n
- Whitehouse, C. M., Dreyer, R. N., Yamashita, M., and Fenn, J. B. (1985). Electrospray interface for liquid chromatographs and mass spectrometers. *Anal. Chem.* 57, 675–679. doi: 10.1021/ac00280a023
- Yamashita, M., and Fenn, J. B. (1984). Electrospray ion-source - another variation on the free-jet theme. *J. Phys. Chem.* 88, 4451–4459. doi: 10.1021/j150664a002

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Bartels and Svatoš. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Cytological and proteomic analyses of horsetail (*Equisetum arvense* L.) spore germination

Qi Zhao^{1†}, Jing Gao^{2†}, Jinwei Suo^{2†}, Sixue Chen³, Tai Wang⁴ and Shaojun Dai^{1*}

¹ Development Center of Plant Germplasm Resources, College of Life and Environmental Sciences, Shanghai Normal University, Shanghai, China, ² Key Laboratory of Saline-alkali Vegetation Ecology Restoration in Oil Field, Ministry of Education, Alkali Soil Natural Environmental Science Center, Northeast Forestry University, Harbin, China, ³ Department of Biology, Interdisciplinary Center for Biotechnology Research, Genetics Institute, Plant Molecular and Cellular Biology Program, University of Florida, Gainesville, FL, USA, ⁴ Institute of Botany, Chinese Academy of Sciences, Beijing, China

OPEN ACCESS

Edited by:

Wagner L. Araújo,
Universidade Federal de Viçosa, Brazil

Reviewed by:

Ján A. Miernyk,
University of Missouri, USA
Kazuhiro Takemoto,
Kyushu Institute of Technology, Japan

*Correspondence:

Shaojun Dai,
College of Life and Environmental
Sciences, Shanghai Normal University,
Guilin Rd. 100, Shanghai 200234,
China
daishaojun@hotmail.com

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Plant Systems and Synthetic Biology,
a section of the journal
Frontiers in Plant Science

Received: 19 March 2015

Accepted: 29 May 2015

Published: 17 June 2015

Citation:

Zhao Q, Gao J, Suo J, Chen S, Wang
T and Dai S (2015) Cytological and
proteomic analyses of horsetail
(*Equisetum arvense* L.) spore
germination. *Front. Plant Sci.* 6:441.
doi: 10.3389/fpls.2015.00441

Spermatophyte pollen tubes and root hairs have been used as single-cell-type model systems to understand the molecular processes underlying polar growth of plant cells. Horsetail (*Equisetum arvense* L.) is a perennial herb species in Equisetopsida, which creates separately growing spring and summer stems in its life cycle. The mature chlorophyllous spores produced from spring stems can germinate without dormancy. Here we report the cellular features and protein expression patterns in five stages of horsetail spore germination (mature spores, rehydrated spores, double-celled spores, germinated spores, and spores with protonemal cells). Using 2-DE combined with mass spectrometry, 80 proteins were found to be abundance changed upon spore germination. Among them, proteins involved in photosynthesis, protein turnover, and energy supply were over-represented. Thirteen proteins appeared as proteoforms on the gels, indicating the potential importance of post-translational modification. In addition, the dynamic changes of ascorbate peroxidase, peroxiredoxin, and dehydroascorbate reductase implied that reactive oxygen species homeostasis is critical in regulating cell division and tip-growth. The time course of germination and diverse expression patterns of proteins in photosynthesis, energy supply, lipid and amino acid metabolism indicated that heterotrophic and autotrophic metabolism were necessary in light-dependent germination of the spores. Twenty-six proteins were involved in protein synthesis, folding, and degradation, indicating that protein turnover is vital to spore germination and rhizoid tip-growth. Furthermore, the altered abundance of 14-3-3 protein, small G protein Ran, actin, and coffeoyl-CoA O-methyltransferase revealed that signaling transduction, vesicle trafficking, cytoskeleton dynamics, and cell wall modulation were critical to cell division and polar growth. These findings lay a foundation toward understanding the molecular mechanisms underlying fern spore asymmetric division and rhizoid polar growth.

Keywords: spore germination, *Equisetum arvense* L., fern, proteomics, single cell, polar growth

Introduction

Sexual reproduction is crucial in plant life cycle. Spermatophyte seeds, pollen grains, and fern spores play central roles in sexual reproduction with common capability of surviving under harsh conditions upon emergence from dormancy. They have physiological resemblance with each other as to features of resurrection and polar growth during the emergence of roots, pollen tubes, and rhizoids, respectively. Previous reports have addressed the physiological and molecular characteristics of pollen and seed germination using genomics and proteomics approaches (Dai et al., 2007a,b; Tan et al., 2013), but comparable information is lacking for fern spore germination.

The similarity of germination among spermatophyte pollen grains, seeds, and fern spores includes mobilization and organization of limited reserves for polar growth within a short time frame. Previously, 33 and 123 genes were found to be commonly expressed when *Ceratopteris richardii* spores were compared with *Arabidopsis thaliana* pollen and seeds, respectively (Bushart and Roux, 2007; Salmi et al., 2007). Some proteins encoded by these genes (e.g., *Rop GTPase*, *Mago nashi*, *calmodulin 2*, *No pollen germination 1*, *phospholipase D*, *synaptobrevin*, and *constitutive photomorphogenic 9*) are mainly involved in calcium signaling, vesicle trafficking, and ubiquitin-mediated protein degradation. Although the protein functions encoded by some of the important genes (e.g., *Rop GTPase*, *calmodulin*, and *phospholipase D*) in pollen grains have been well-studied (Malhó et al., 2006), their roles in fern spore germination are not clear. In spite of polar growth similarities, pollen and fern spores are evolutionarily distinct. The pollen grains are reduced male gametophytes with two/three cells to reach ovary by tip-growing pollen tube after recognition on appropriate stigma, while the mononuclear fern spores can generate live-independent gametophyte through germination (Dai et al., 2008). Fern spore germination is somewhat more complex in that they undergo extensive cell division and differentiation during the emergences of rhizoid and prothallus under various environmental conditions (Bushart and Roux, 2007). Fern spores represent a new single-celled model for investigating asymmetric cell division, differentiation, and polar growth (Chatterjee et al., 2000; Bushart and Roux, 2007).

Previous physiological studies have reported that spore germination of more than 200 fern species was modulated by various environmental factors, such as light, gravity, calcium, phytohormones, and temperature (Suo et al., 2015). Recently, gene function analyses revealed that several signaling pathways

are crucial for fern spore germination. Phytochrome and cryptochrome signaling is important for the first spore cell mitosis (Suetsugu and Wada, 2003; Kamachi et al., 2004; Tsuboi et al., 2012). Besides, gravity and calcium signaling determines polarity establishment, cell asymmetry division, and the direction of rhizoid elongation (Chatterjee et al., 2000; Bushart and Roux, 2007; Bushart et al., 2014), and nitric oxide functions as a signal molecule in the regulation of gravity-directed fern spore polarity (Salmi et al., 2007). Moreover, gibberellin and antheridiogen can initiate and promote spore germination in many species, but abscisic acid, jasmonic acid, and ethylene have only minor promoting effects (Suo et al., 2015). In addition, the enzymes involved in glyoxylate cycle (e.g., isocitrate lyase and malate synthase) and genes encoding aconitase in the conversion from citrate to isocitrate were induced in germinating spores from *Onoclea sensibilis* (DeMaggio and Stetler, 1980), *Anemia phyllitidis* (Gemmrich, 1979), and *C. richardii* (Banks, 1999), indicating the importance of lipid reserve degradation and mobilization during spore germination. However, the molecular regulatory mechanisms in these processes are still unknown.

Germination of fern spores, especially the chlorophyllous spores (green spores), is a sophisticated signaling and metabolic process. Spores from *Equisetum* species are chlorophyll-bearing and of short viability (only a few weeks) (Ballesteros et al., 2011). *Equisetum* spores can germinate immediately under appropriate conditions with high humidity (Lebkuecher, 1997), therefore, they are good materials for studying the chlorophyllous spore germination. *Equisetum* is the oldest living genus of vascular plants, containing 15–25 extant hollow-stemmed taxa (Guillon, 2007). Most *Equisetum* species are regarded as persistent weeds in wetlands (Large et al., 2006). The reproduction of *Equisetum* species is mainly dependent on the growth of the rhizomes underground, but not the spore germination. Large-scale comparative proteomics of developing rhizomes of *Equisetum hyemale* have revealed that 1911 and 1860 proteins in rhizomes apical tip and elongation zone, respectively (Balbuena et al., 2012). However, the cellular and proteomic features of spore germination are still lacking. In the present study, we carried out cellular and 2-DE based proteomics analysis of horsetail (*Equisetum arvense* L.) spore germination to reveal the signaling and metabolic characteristics.

Materials and Methods

Collection and Germination of Mature Horsetail Spores

The mature fertile sporophylls (the separate stalks in the spring) of horsetail (*E. arvense*) were collected in suburb of Harbin (45° 27' N, 127° 52' E), Heilongjiang province, China. The mature spores (MS) were released by its elaters moving from the sporangia during dehydration at room temperature (25°C). To synchronize spore germination, the collected fresh MS were soaked in deionized H₂O overnight in the dark. Then, the rehydrated spores (RS) were transferred into a liquid germination medium (4.6 mM Ca(NO₃)₂·4H₂O, 2 mM KNO₃, 1.5 mM KH₂PO₄, 0.8 mM MgSO₄·7H₂O) and cultured in an

Abbreviations: APX, ascorbate peroxidase; ARF, ADP-ribosylation factor; CCoAOMT, caffeoyl-CoA O-methyltransferase; CDSP32, chloroplast drought-induced stress protein of 32 kDa; DAB, differentially abundant protein; DCS, double-celled spores; DHAR, dehydroascorbate reductase; EF, elongation factor; eIF, eukaryotic translation initiation factor; GO, Gene Ontology; GS, germinated spores; HAI, hours after illumination; HSP, heat shock protein; IPMDH, 3-isopropylmalate dehydrogenase; MS, mature spores; PDX, pyridoxal biosynthesis protein; Prx, peroxiredoxin; RAN, GTP-binding nuclear protein Ran; ROS, reactive oxygen species; RS, rehydrated spores; RuBisCO, ribulose-1,5-bisphosphate carboxylase/oxygenase; SPC, spores with protonemal cells; TCA, tricarboxylic acid; TCP, T-complex protein; Trx, thioredoxin; TypA, tyrosine phosphorylated protein A.

environmentally controlled chamber at $25 \pm 2^\circ\text{C}$, 24 h light at $60 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. The amount of germinating spores at various stages of RS, double-celled spores (DCS), germinated spores (GS), and spores with protonemal cells (SPC) was examined under a microscope. All these spores were collected by centrifugation at $500 \times g$ for 5 min, and used immediately after collection for protein extraction or storage at -80°C .

Observation of Spore Morphology Upon Germination

Morphological characteristics of spores upon germination were examined under a Axioskop 40 fluorescence microscope (Zeiss, Oberkochen, Germany) without or with staining using $0.2 \mu\text{g}\cdot\mu\text{L}^{-1}$ 4,6-diamidino-2-phenylindole (Molecular Probes, Carlsbad, USA). Observation of MS was performed under HITACHI S-520 scanning electron microscope (Hitachi, Tokyo, Japan) (Dai et al., 2002). Living MS were prepared by standard techniques for scanning electron microscope observation. Spores were fixed, gradually dehydrated, and critically point dried by Balt-Tec CPD-030 critical point dryer (Balt-Tec AG, Balzers, Liechtenstein) according to Dai et al. (2002). Dry spores were mounted on aluminum stubs using double-sided tape and coated with gold-palladium using a Denton Vacuum Desk II sputter coater (Denton Vacuum Inc., Cherry Hill, USA).

Protein Extraction and Quantification

For protein preparation, 0.5 g spores in the five stages of germination (MS, RS, DCS, GS, and SPC) were ground to powder in liquid nitrogen using chilled mortar and pestle. Total protein of spores was extracted according to the method of Wang et al. (2010). Protein samples were prepared independently from three different batches of plants, considered as three biological replicates. Protein concentration was determined using a Quant-Kit according to manufacturer's instructions (GE Healthcare, Salt Lake City, USA).

2-DE, Gel Staining and Protein Abundance Analysis

The protein samples were separated and visualized using 2-DE according to Dai et al. (2006). An aliquot of 1.3 mg total protein was diluted with rehydration buffer (7 M urea, 2 M thiourea, 0.5% CHAPS, 20 mM DTT, 0.5% immobilized pH gradient IPG buffer 4–7, and 0.002% bromphenol blue) to a final volume of 450 μL and loaded onto an IPG strip holder containing a 24 cm, pH 4–7 linear gradient IPG strip (GE Healthcare, Salt Lake City, USA). Isoelectric focusing was performed in the Ettan IPGphor isoelectric focusing system (GE Healthcare, Salt Lake City, USA) following the protocol of the manufacturer. For SDS-PAGE, the equilibrated IPG strips were transferred onto 12.5% acrylamide gels by using an Ettan DALT Six Electrophoresis Unit (GE Healthcare, Salt Lake City, USA). The gels were stained by Coomassie Brilliant Blue. Gel image acquisition and analysis were conducted as previously described (Wang et al., 2010). Images were acquired by scanning each stained gel using an ImageScanner (GE Healthcare, Salt Lake City, USA) at a resolution of 300 dpi and 16-bit grayscale pixel depth, and then analyzed using ImageMaster 2D version 5.0 (GE Healthcare,

Salt Lake City, USA). The experimental molecular weight of each protein was estimated by comparison with the coseparated molecular weight markers. The experimental pI of each protein was determined by its migration on IPG linear strips. For quantitative analysis, the average vol% values were calculated from three biological replicates. Protein spots with reproducible and statistically significant changes in intensity (greater than 1.5-fold and $p < 0.05$) were considered to be differentially abundant protein (DAP) spots.

Protein Identification Using Mass Spectrometry and Database Searching

The DAP spots were manually excised from the 2-DE gels, and the in-gel digestion was performed as described previously (Dai et al., 2006). Tandem mass spectrometry spectra were acquired on a ESI-Q-TOF mass spectrometry (QSTAR XL) and a ESI-Q-Trap mass spectrometry (Applied Biosystems, Foster City, USA) (Wang et al., 2010; Yu et al., 2011). The tandem mass spectrometry spectra were searched against the NCBI non-redundant protein database (<http://www.ncbi.nlm.nih.gov/>) using Mascot software (Matrix Science, London, UK), according to the searching criteria described previously (Yu et al., 2011). The taxonomic category was green plants (3,019,757 sequence entries), mass accuracy was 0.3 Da, and the maximum number of missed cleavages was set to one. To obtain highly confident identification, proteins had to meet the following criteria: (1) the top hits on the database searching report, (2) a probability-based MOWSE score greater than 43 ($p < 0.05$), and (3) more than two peptides matched with nearly complete y-ion series and complementary b-ion series. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (Vizcaíno et al., 2014) via the PRIDE partner repository with the dataset identifier PXD002218.

Protein Classification and Hierarchical Clustering Analysis

For function classification of DAPs, the peptide sequences of each DAP were aligned against the Gene Ontology (GO) protein database (<http://geneontology.org/>) following the policies and procedures provided by the GO Consortium (<http://geneontology.org/>). The proteins were classified according to their cellular component, molecular function, and biological process. Also, PSI and PHI-BLAST programs (<http://www.ncbi.nlm.nih.gov/BLAST/>) were used to search against the NCBI non-redundant protein database for protein functional domain annotation. Besides, the biological function of protein was obtained from the Kyoto Encyclopedia of Genes and Genomes pathway database (<http://www.kegg.jp/kegg/>). In addition, the conservative protein function during spore germination was predicted from previous publications on the germinating seeds and pollen. Finally, by integrative analysis of all the information collected from aforementioned processes, each DAP was classified into certain functional category defined by us. The definition of functional category is referred from literatures on pollen and seed germination (Supplementary Figure S1). Log (base 2) transformed ratios were used for hierarchical clustering analysis using Cluster 3.0 available on

the Internet (<http://bonsai.hgc.jp/~mdehoon/software/cluster/software.htm>). Protein abundance ratio was calculated as protein abundance at MS stage divided by abundance at each stage. Using a tree algorithm, these DAPs were organized based on similarities in the expression profile. These proteins can be joined by very short branches if they are very similar to each other, and by increasingly longer branches as their similarity decreases. Java TreeView (<http://jtreeview.sourceforge.net/>) was used for data visualization.

Protein Subcellular Location and Protein-protein Interaction Analysis

The subcellular location of the identified proteins was predicted using five internet tools: (1) YLoc (<http://abi.inf.uni-tuebingen.de/Services/YLoc/webloc.cgi>), confidence score ≥ 0.4 ; (2) LocTree3 (<https://roslab.org/services/loctree3/>), expected accuracy $\geq 80\%$; (3) ngLOC (<http://genome.unmc.edu/ngLOC/index.html>), probability $\geq 80\%$; (4) TargetP (<http://www.cbs.dtu.dk/services/TargetP/>), reliability class ≤ 3 ; (5) Plant-mPLOC (<http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/>), no threshold value in Plant-mPLOC. Only the consistent predictions from at least two tools were accepted as a confident result. For the inconsistent prediction results among five tools, subcellular localizations for corresponding proteins were predicted based on literatures.

The protein-protein interactions were predicted using the web-tool STRING 9.1 (<http://string-db.org>). The DAPs homologs in Arabidopsis were found by sequence BLASTing in TAIR database (<http://www.arabidopsis.org/Blast/index.jsp>). The homologs were subjected to the molecular interaction tool of STRING 9.1 for creating the proteome-scale interaction network.

Statistical Analysis

All the results were presented as means \pm standard deviation of three biological replicates. Data were analyzed by One-Way ANOVA using the statistical software SPSS 17.0 (SPSS Inc., Chicago, USA). The mean values from different stages of spore germination were compared by least significant difference *post-hoc* test. A *p*-value less than 0.05 was considered statistically significant.

Results and Discussion

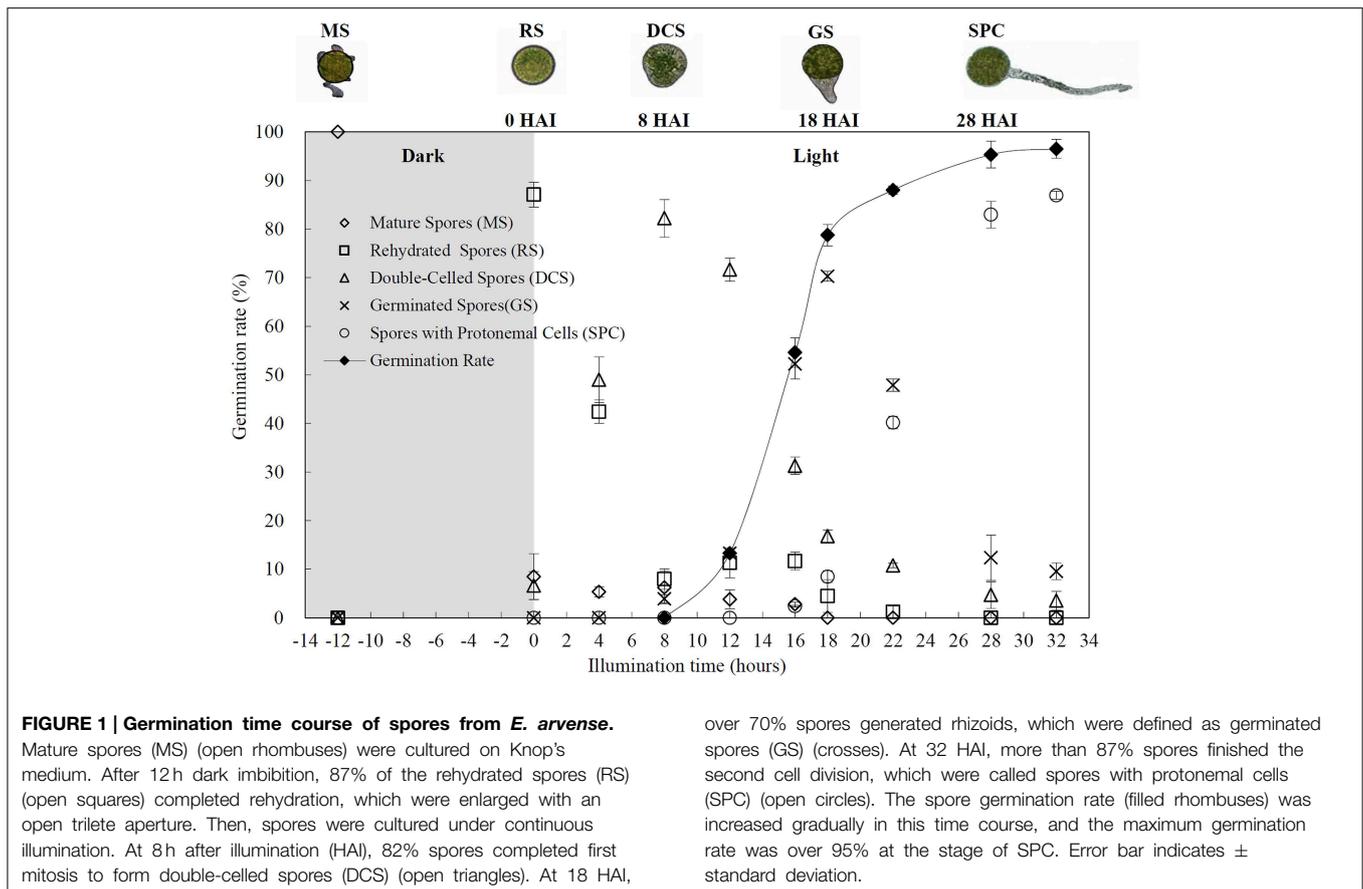
Horsetail Chlorophyllous Spore Germination Process

Chlorophyllous spores from fern species in a few unrelated taxa in Pteridophyta can germinate in less than 3 days and their viability lasts about 1 year or less (Lloyd and Klekowski, 1970). *Equisetum* spores are typical chlorophyllous spores with short viability, which can germinate immediately under humid conditions and remain viable for about 2 weeks (Lebkuecher, 1997). In this study, the spores were sown and cultured on Knop's medium under two different illumination levels (i.e., $30 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ and $60 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) after 12 h dark imbibition in deionized H₂O at room temperature. We found that the light illumination level has obvious effects on horsetail spore germination rate. At 16 h after illumination (HAI), the

germination rate of spores under higher illumination level ($60 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) was over 50%, but the germination rate of spores under lower illumination level ($30 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) was only 20%. The spores under the higher light intensity reached to the maximum germination rate of 95% at 32 HAI, but the maximum germination rate of spores under $30 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ was only 35% (Supplementary Figure S2). Thus, the $60 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ light was used for horsetail spore germination. Actually, the horsetail MS have initiated germination during the process of dark imbibition. Eighty-seven percent of the RS finished their cell nucleus polar migration toward the gravity, and the nucleus of spore started their division at 12 HAI for the preparation of asymmetry cell mitosis (Figure 1). At 8 HAI, 82% spores finished the first mitosis to generate DCS containing a larger cell and a smaller cell, and the smaller cell of 4% spores elongated to form an obvious rhizoid. At this time point, we started to calculate the spore germination rate that is defined as the ratio of GS number to total spore number. At 18 HAI, more than 70% of spores generated the rhizoids, being defined as GS. At 32 HAI, spores obtained the maximum germination rate of over 95%, and the larger cell (protonemal cell) of over 87% spores have finished the second mitosis to give rise to the photosynthetic prothallus. The spores at this stage were defined as SPC (Figure 1).

Cytological Characteristics of Horsetail Germinating Spores

The horsetail chlorophyllous spores were generated in sporangium from strobili. The MS were released from partially dried strobilus. Horsetail spores are unusual in the morphological and physiological aspects. The spores are typically about 25–35 μm in diameter with four flexible ribbon-like elaters about 100 μm each (Figures 2A,E,I,Q). The elaters initially wrapping around the spore body can deploy upon dehydration and fold back in humid air. The elater movement driven by humidity variations led to the spore exiting from sporangium, and especially can catch the wind again when they were jumping from the ground, which is believed to be a novel type of efficient self-propelled dispersal mechanism (Marmottant et al., 2013). The single nuclear was visible clearly in the center of the spores (Figures 2E,I). The elaters were lost when spores were cultured in liquid medium and cell nucleus has migrated toward gravity (Figures 2N,P). After 12 h dark imbibition, the RS were swelled to the diameter of about 50 μm . In RS, the nucleus has completed the first mitosis and two nuclei in the bottom side of the spore can be clearly observed (Figures 2B,E,J). Subsequently, at 8 HAI, the novel cell wall was formed between the two nuclei to generate an asymmetry DCS, containing a larger cell and a smaller cell (Figures 2C,G,K). The smaller cell started to elongate out of the spore and an approximate 30 μm long rhizoid emerged at 18 HAI to form GS (Figures 2D,H,L). In the GS, the cell nucleus of rhizoid still left inside the spores (Figures 2H,L). As rhizoid elongated, its nucleus was moving outside from the spore to the center of the rhizoid at 28 HAI (Figures 2O,R). At the same time, the larger cell inside the spore can complete the second mitosis to form a new protonemal cell (Figures 2O,R). Thus, the spore finish its germination by the formation of



spores with a rhizoid cell and two protonemal original cells (Figures 2M,O,R).

DAPs Upon Spore Germination

To determine DAPs in mature and germinating spores, protein samples collected at the five stages (MS, RS, DCS, GS and SPC) were subjected to 2-DE analysis. On the Coomassie Brilliant Blue-stained gels (pH 4–7, 24 cm IPG strip), 1243 \pm 47, 1247 \pm 34, 1254 \pm 31, 1229 \pm 57, and 1234 \pm 40 spots from MS, RS, DCS, GS, and SPC were detected, respectively (Figure 3, Supplementary Figure S3). Among them, 139 protein spots showed differential abundances in five distinct stages of spore germination (>1.5 -fold, $p < 0.05$). A total of 131 spots were identified using tandem mass spectrometry and Mascot database searching, and eight spots were not matched in database. Among the proteins, 28 spots contained only single peptide match, which were considered as un-identified according to our criteria. In the remaining 103 spots, 80 spots contained a single protein each (Table 1, Supplementary Table S1) and 23 spots contained more than one protein each (Supplementary Table S2). Thus, the 80 proteins were DAPs during spore germination.

The 80 DAPs were classified into eleven groups, including photosynthesis (17), carbohydrate and energy metabolism (9), other metabolisms (5), signaling and vesicle trafficking (8), cell

structure (5), cell cycle (2), transcription related (1), protein synthesis (10), protein folding and processing (8), protein degradation (8), stress and defense (7) (Table 1). Among them, proteins involved in photosynthesis (21%) and protein synthesis (13%) were over-represented (Figure 4A), indicating active photosynthesis and *de novo* protein synthesis are pivotal for the germinating chlorophyllous spores.

Interestingly, the 80 DAPs represented 48 unique proteins. Thirteen proteins (16.3%) had multi-proteoforms, which were mainly involved in photosynthesis, tricarboxylic acid (TCA) cycle, vesicle trafficking, protein synthesis, folding and turnover, as well as reactive oxygen species (ROS) scavenging (Table 1, Supplementary Table S3). These proteoforms might be generated from alternative splicing and various post-translational modifications.

The subcellular localization of the 80 DAPs was predicted based on five internet tools (i.e., YLoc, LocTree3, ngLOC, TargetP, and Plant-mPLOC) and literature. In total, 32 DAPs were predicted to be localized in chloroplast, 27 in cytoplasm, one in Golgi apparatus, two in mitochondria, and six in nucleus. Besides, seven proteins were predicted to be localized in two organelles, and five proteins in four organelles (Figure 4B, Supplementary Table S4).

To better understand the protein expression characteristics during the five stages, hierarchical clustering analysis was

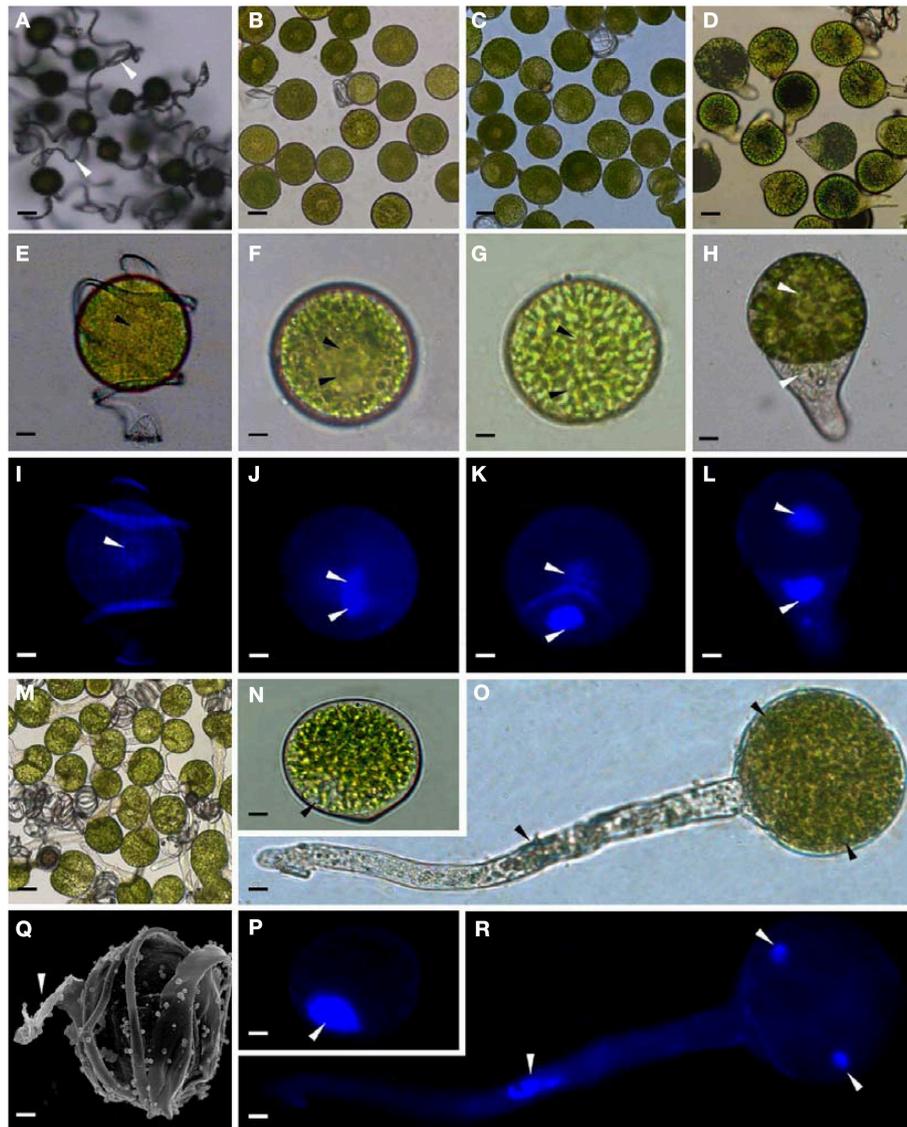


FIGURE 2 | Morphology of mature and germinating spores from *E. arvense*. (A,E,I,N,P,Q) Mature spores (MS). (A) MS groups, arrows show elaters, bar = 20 μm ; (E,I) nucleus center-localized MS with elaters under bright-field microscope (BFM) and fluorescence microscope (FM), arrows show nuclei, bar = 5 μm ; (N,P) nucleus-migrated MS without elaters under BFM and FM, arrows show nuclei, bar = 5 μm ; (Q) MS with elaters under scanning electron microscope, arrows show elaters, bar = 8 μm . (B,F,J)

Rehydrated spores (RS). (B) RS groups, bar = 15 μm ; (F,J) RS under BFM and FM, arrows show nuclei, bar = 5 μm . (C,G,K) Double-celled spores (DCS). (C) DCS groups, bar = 15 μm ; (G,K) DCS under BFM and FM, arrows show nuclei, bar = 5 μm . (D,H,L) Germinated spores (GS). (D) GS groups, bar = 15 μm ; (H,L) GS under BFM and FM, arrows show nuclei, bar = 5 μm . (M,O,R) Spores with protonemal cells (SPC). (M) SPC groups, bar = 15 μm ; (O,R) SPC under BFM and FM, arrows show nuclei, bar = 5 μm .

applied to the 80 proteins, which revealed three main clusters. Cluster I contained 21 proteins, which were not expressed in MS, but induced in other stages of germination. They were involved in photosynthesis, glycolysis, protein folding and turnover, signal transduction, and ROS scavenging, implying these pathways were specially induced in certain germination stages (Figure 4C). Cluster II included 35 proteins that were mainly decreased during spore germination. These proteins were divided into two subclusters. Subcluster II-1 contained the proteins mainly decreased during germination,

and subcluster II-2 included proteins induced at the stage of GS but obviously reduced in SPC. The proteins in cluster II covered ten function categories, in addition to category of transcription (Figure 4D). The rest 24 proteins were grouped into cluster III, representing increased proteins during spore germination. Cluster III contained two subclusters. Subcluster III-1 contained five proteins induced in stages of RS and DCS, but reduced in GS and SPC, and subcluster III-2 included 19 proteins significantly increased at all stages of germination (Figure 4D).

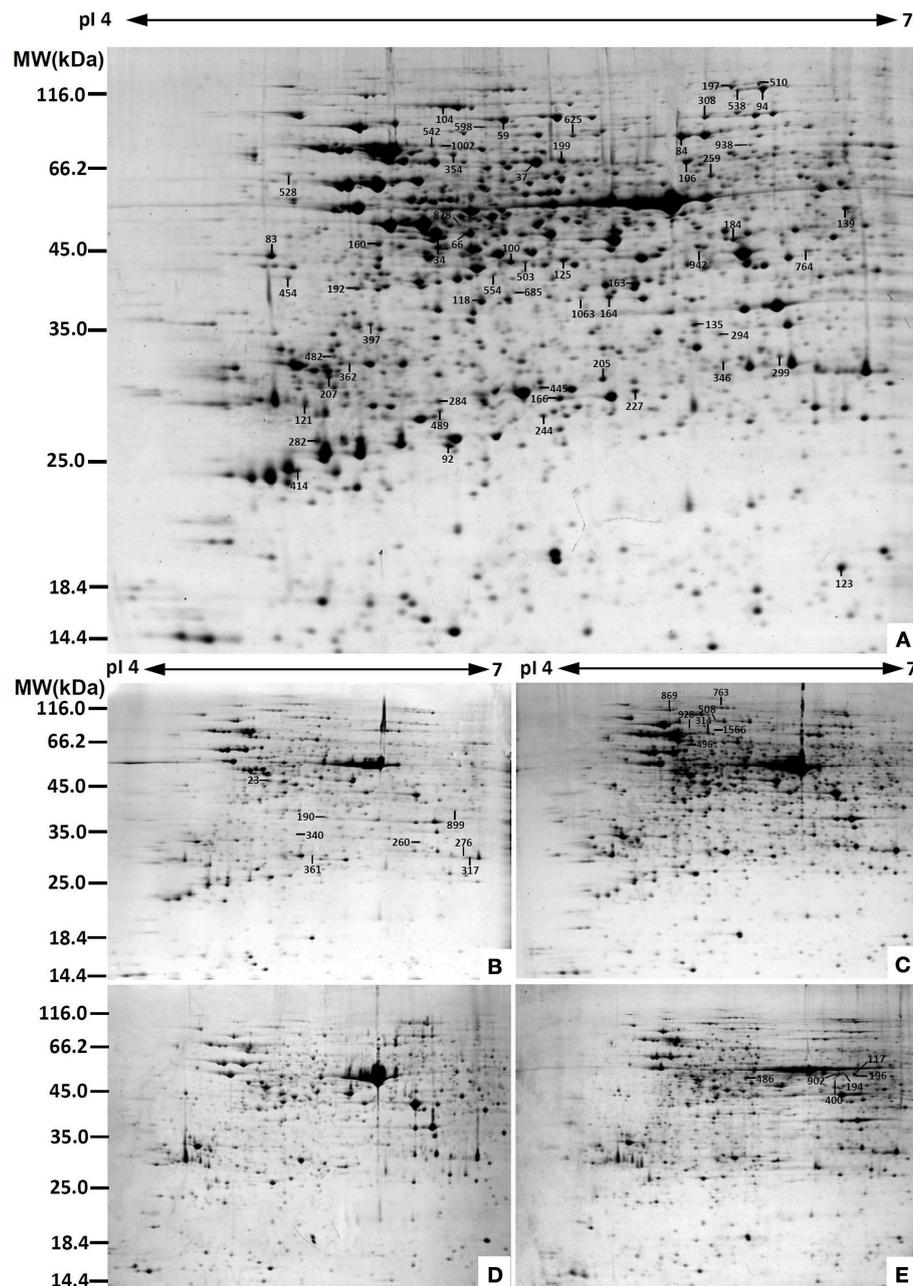


FIGURE 3 | Representative 2-DE images of proteins in various germination stages from *E. arvense* spores. (A) Mature spores. (B) Rehydrated spores. (C) Double-celled spores. (D) Germinated spores. (E) Spores with protonemal cells. Proteins were separated on 24 cm IPG strips (pH 4–7 linear gradient) using IEF in the first dimension, followed by 12.5% SDS-PAGE gels in the second dimension. The 2-DE gel was

stained with Coomassie Brilliant Blue. A total of 80 differentially abundant proteins identified by ESI-Q-TOF and ESI-Q-Trap tandem mass spectrometry are marked with numbers on the gels. Molecular weight (MW) in kDa and pI of proteins are indicated on the left and top of the gels, respectively. Detailed information can be found in **Table 1** and Supplementary Table S1.

Photosynthesis and Reserve Mobilization are Active in the Germinating Spores

Mature chlorophyllous spores of *Equisetum* species contain chloroplasts and water, making them metabolically active and short-lived. After released from sporangium, the green spores

can survive desiccation for less than 2 weeks (Lebkuecher, 1997). It has been found that the *E. hyemale* spores can tolerate desiccation of 2% relative humidity for 24 h. Upon rehydration, they can rapidly regain photosynthetic competence (Lebkuecher, 1997). In this study, we found the viability of

TABLE 1 | Differentially abundant proteins during *E. arvense* spore germination.

Spot No. ^(a)	Protein name ^(b)	Subcellular location ^(c)	Plant species ^(d)	Gi No. ^(e)	Thr. MW(Da) /pI ^(f)	Exp. MW(Da) /pI ^(g)	Cov (%) ^(h)	Sco ⁽ⁱ⁾	QM ^(j)	V% ± SD ^(k) MS RS DCS GS SPC
PHOTOSYNTHESIS (17)										
121	Chlorophyll a/b binding protein (CAB)	Chl	<i>Hedera helix</i>	12,582	20,759/ 4.83	25,376/ 4.71	10	57	3	
346	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (RBCL)	Chl	<i>Eryngium bourgatii</i>	1,292,976	53,093/ 5.56	28,686/ 6.28	6	59	3	
763	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (RBCL)	Chl	<i>Donatia fascicularis</i>	1,304,292	49,896/ 6.32	134,736/ 5.48	8	55	5	
902	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (RBCL)	Chl	<i>Grammitis diminita</i>	340,031,166	45,938/ 6.26	50,502/ 6.31	5	61	2	
117	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (RBCL)	Chl	<i>Equisetum telmateia</i>	16,565,336	48,750/ 6.26	52,974/ 6.47	15	57	7	
196	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (RBCL)	Chl	<i>E. telmateia</i>	16,565,336	48,750/ 6.26	51,502/ 6.47	16	52	6	
194	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (RBCL)	Chl	<i>E. telmateia</i>	16,565,336	48,750/ 6.26	51,418/ 6.33	13	53	6	
400	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (RBCL)	Chl	<i>E. telmateia</i>	16,565,336	48,750/ 6.26	49,452/ 6.31	11	53	5	
23	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (RBCL)	Chl	<i>E. bourgatii</i>	1,292,976	53,093/ 5.56	50,872/ 5.02	4	78	2	
528	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (RBCL)	Chl	<i>Equisetum arvense</i>	1,352,773	52,493/ 5.86	67,620/ 4.64	7	60	3	
100	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (RBCL)	Chl	<i>Isoetes capensis</i>	83,032,384	47,561/ 6.30	42,425/ 5.47	6	55	3	
496	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit-binding protein subunit beta (RBP)	Chl	<i>Solanum lycopersicum</i>	460,379,814	64,527/ 5.46	71,249/ 5.21	4	61	2	
34	Ribulose-1,5-bisphosphate carboxylase/oxygenase activase (RCA)	Chl	<i>Gossypium hirsutum</i>	12,620,883	48,609/ 5.06	48,999/ 5.19	11	98	4	
184	Ribulose-1,5-bisphosphate carboxylase/oxygenase activase (RCA)	Chl	<i>Hordeum vulgare</i>	100,614	47,496/ 5.64	47,356/ 6.31	11	94	4	

(Continued)

TABLE 1 | Continued

Spot No. ^(a)	Protein name ^(b)	Subcellular location ^(c)	Plant species ^(d)	Gi No. ^(e)	Thr. MW(Da) /pI ^(f)	Exp. MW(Da) /pI ^(g)	Cov (%) ^(h)	Sco ⁽ⁱ⁾	QM ^(j)	V% ± SD ^(k) MS RS DCS GS SPC
125	Ribulose-1,5-bisphosphate carboxylase/oxygenase activase (RCA)	Chl	<i>Musa acuminata</i> subsp. <i>malaccensis</i>	695,062,479	47,898/6.18	42,859/5.67	8	75	3	
308	Transketolase (TK)	Chl	<i>Spinacia oleracea</i>	2,529,342	80,744/6.20	95,817/6.20	8	156	5	
84	Transketolase (TK)	Chl	<i>S. lycopersicum</i>	460,388,792	80,615/6.26	85,056/6.11	8	90	7	
CARBOHYDRATE AND ENERGY METABOLISM (9)										
685	Unknown, pyruvate dehydrogenase E1 component subunit beta* (PDH)	Chl	<i>Glycine max</i>	255,647,166	44,596/6.28	27,410/6.46	4	60	2	
160	Malate dehydrogenase (MDH)	#Chl, Cyt, Mit, Pox	<i>Arabidopsis thaliana</i>	15,219,721	35,890/6.11	46,411/4.98	12	74	3	
503	Malate dehydrogenase (MDH)	#Chl, Cyt, Mit, Pox	<i>A. thaliana</i>	15,219,721	35,890/6.11	42,425/5.53	6	57	2	
260	Malate dehydrogenase (MDH)	#Chl, Cyt, Mit, Pox	<i>A. thaliana</i>	11,133,509	35,548/6.11	29,729/6.41	6	59	3	
164	Malate dehydrogenase (MDH)	#Chl, Cyt, Mit, Pox	<i>Beta vulgaris</i> subsp. <i>vulgaris</i>	731,361,010	41,677/5.74	36,353/5.84	12	193	4	
163	Malate dehydrogenase (MDH)	Chl	<i>Brachypodium distachyon</i>	357,147,942	41,864/6.97	38,560/5.94	10	60	4	
486	Enolase	Cyt	<i>Tarenaya hassleriana</i>	729,317,446	51,639/5.91	48,823/5.62	3	54	2	
489	6-phosphogluconate dehydrogenase (6-PGDH)	#Chl, Cyt	<i>G. max</i>	356,513,305	54,116/6.25	24,817/5.21	12	53	5	
192	Fructokinase (FK)	Chl	<i>Lycopersicon esculentum</i>	23,476,263	40,620/5.41	37,710/4.91	8	129	3	
OTHER METABOLISMS (5)										
135	Enoyl-acyl carrier protein reductase (EAR)	Chl	<i>Oryza sativa</i> subsp. <i>japonica</i>	75,225,229	39,277/8.81	32,651/6.16	5	21	2	
197	Hypothetical protein PHAVU_001G035500g, glycine decarboxylase* (GDC)	Mit	<i>Phaseolus vulgaris</i>	593,795,946	116,167/6.65	117,088/6.30	11	157	12	
362	3-isopropylmalate dehydrogenase (IPMDH)	Chl	<i>A. thaliana</i>	15,241,338	44,305/5.75	28,512/4.87	23	67	7	

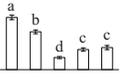
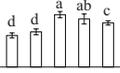
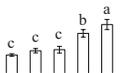
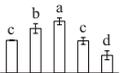
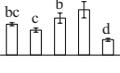
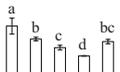
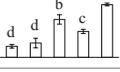
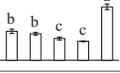
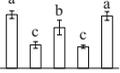
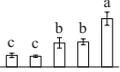
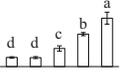
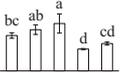
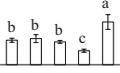
(Continued)

TABLE 1 | Continued

Spot No. ^(a)	Protein name ^(b)	Subcellular location ^(c)	Plant species ^(d)	Gi No. ^(e)	Thr. MW(Da) /pI ^(f)	Exp. MW(Da) /pI ^(g)	Cov (%) ^(h)	Sco ⁽ⁱ⁾	QM ^(j)	V% ± SD ^(k) MS RS DCS GS SPC
445	Hypothetical protein SELMODRAFT_406755, containing cd00517 ATP-sulfurylase domain* (ATPS)	Chl	<i>Selaginella moellendorffii</i>	302,763,978	56,179/6.69	26,370/5.59	5	60	3	
190	Predicted protein, pyridoxal biosynthesis protein PDX1*	Cyt	<i>Physcomitrella patens</i>	168,019,502	33,769/6.03	29,893/5.82	15	324	8	
SIGNALING AND VESICLE TRAFFICKING (8)										
282	Unknown, containing PLN02804 chalcone isomerase domain* (CHI)	Chl	<i>Picea sitchensis</i>	116,784,316	23,688/5.23	23,141/4.75	7	56	2	
1063	Hypothetical protein SELMODRAFT_151778, containing cd00200 WD40 domain* (WD40)	Nuc	<i>S. moellendorffii</i>	302,791,020	37,267/5.65	36,189/5.74	6	84	2	
207	14-3-3 protein	Cyt	<i>S. oleracea</i>	440,573,600	30,029/4.84	27,277/4.79	16	79	6	
276	Unknown, containing cd00877 Ran GTPase domain* (RAN)	Nuc	<i>P. sitchensis</i>	116,794,384	25,374/6.30	27,727/6.65	20	52	4	
317	GTP-binding nuclear protein Ran/TC4 (RAN)	Nuc	<i>Vicia faba</i>	585,783	25,274/6.39	27,750/6.73	27	187	6	
340	Unknown, containing cd00877 Ran GTPase domain* (RAN)	Nuc	<i>P. sitchensis</i>	116,794,384	25,374/6.30	30,115/5.59	24	53	5	
764	Unknown, containing cd00877 Ran GTPase domain* (RAN)	Nuc	<i>P. sitchensis</i>	116,794,384	25,374/6.30	44,900/6.58	30	51	6	
123	ADP-ribosylation factor (ARF)	Gol	<i>Chlamydomonas reinhardtii</i>	1,703,374	20,747/6.92	19,121/6.71	29	92	4	
CELL STRUCTURE (5)										
397	Os03g0718100, actin*	Cyt	<i>O. sativa</i> subsp. <i>japonica</i>	115,454,971	42,014/5.30	32,585/4.95	9	82	3	
554	Reversibly glycosylated polypeptide (RGP)	Cyt	<i>Ricinus communis</i>	223,546,230	41,557/5.82	39,973/5.41	14	105	5	
118	Reversibly glycosylated polypeptide (RGP)	Cyt	<i>Solanum tuberosum</i>	34,582,499	42,146/5.71	35,827/5.36	24	74	8	
299	Unknown protein, rhamnose biosynthetic enzyme 1* (RBE)	#Cyt	<i>A. thaliana</i>	8,493,590	33,861/5.73	29,119/6.48	7	127	4	
482	Caffeoyl-CoA O-methyltransferase (CCoAOMT)	#Cyt	<i>Eucalyptus gunnii</i>	3,023,419	28,010/5.02	29,093/4.81	11	73	2	

(Continued)

TABLE 1 | Continued

Spot No. ^(a)	Protein name ^(b)	Subcellular location ^(c)	Plant species ^(d)	Gi No. ^(e)	Thr. MW(Da) /pI ^(f)	Exp. MW(Da) /pI ^(g)	Cov (%) ^(h)	Sco ⁽ⁱ⁾	QM ^(j)	V% ± SD ^(k) MS RS DCS GS SPC
CELL CYCLE (2)										
104	Cell division cycle protein 48 homolog (CDC48)	Cyt	<i>T. hassleriana</i>	729,396,339	89,888/5.09	103,252/5.22	21	119	14	
139	Proliferation-associated protein 2G4-like (PA2G4)	Nuc	<i>Nicotiana tomentosiformis</i>	697,180,533	43,837/5.96	55,758/6.73	5	52	2	
TRANSCRIPTION RELATED (1)										
938	Predicted protein, containing cd00771 threonyl-tRNA synthetase class II core catalytic domain* (ThrRS)	Cyt	<i>P. patens</i>	168,012,416	71,002/5.96	80,637/6.37	3	54	2	
PROTEIN SYNTHESIS (10)										
83	Predicted protein, 40S ribosomal protein SA (RPSA)	Cyt	<i>P. patens</i>	168,017,628	31,539/4.72	43,804/4.58	7	101	2	
454	Hypothetical protein AMTR_s00087p00135370, containing cd08065 eukaryotic translation initiation factor 3* (eIF3)	#Chl, Cyt, Mit, Nuc	<i>Amborella trichopoda</i>	548,842,775	44,710/5.17	39,587/4.64	5	82	2	
66	Hypothetical protein, eukaryotic initiation factor 4A* (eIF4A)	Cyt	<i>P. patens</i>	168,026,095	47,119/5.46	48,999/5.32	25	162	11	
878	Eukaryotic initiation factor 4A (eIF4A)	Cyt	<i>Nicotiana tabacum</i>	1,170,511	47,098/5.37	49,134/5.23	14	118	6	
94	OSJNBa0020P07.3, elongation factor 2* (EF2)	Cyt	<i>O. sativa</i> subsp. <i>japonica</i>	38,344,860	94,939/5.85	115,425/6.42	9	111	7	
538	Hypothetical protein SELMODRAFT_411087, elongation factor 2* (EF2)	Cyt	<i>S. moellendorffii</i>	302,773,640	94,568/6.00	114,195/6.33	5	167	5	
510	Hypothetical protein SELMODRAFT_143627, elongation factor G* (EF-G)	#Chl, Mit	<i>S. moellendorffii</i>	302,765,284	75,483/5.31	119,422/6.41	6	75	4	
942	Elongation factor Tu (EF-Tu)	Mit	<i>Setaria italica</i>	514,812,465	48,530/5.99	44,724/6.18	16	63	8	
205	Uncharacterized protein LOC105056625, containing cd14275 elongation factor Ts domain* (EF-Ts)	Chl	<i>Elaeis guineensis</i>	743,840,139	125,269/4.92	27,102/5.82	2	130	5	
625	Tyrosine phosphorylated protein A (TypA)	Chl	<i>Suaeda salsa</i>	162,424,768	75,474/6.71	85,342/5.71	6	205	4	

(Continued)

TABLE 1 | Continued

Spot No. ^(a)	Protein name ^(b)	Subcellular location ^(c)	Plant species ^(d)	Gi No. ^(e)	Thr. MW(Da) /pI ^(f)	Exp. MW(Da) /pI ^(g)	Cov (%) ^(h)	Sco ⁽ⁱ⁾	QM ^(j)	V% ± SD ^(k) MS RS DCS GS SPC
PROTEIN FOLDING AND PROCESSING (8)										
542	Hypothetical protein SELMODRAFT_440382, containing pfam00012 heat shock protein 70 domain* (HSP70)	Cyt	<i>S. moellendorffii</i>	302,770,212	71,931/5.17	80,237/5.18	13	295	7	
1002	Heat shock protein 70 (HSP70)	Cyt	<i>Populus trichocarpa</i>	224,098,390	71,620/5.14	80,105/5.21	16	135	8	
1566	Heat shock protein 70 (HSP70)	Cyt	<i>Petunia × hybrida</i>	20,559	71,137/5.07	106,146/6.34	4	92	2	
314	Heat shock protein 70 (HSP70)	Cyt	<i>Dactylis glomerata</i>	188,011,548	72,002/5.03	86,310/5.38	21	338	13	
928	Heat shock protein 70 like protein (HSP70)	Cyt	<i>S. lycopersicum</i>	460,394,037	72,308/5.16	91,122/5.25	7	122	5	
598	Predicted protein, containing pfam00183 heat shock protein 90 domain* (HSP90)	Cyt	<i>P. patens</i>	168,034,606	79,652/4.93	87,379/5.25	5	89	3	
259	T-complex protein 1 subunit alpha (TCP1α)	Cyt	<i>A. thaliana</i>	135,535	59,477/5.93	66,881/6.22	8	246	4	
106	Hypothetical protein VITISV_000290, T-complex protein 1 subunit gamma* (TCP1γ)	Cyt	<i>V. vinifera</i>	147,784,740	61,064/6.06	72,989/6.14	5	94	3	
PROTEIN DEGRADATION (8)										
284	Alpha7 proteasome subunit (PSA7)	Cyt, Nuc	<i>N. tabacum</i>	14,594,925	27,466/6.11	25,449/5.20	8	79	2	
199	Zinc dependent protease (ZDP)	Chl	<i>Trifolium pratense</i>	84,468,286	74,746/5.82	73,106/5.66	12	106	7	
227	Zinc dependent protease (ZDP)	Chl	<i>T. pratense</i>	84,468,286	74,746/5.82	26,150/5.94	10	87	7	
37	Zinc dependent protease (ZDP)	Chl	<i>T. pratense</i>	84,468,286	74,746/5.82	72,520/5.57	5	93	3	
869	Zinc metalloprotease (ZMP)	Chl, Mit	<i>A. thaliana</i>	10,120,424	121,539/5.39	131,816/5.10	1	44	2	
354	Predicted protein, containing pfam06480 FtsH extracellular domain* (FtsH)	Chl	<i>Micromonas pusilla</i> CCMP1545	303,275,720	77,421/5.29	75,882/5.26	6	113	4	

(Continued)

TABLE 1 | Continued

Spot No. ^(a)	Protein name ^(b)	Subcellular location ^(c)	Plant species ^(d)	Gi No. ^(e)	Thr. MW(Da) /pI ^(f)	Exp. MW(Da) /pI ^(g)	Cov (%) ^(h)	Sco ⁽ⁱ⁾	QM ^(j)	V% ± SD ^(k)				
										MS	RS	DCS	GS	SPC
508	Predicted protein, ATP-dependent zinc metalloprotease FtsH* (FtsH)	Chl	<i>P. patens</i>	168,001,910	68,933/5.23	108,174/5.50	3	75	2					
59	ATP-dependent Clp protease ATP-binding subunit ClpC (CLPC)	Chl	<i>A. thaliana</i>	9,758,239	103,616/6.36	94,506/5.45	9	415	9					
STRESS AND DEFENSE (7)														
414	Hypothetical protein MIMGU_mgv1a021611mg, containing cd03015 2-cys peroxiredoxin domain* (Prx)	Chl, Cyt	<i>Erythranthe guttata</i>	604,334,612	21,153/4.98	21,812/4.68	17	132	4					
92	2-cys peroxiredoxin-like protein (Prx)	Chl, Cyt	<i>Hyacinthus orientalis</i>	47,027,073	21,956/4.93	22,930/5.24	12	58	3					
166	Ascorbate peroxidase (APX)	Cyt	<i>Eucalyptus grandis</i>	702,241,628	27,613/6.07	25,671/5.66	7	104	2					
244	Dehydroascorbate reductase-like protein (DHAR)	Cyt	<i>S. tuberosum</i>	76,573,291	23,610/6.32	24,534/5.60	11	147	3					
361	Dehydroascorbate reductase-like protein (DHAR)	Cyt	<i>S. tuberosum</i>	76,573,291	23,610/6.32	24,783/5.64	11	105	3					
899	Dehydroascorbate reductase-like protein (DHAR)	Cyt	<i>S. tuberosum</i>	76,160,951	23,596/6.09	44,783/6.60	11	92	3					
294	Chloroplast drought-induced stress protein of 32 kDa (CDSP32)	Chl, Cyt	<i>S. tuberosum</i>	2,582,822	33,779/8.07	31,436/6.26	8	87	3					

^aAssigned spot number as indicated in **Figure 3**.

^bThe name and functional category of the proteins identified by ESI-Q-TOF and ESI-Q-Trap tandem mass spectrometry. Protein names marked with an asterisk (*) have been edited by us depending on searching against NCBI non-redundant protein database for functional domain. The abbreviations for the protein names are indicated in the bracket after protein names.

^cProtein subcellular localization predicted by softwares (YLoc, LocTree3, Plant-mPLoc, ngLOC, and TargetP). Only the consistent predictions from at least two tools were accepted as a confident result. Pounds (#) indicate prediction results were inconsistent among five tools. The subcellular localizations were predicted based on literature listed in Supplementary Table S4. Chl, chloroplast; Cyt, cytoplasm; Gol, Golgi apparatus; Mit, mitochondria; Nuc, nucleus; Pox, peroxisome.

^dThe plant species that the peptides matched from.

^eDatabase accession number from NCBI non-redundant protein database.

^{f,g}Theoretical (f) and experimental (g) molecular weight (Da) and pI of identified proteins. Theoretical values were retrieved from the protein database. Experimental values were calculated using ImageMaster 2D version 5.0.

^hThe amino acid sequence coverage for the identified proteins.

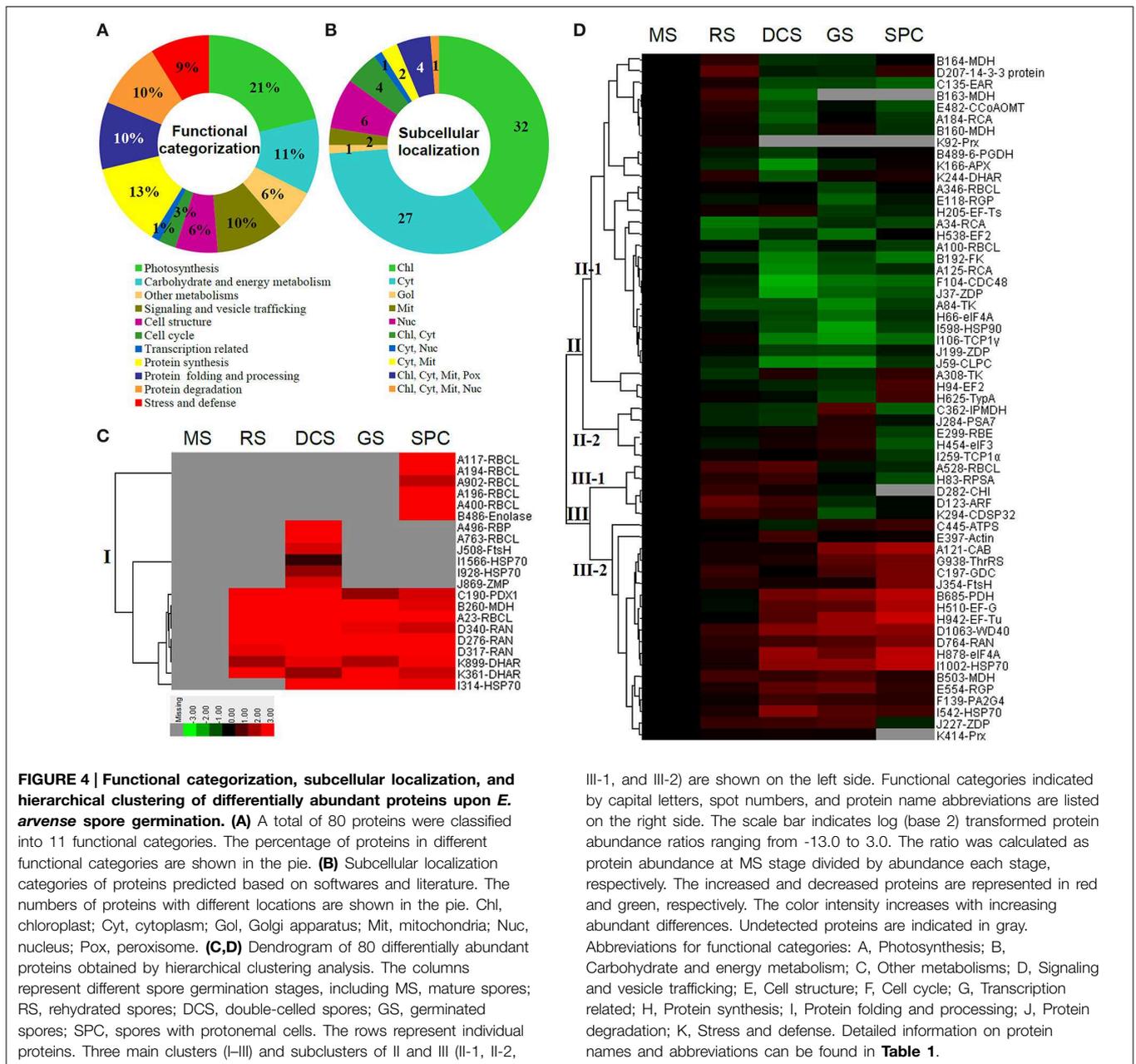
ⁱThe Mascot score obtained after searching against the NCBI non-redundant protein database.

^jThe number of matched peptides for each protein.

^kThe mean values of protein spot volumes relative to total volume of all the spots. Five spore germination stages, MS, mature spores; RS, rehydrated spores, DCS, double-celled spores, GS, germinated spores; and SPC, spores with protonemal cells were performed. Error bar indicates ± standard deviation (SD). Letters indicate statistically significant differences ($p < 0.05$) among five stages of spore germination as determined by One-Way ANOVA.

spores from *E. arvense* lasted less than 3 weeks, and the fresh spores germinated in 32 h (**Figure 1**). The extremely short viability and rapid germination were mainly due to active photosynthesis and respiration in the chlorophyllous spores. This is different from non-green spores with long

dormancy and viability (Lloyd and Klekowski, 1970). Upon germination, chlorophyllous spores also exhibited features distinct from the non-green spores. In our proteomic results, we found that 17 DAPs were photosynthesis-related proteins, including a chlorophyll a/b-binding protein, 10 proteoforms of



ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) large subunit, a RuBisCO large subunit binding protein, three proteoforms of RuBisCO activase, and two proteoforms of chloroplast transketolase (**Table 1**). The increased chlorophyll a/b-binding protein indicated that the photosystem II was enhanced upon spore germination. The multiple proteoforms of RuBisCO large subunit, RuBisCO activase, and transketolase changed in abundance possibly due to protein phosphorylation (Guitton and Mache, 1987). This implies that the carbon assimilation is active in germinating spores. Importantly, we also found nine enzymes involved in carbohydrate and energy metabolism. They were six TCA cycle enzymes (i.e., a

pyruvate dehydrogenase and five malate dehydrogenases), an enolase involved in glycolysis, a cytosolic 6-phosphogluconate dehydrogenase (an enzyme in pentose phosphate pathway), and a fructokinase in charge of sucrose and fructose metabolism. All these enzymes, taking up 32% of DAPs, were considered as key enzymes for carbon and energy supplies during spore germination. Their variations also indicate that the heterotrophic metabolism is crucial for chlorophyllous spore germination.

We also found several proteins involved in fatty acid synthesis, amino acid metabolism, sulfur assimilation, and secondary metabolism (**Table 1**). Among them, enoyl-acyl carrier protein reductase is a key enzyme of the type II fatty acid synthesis system

in plastids. Enoyl-acyl carrier protein reductase was proved to be important for fatty acid deposition in developing seeds (de Boer et al., 1999) and pollen grains (Poghosyan et al., 2005). In our results, enoyl-acyl carrier protein reductase was decreased at the stages of DCS, GS and SPC during germination. This indicates that fatty acid synthesis is reduced upon spore germination, which is consistent with the notion that mobilization of storage lipids has been triggered for energy supply in germinating spores (DeMaggio and Stetler, 1985). Besides, we found three increased enzymes at certain stages involved in amino acid metabolism, including glycine decarboxylase, 3-isopropylmalate dehydrogenase (IPMDH), and ATP-sulfurylase. Glycine decarboxylase participates in glycine, serine and threonine metabolism, and was found to be abundant in mitochondria of C3 leaves that functions in photorespiratory carbon recovery (Timm et al., 2012). Whether glycine decarboxylase is involved in photorespiratory in germinating fern spores needs to be further investigated. IPMDH catalyzes the oxidative decarboxylation of 3-isopropylmalate in leucine biosynthesis, and was proved to be essential for Arabidopsis pollen development (He et al., 2011). In our results, IPMDH was obviously induced in the stage of GS, suggesting that IPMDH is also pivotal for spore germination. Interestingly, ATP-sulfurylase acts as the metabolic entry point into the sulfur assimilation pathway. It was found that the increase of ATP-sulfurylase level during soybean seed development could lead to an increase in the availability of sulfur amino acids (Phartiyal et al., 2006). The induced ATP-sulfurylase in GS and SPC may facilitate the synthesis of sulfur rich amino acids for spore germination. Additionally, pyridoxal biosynthesis protein (PDX) was specially expressed in germinating spores, but not found in MS. PDX family in plants is primarily known for its role in vitamin B6 biosynthesis. Recently, PDX1.2 was proved to be critically required for hypocotyl elongation and primary root growth (Leuendorf et al., 2014). The increase of PDX in germinating spores implies that it probably functions in rhizoid elongation.

Signaling and Vesical Trafficking are Important for Spore Germination

The critical roles of hormone signaling and vesical trafficking in germinating seeds and pollen grains have been well-studied (Ellis and Turner, 2002; Dai et al., 2007a), but little information is available for fern spore germination. In this study, we found some novel factors in signal transduction and vesical trafficking upon fern spore germination, including chalcone isomerase, WD40, 14-3-3 protein, GTP-binding nuclear protein Ran (RAN), and ADP-ribosylation factor (ARF) (Table 1). Among them, chalcone isomerase involved in flavonoid biosynthesis was increased at the stage of RS. Flavonoid participates in auxin signaling and facilitates pollen-tube growth (Falcone Ferreyra et al., 2012). In maize (*Zea mays*) and *Petunia hybrida* mutants, the flavonoid-deficient pollen failed to produce a functional pollen tube (Mo et al., 1992). The increase of chalcone isomerase in spores implies that flavonoids may function in fern spore germination as well. Interestingly, flavonoid biosynthesis is regulated by conserved WD40 domain-contained transcription factors (Falcone Ferreyra et al., 2012). Members of WD40 protein superfamily are known

as key regulators of multi-cellular processes (e.g., cell division, light signaling, protein trafficking, cytoskeleton dynamics, nuclear export, RNA processing, chromatin modification, and transcriptional mechanism), acting as scaffolding molecules assisting proper activity of other proteins (Stirnemann et al., 2010). Moreover, some WD40 proteins have been found to be crucial for Arabidopsis seed germination (Gachomo et al., 2014) and pollen viability in flax (*Linum usitatissimum* L.) (Kumar et al., 2013). In our results, the nucleus-localized WD40 was induced in horsetail germinating spores, indicating its regulatory function for spore germination and rhizoid tip-growth. Similarly, 14-3-3 protein family is also a highly conserved eukaryotic proteins with multiple molecular and cellular functions by binding to phosphorylated client proteins to modulate their function (Denison et al., 2011). In lily (*Lilium longiflorum*) pollen, 14-3-3 protein was involved in the regulation of plasma membrane H⁺-ATPase via modulation of its activity, which is essential for germination and tube elongation (Pertl et al., 2010). In this study, we found a 14-3-3 protein was induced in the stages of RS and SPC, indicating it would play roles in cell nuclear migration and rhizoid tip-growth. Importantly, we also found four proteoforms of RAN were significantly increased in germinating spores. RAN, a primarily nucleus-localized small GTPase, is essential for nuclear transport and assembly, mRNA processing, and cell cycle (Ciciarello et al., 2007). It was reported that overexpression of *RAN1* in rice (*Oryza sativa*) and Arabidopsis altered the mitotic progress and sensitivity to auxin (Wang et al., 2006). The multi-proteoforms of RAN induced in germinating horsetail spores were probably due to the different phosphorylation levels, implying their probable key roles in cell division and polar growth during spore germination. Besides of RAN, we found another important small GTPase, ARF, was induced at the stages of RS and DCS. ARF contributes to the regulation of multiple trafficking routes with respect to Golgi organization, endocytic cycling, cell polarity and cytokinesis (Yorimitsu et al., 2014). It was found that ARF played essential roles for endosomal recycling during Arabidopsis pollen and root hair polarized tip growth (Richter et al., 2011). The induced ARF in horsetail spores meets the specific requirement of early-secretory and polar vesical recycling to facilitate rhizoids tip growth.

Cytoskeleton and Cell Wall Dynamics are Necessary for Spore Germination

Actin cytoskeleton directs the flow of vesicles to the apical domain, where they fuse with the plasma membrane and contribute their contents to the expanding cell wall (Hepler et al., 2013). The local changes in contents and viscosity of the apical wall control the local expansion rate and cell elongation. Precise mechanisms in the organization of actin cytoskeleton and cell wall dynamics have been well-studied in growing pollen tubes (Hepler et al., 2013; Qu et al., 2015), but the mechanisms remain to be further elucidated in fern spores. Here we found the levels of actin, reversibly glycosylated polypeptide, rhamnose biosynthetic enzyme 1, and caffeoyl-CoA O-methyltransferase (CCoAOMT) were altered during horsetail spore germination. Actin is known to be the main content of actin cytoskeleton, and the other three enzymes are essential

for cell wall modulation. Reversibly glycosylated polypeptide is implicated in polysaccharide biosynthesis (Langeveld et al., 2002), and may function in cell wall construction in pollen from rice and *Picea meyeri* (Dai et al., 2006; Chen et al., 2009). Besides, rhamnose biosynthetic enzyme catalyzes the synthesis of L-rhamnose, an important constituent of pectic polysaccharides in cell wall of pollen tube (Yue et al., 2014). In addition, CCoAOMT has a proven role in lignin monomer biosynthesis, which is crucial for pollen wall development (Arnaud et al., 2012). In our results, rhamnose biosynthetic enzyme was increased in GS and decreased in SPC, CCoAOMT was induced in RS and reduced in the stages of DCS and SPC, and two proteoforms reversibly glycosylated polypeptide displayed different changes. All these alterations would modulate the biosynthesis of cell wall components (e.g., polysaccharide, rhamnose, and lignin), leading to the dynamics of rhizoid cell wall rigidity for sustained polarized growth.

Rapid Protein Synthesis, Processing, and Turnover are Essential for Fern Spore Germination

The germinating pollen and fern spores exhibit quick switches from metabolic quiescent state to active state. The substance and energy supply for rapid cell division and polar tip-growth need to be triggered in a short time period. Although it has been found that mature pollen grains have pre-synthesized mRNA and proteins for germination and tube growth (Taylor and Hepler, 1997; Dai et al., 2006), *de novo* protein synthesis is necessary for pollen tube elongation (Dai et al., 2007a,b). For fern species, the germination of green spores from *O. sensibilis* and non-green spores from *A. phyllitidis*, *Marsilea vestita*, *Pteridium aquilinum*, and *Pteris vittata* were not inhibited by actinomycin D (Raghavan, 1977, 1991, 1992; Kuligowski et al., 1991). In horsetail spores, similar metabolic features were discovered from our proteomics data. We found eleven DAPs involved in transcription and protein synthesis, including a threonyl-tRNA synthetase, a 40S ribosomal protein SA, three eukaryotic translation initiation factors (eIFs) (a eIF3 and two eIF4A), five elongation factors (EFs) (two proteoforms of EF2, a EF-G, a EF-Tu, and a EF-Ts), and a protein translation-related GTP-binding protein tyrosine phosphorylated protein A (TypA) (Table 1). The increases of threonyl-tRNA synthetase in GS and SPC and increases of 40S ribosomal protein SA in RS and DCS indicate that the protein synthesis machinery is enhanced during spore germination, while the changes in multi-proteoforms of eIF and EF imply that certain specific protein synthesis is regulated in diverse modes in different stages of germinating spores. It is interesting to note that these proteins are localized in cytoplasm, mitochondria, and chloroplast, respectively (Table 1). This indicates that not only the synthesis of nuclear gene-encoding proteins are necessary, but also the protein synthesis machineries in chloroplast and mitochondria are all triggered for the active metabolism in germinating spores.

Molecular chaperones not only control house-keeping processes, but also regulate protein functional folding and assembly which are necessary for activating/inhibiting various signaling pathways. In horsetail germinating spores, we found

eight chaperones, including five proteoforms of heat shock protein 70 (HSP70), HSP90, T-complex protein 1 (TCP1) subunit alpha, and TCP1 subunit gamma. Among these HSP members, HSP70 was found to bind microtubules and interact with kinesin in tobacco (*Nicotiana tabacum*) pollen tubes (Parrotta et al., 2013). Besides, Arabidopsis HSP90 was found to be active in mature and germinating pollen grains (Prasinos et al., 2005). Moreover, TCP1 plays a pivotal role in the folding and assembly of cytoskeleton proteins as an individual or complex with other subunits (Sternlicht et al., 1993; Bhaskar et al., 2012). The variations of chaperones in spores highlight that protein processing is important for diverse processes (e.g., photosynthesis, cytoskeleton, protein turnover, and various metabolisms) upon horsetail spore germination.

In addition to folding and assembly, active protein turnover occurred in germinating horsetail spores, which was reflected by the changes of eight degradation-related proteins (Table 1). These proteins included an alpha 7 proteasome subunit, three proteoforms of zinc dependent protease, a zinc metalloprotease, two proteoforms of FtsH protease, and an ATP-dependent Clp protease (Table 1). Among them, all the zinc dependent proteases, zinc metalloprotease, and FtsH are all ATP-dependent metalloproteases in chloroplasts, which play a major role in assembly and maintenance of the plastidic membrane system. The Clp protease system plays essential role in plastid development through selective removal of miss-folded, aggregated (Nishimura and van Wijk, 2014), or unwanted proteins. All these imply that active protein degradation and turnover in chlorophyllous spores are crucial for spore germination.

ROS Homeostasis is Crucial for Spore Germination

In the tip-growing pollen tubes and root hairs, ROS act as regulators in diverse signal and metabolic pathways, modulating kinase cascades, ion channels, cell wall properties, nitric oxide levels, and G-protein activities (Wilson et al., 2008; Swanson and Gilroy, 2010). Interestingly, in fern spores, nitric oxide was shown to be a positive regulator for tip growth (Bushart and Roux, 2007). However, the ROS function in fern spores is poorly understood. Our proteomics results revealed that seven proteins in germinating spores function as ROS scavengers, including two proteoforms of 2-cys peroxiredoxin (Prx), a chloroplast drought-induced stress protein of 32 kDa (CDSP32), ascorbate peroxidase (APX), and three proteoforms of dehydroascorbate reductase (DHAR). The proteoforms of Prx, CDSP32, DHAR, and APX are predicted to be localized in chloroplasts and/or cytoplasm in horsetail spores (Table 1). CDSP32 is composed of two thioredoxin (Trx) modules and has been found to be involved in the protection of the photosynthetic apparatus against oxidative damage (Broin et al., 2002). PrxR/Trx pathway is a central antioxidant defense system in plants, in which PrxRs employ a thiol-based catalytic mechanism to reduce H₂O₂ and is regenerated using Trxs as electron donors. APX and DHAR are the key members in glutathione-ascorbate cycle. In this cycle, H₂O₂ is reduced to water by APX using ascorbate as the electron donor. The oxidized ascorbate is still a radical,

which can be converted into dehydroascorbate spontaneously or reduced to ascorbate by monodehydroascorbate reductase. Dehydroascorbate is then reduced to ascorbate by DHAR at the expense of glutathione, yielding oxidized glutathione. Finally, oxidized glutathione is reduced by glutathione reductase using NADPH as electron donor (Horling et al., 2003). Previous proteomics studies have revealed that the abundances of APX and Trx were changed in germinating pollen grains from *O. sativa* (Dai et al., 2007a), *Arabidopsis* (Ge et al., 2011), *Brassica napus* (Sheoran et al., 2009), *P. meyeri* (Chen et al., 2009), *Picea wilsonii* (Chen et al., 2012), and *Pinus strobus* (Fernando, 2005). This indicated that PrxR/Trx pathway and glutathione-ascorbate cycle were employed in germinating pollen grains. Thus, our results indicate chlorophyllous fern spores with active photosynthesis and respiration modulate ROS homeostasis by these enzymes during germination.

Prediction of Protein-protein Interaction Upon Spore Germination

The fern spore germination is a fine-tuned process regulated by temporal and spatial expression and interaction of a number of genes/proteins. To discover the relationship of DAPs during horsetail spore germination, the protein-protein interaction networks were generated by the web-tool STRING 9.1 (<http://string-db.org>). The DAPs homologs in *Arabidopsis* were found by sequence BLASTing in TAIR database (<http://www.arabidopsis.org/Blast/index.jsp>) (Supplementary Table S5),

and then the homologs were subjected to the molecular interaction tool of STRING 9.1 for creation of proteome-scale interaction network. Among all the 80 DAPs, 63 proteins identified in germinating spores and represented by 38 unique homologous proteins from *Arabidopsis*, were depicted in the STRING database (Figure 5), according to the information from the published literature, genome analysis based on domain fusion, phylogenetic profiling/homology, gene neighborhood, co-occurrence, co-expression, and other experimental evidence (Supplementary Figure S4). Four functional modules are apparently illuminated in the network, which form tightly connected clusters (Figure 5). In the protein networks, stronger associations are represented by thicker lines (Figure 5). In Module 1 (green nodes), photosynthetic proteins (chlorophyll a/b-binding protein, RuBisCO large subunit, RuBisCO activase, and RuBisCO large subunit binding protein), three members of chloroplastic protein synthesis machine (EF-Ts, EF-G, and TypA), chloroplast-located Prx/Trx and zinc dependent protease, as well as a mitochondria-located glycine decarboxylase appeared linked closely. This implies that photosynthesis, photorespiration, chloroplastic protein synthesis and turnover, as well as ROS scavenging are active and cooperated closely in horsetail chlorophyllous spores. Besides, multiple metabolic enzymes (e.g., two malate dehydrogenases, enolase, transketolase, pyruvate dehydrogenase, and 6-phosphogluconate dehydrogenase), cytoplasmic members of protein synthesis machine (40S ribosomal protein SA,

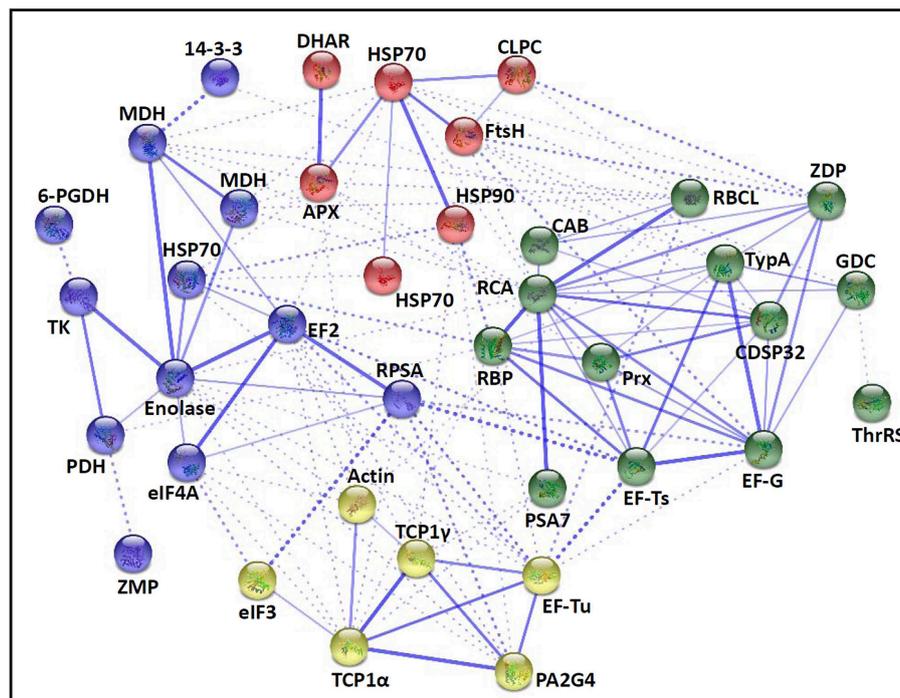


FIGURE 5 | The protein-protein interaction (PPI) network of proteins in *E. arvense* spores based on STRING analysis. A total of 63 differentially abundant proteins represented by 38 unique homologous proteins from *Arabidopsis* are shown in PPI network. Nodes in different

colors belong to four main groups. The PPI network is shown in the confidence view generated by STRING database. Strong associations are represented by thicker lines. Detailed information on protein names and abbreviations can be found in Table 1.

eIF4A, EF2), a molecular chaperone (HSP70), a protease (zinc metalloprotease), and 14-3-3 protein were assigned in Module 2 (blue nodes). These linked proteins show that diverse metabolic pathways (e.g., TCA cycle, glycolysis, pentose phosphate pathway, and Calvin cycle) formed a synergistic system for carbon and energy supplies during germination. Moreover, these metabolic activities were controlled by the levels and activities of key enzymes that were modulated through protein synthesis, folding, and turnover, while 14-3-3 proteins acted as a crucial regulator of these metabolic processes (e.g., TCA cycle) (Diaz et al., 2011). Interestingly, actin is linked with TCP1, proliferation-associated protein 2G4, and two members of protein synthesis machine (eIF3 and EF-Tu) in Module 3 (yellow nodes), indicating the synthesis and processing of cytoskeletal proteins are pivotal for the rapid cell division and cell cycle upon spore germination. Furthermore, proteins involved in protein folding (HSP70 and HSP90) and degradation (FtsH and ATP-dependent Clp protease), as well as ROS scavengers (APX and DHAR) were fitted into Model 4 (red nodes). This indicates that the protein conformational changes determine their fates and are regulated by ROS homeostasis in germinating spores.

Conclusion and Remarks

Although the molecular mechanism of pollen germination has been well-studied as a model for cell polar growth, our knowledge of fern chlorophyllous spore germination is lacking. In this proteomics study, we found some pollen homologous proteins and several novel components are pivotal for fern spore germination. The dynamics of photosynthesis, TCA cycle, glycolysis, and pentose phosphate pathway, as well as the variations of reserve mobilization pathways (fatty acid synthesis, amino acid metabolism, sulfur assimilation, and secondary metabolism) indicate that both heterotrophic and autotrophic

metabolisms are triggered in chlorophyllous spores, which is obviously distinct with non-green spores and pollen grains. Besides, a number of proteins are suspected to be necessary for the cell nuclear migration, cytoskeleton dynamics, and cell wall modulation during fern spore germination. Importantly, the protein synthesis machines, protein processing, and proteasome-dependent protein degradation in cytoplasm and chloroplasts are active for the rapid protein synthesis and turnover. In addition, several members in ROS signaling and G protein-involved vesical trafficking are crucial for polar rhizoid growth. All these provide invaluable information, however, further validation and characterization of these proteins in a model system (i.e., *C. richardii*), as well as the post-translational modification analysis are still necessary for ultimately discovering protein functions and interactions toward understanding of the underlying sophisticated cellular and molecular processes in fern germinating spores.

Acknowledgments

The project was supported by the National Natural Science Foundation of China (No. 31071194), the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2011), Capacity Construction Project of Local Universities, Shanghai, China (No.14390502700) to SD, and by Funding Program for Young Teachers at Universities and Colleges of Shanghai (2014), and General Scientific Research Project of Shanghai Normal University (No. SK201419) to QZ.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2015.00441/abstract>

References

- Arnaud, D., Déjardin, A., Leplé, J. C., Lesage-Descauses, M. C., Boizot, N., Villar, M., et al. (2012). Expression analysis of *LIM* gene family in poplar, toward an updated phylogenetic classification. *BMC Res. Notes* 5:102. doi: 10.1186/1756-0500-5-102
- Balbuena, T. S., He, R., Salvato, F., Gang, D. R., and Thelen, J. J. (2012). Large-scale proteome comparative analysis of developing rhizomes of the ancient vascular plant *Equisetum hyemale*. *Front. Plant Sci.* 3:131. doi: 10.3389/fpls.2012.00131
- Ballesteros, D., Estrelles, E., Walters, C., and Ibars, A. M. (2011). Effect of storage temperature on green spore longevity for the ferns *Equisetum ramosissimum* and *Osmunda regalis*. *Cryoletters* 32, 89–98.
- Banks, J. A. (1999). Gametophyte development in ferns. *Annu. Rev. Plant Biol.* 50, 163–186. doi: 10.1146/annurev.arplant.50.1.163
- Bhaskar, Kumari, N., and Goyal, N. (2012). Cloning, characterization and sub-cellular localization of gamma subunit of T-complex protein-1 (chaperonin) from *Leishmania donovani*. *Biochem. Biophys. Res. Commun.* 429, 70–74. doi: 10.1016/j.bbrc.2012.10.090
- Broin, M., Cui, S., Eymery, F., and Rey, P. (2002). The plastidic 2-cysteine peroxiredoxin is a target for a thioredoxin involved in the protection of the photosynthetic apparatus against oxidative damage. *Plant Cell* 14, 1417–1432. doi: 10.1105/tpc.001644
- Bushart, T. J., Cannon, A., Clark, G., and Roux, S. J. (2014). Structure and function of CrACA1, the major PM-type Ca²⁺-ATPase, expressed at the peak of the gravity-directed trans-cell calcium current in spores of the fern *Ceratopteris richardii*. *Plant Biol.* 16, 151–157. doi: 10.1111/plb.12107
- Bushart, T. J., and Roux, S. J. (2007). Conserved features of germination and polarized cell growth: a few insights from a pollen-fern spore comparison. *Ann. Bot.* 99, 9–17. doi: 10.1093/aob/mcl159
- Chatterjee, A., Porterfield, D. M., Smith, P. S., and Roux, S. J. (2000). Gravity-directed calcium current in germinating spores of *Ceratopteris richardii*. *Planta* 210, 607–610. doi: 10.1007/s004250050050
- Chen, T., Wu, X., Chen, Y., Li, X., Huang, M., Zheng, M., et al. (2009). Combined proteomic and cytological analysis of Ca²⁺-calmodulin regulation in *Picea meyeri* pollen tube growth. *Plant Physiol.* 149, 1111–1126. doi: 10.1104/pp.108.127514
- Chen, Y., Liu, P., Hoehenwarter, W., and Lin, J. (2012). Proteomic and phosphoproteomic analysis of *Picea wilsonii* pollen development under nutrient limitation. *J. Proteome Res.* 11, 4180–4190. doi: 10.1021/pr300295m
- Ciciarello, M., Mangiacasale, R., and Lavia, P. (2007). Spatial control of mitosis by the GTPase Ran. *Cell. Mol. Life Sci.* 64, 1891–1914. doi: 10.1007/s00018-007-6568-2
- Dai, S., Chen, T., Chong, K., Xue, Y., Liu, S., and Wang, T. (2007a). Proteomics identification of differentially expressed proteins associated with pollen

- germination and tube growth reveals characteristics of germinated *Oryza sativa* pollen. *Mol. Cell. Proteomics* 6, 207–230. doi: 10.1074/mcp.M600146-MCP200
- Dai, S., Gao, G., Mu, H. F., and Song, Y. Y. (2008). Comparison of germination between fern spores and spermatophyte pollen. *Chin. Bull. Bot.* 25, 139–148.
- Dai, S., Li, L., Chen, T., Chong, K., Xue, Y., and Wang, T. (2006). Proteomic analyses of *Oryza sativa* mature pollen reveal novel proteins associated with pollen germination and tube growth. *Proteomics* 6, 2504–2529. doi: 10.1002/pmic.200401351
- Dai, S., Wang, Q., Bao, W., and Shing, K. H. (2002). Spore morphology of pteridophytes from China III. Thelypteridaceae 1. *Cyclosorus* Link. *Acta Phyto. Sin.* 40, 334–344.
- Dai, S., Wang, T., Yan, X., and Chen, S. (2007b). Proteomics of pollen development and germination. *J. Proteome Res.* 6, 4556–4563. doi: 10.1021/pr070474y
- de Boer, G. J., Testerink, C., Pielage, G., Nijkamp, H. J., and Stuitje, A. R. (1999). Sequences surrounding the transcription initiation site of the *Arabidopsis* enoyl-acyl carrier protein reductase gene control seed expression in transgenic tobacco. *Plant Mol. Biol.* 39, 1197–1207. doi: 10.1023/A:1006129924683
- DeMaggio, A. E., and Stetler, D. A. (1980). Storage products in spores of *Onoclea sensibilis* L. *Amer. J. Bot.* 67, 452–455. doi: 10.2307/2442284
- DeMaggio, A. E., and Stetler, D. A. (1985). Mobilization of storage reserves during fern spore germination. *P. Roy. Soc. B Biol. Sci.* 86, 195–202. doi: 10.1017/S0269727000008137
- Denison, F. C., Paul, A. L., Zupanska, A. K., and Ferl, R. J. (2011). 14-3-3 proteins in plant physiology. *Semin. Cell Dev. Biol.* 22, 720–727. doi: 10.1016/j.semcdb.2011.08.006
- Diaz, C., Kusano, M., Sulpice, R., Araki, M., Redestig, H., Saito, K., et al. (2011). Determining novel functions of *Arabidopsis* 14-3-3 proteins in central metabolic processes. *BMC Syst. Biol.* 5:192. doi: 10.1186/1752-0509-5-192
- Ellis, C., and Turner, J. G. (2002). A conditionally fertile coil allele indicates cross-talk between plant hormone signalling pathways in *Arabidopsis thaliana* seeds and young seedlings. *Planta* 215, 549–556. doi: 10.1007/s00425-002-0787-4
- Falcone Ferreyra, M. L., Rius, S. P., and Casati, P. (2012). Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Front. Plant Sci.* 3:222. doi: 10.3389/fpls.2012.00222
- Fernando, D. D. (2005). Characterization of pollen tube development in *Pinus strobus* (Eastern white pine) through proteomic analysis of differentially expressed proteins. *Proteomics* 5, 4917–4926. doi: 10.1002/pmic.200500009
- Gachomo, E. W., Jimenez-Lopez, J. C., Baptiste, L. J., and Kotchoni, S. O. (2014). GIGANTUS1 (GTS1), a member of Transducin/WD40 protein superfamily, controls seed germination, growth and biomass accumulation through ribosome-biogenesis protein interactions in *Arabidopsis thaliana*. *BMC Plant Biol.* 14:37. doi: 10.1186/1471-2229-14-37
- Ge, W., Song, Y., Zhang, C., Zhang, Y., Burlingame, A. L., and Guo, Y. (2011). Proteomic analyses of apoplastic proteins from germinating *Arabidopsis thaliana* pollen. *Biochim. Biophys. Acta Prot. Proteom.* 1814, 1964–1973. doi: 10.1016/j.bbapap.2011.07.013
- Gemmrich, A. (1979). Isocitrate lyase in germinating spore of the fern *Anemia phyllitidis*. *Phytochemistry* 18, 1143–1146. doi: 10.1016/0031-9422(79)80122-3
- Guillon, J. M. (2007). Molecular phylogeny of horsetails (*Equisetum*) including chloroplast atpB sequences. *J. Plant Res.* 120, 569–574. doi: 10.1007/s10265-007-0088-x
- Guitton, C., and Mache, R. (1987). Phosphorylation *in vitro* of the large subunit of the ribulose-1,5-bisphosphate carboxylase and of the glyceraldehyde-3-phosphate dehydrogenase. *Eur. J. Biochem.* 166, 249–254. doi: 10.1111/j.1432-1033.1987.tb13509.x
- He, Y., Chen, L., Zhou, Y., Mawhinney, T. P., Chen, B., Kang, B. H., et al. (2011). Functional characterization of *Arabidopsis thaliana* isopropylmalate dehydrogenases reveals their important roles in gametophyte development. *New Phytol.* 189, 160–175. doi: 10.1111/j.1469-8137.2010.03460.x
- Hepler, P. K., Rounds, C. M., and Winship, L. J. (2013). Control of cell wall extensibility during pollen tube growth. *Mol. Plant* 6, 998–1017. doi: 10.1093/mp/sst103
- Horling, F., Lamkemeyer, P., Konig, J., Finkemeier, I., Kandlbinder, A., Baier, M., et al. (2003). Divergent light-, ascorbate-, and oxidative stress-dependent regulation of expression of the peroxiredoxin gene family in *Arabidopsis*. *Plant Physiol.* 131, 317–325. doi: 10.1104/pp.010017
- Kamachi, H., Matsunaga, E., Noguchi, M., and Inoue, H. (2004). Novel mutant phenotypes of a dark-germinating mutant *dkg1* in the fern *Ceratopteris richardii*. *J. Plant Res.* 117, 163–170. doi: 10.1007/s10265-004-0143-9
- Kuligowski, J., Ferrand, M., and Chenou, E. (1991). Stored mRNA in early embryos of a fern *Marsilea vestita*: a paternal and maternal origin. *Mol. Reprod. Dev.* 30, 27–33. doi: 10.1002/mrd.1080300104
- Kumar, S., Jordan, M. C., Datla, R., and Cloutier, S. (2013). The LuWD40-1 gene encoding WD repeat protein regulates growth and pollen viability in flax (*Linum usitatissimum* L.). *PLoS ONE* 8:e69124. doi: 10.1371/journal.pone.0069124
- Langeveld, S. M., Vennik, M., Kottenhagen, M., Van Wijk, R., Buijk, A., Kijne, J. W., et al. (2002). Glucosylation activity and complex formation of two classes of reversibly glycosylated polypeptides. *Plant Physiol.* 129, 278–289. doi: 10.1104/pp.010720
- Large, M. F., Blanchon, D. J., and Angus, M. L. (2006). Devitalisation of imported horsetail (*Equisetum hyemale*). *New Zeal. J. Crop Hort. Sci.* 34, 151–153. doi: 10.1080/01140671.2006.9514400
- Lebkuecher, J. (1997). Desiccation-time limits of photosynthetic recovery in *Equisetum hyemale* (Equisetaceae) spores. *Am. J. Bot.* 84, 792. doi: 10.2307/2445815
- Leuendorf, J. E., Mooney, S. L., Chen, L., and Hellmann, H. A. (2014). *Arabidopsis thaliana* PDX1.2 is critical for embryo development and heat shock tolerance. *Planta* 240, 137–146. doi: 10.1007/s00425-014-2069-3
- Lloyd, R. M., and Klekowski, E. J. Jr. (1970). Spore germination and viability in pteridophyta: evolutionary significance of chlorophyllous spores. *Biotropica* 2, 129–137. doi: 10.2307/2989770
- Malhó, R., Liu, Q., Monteiro, D., Rato, C., Camacho, L., and Dinis, A. (2006). Signalling pathways in pollen germination and tube growth. *Protoplasma* 228, 21–30. doi: 10.1007/s00709-006-0162-6
- Marmottant, P., Ponomarenko, A., and Bienaimé, D. (2013). The walk and jump of *Equisetum* spores. *Proc. R. Soc. B Biol. Sci.* 280:1465. doi: 10.1098/rspb.2013.1465
- Mo, Y., Nagel, C., and Taylor, L. P. (1992). Biochemical complementation of chalcone synthase mutants defines a role for flavonols in functional pollen. *Proc. Natl. Acad. Sci. U.S.A.* 89, 7213–7217. doi: 10.1073/pnas.89.15.7213
- Nishimura, K., and van Wijk, K. J. (2014). Organization, function and substrates of the essential Clp protease system in plastids. *Biochim. Biophys. Acta*. doi: 10.1016/j.bbabi.2014.11.012. [Epub ahead of print]. Available online at: <http://www.sciencedirect.com/science/article/pii/S0005272814006574#>
- Parrotta, L., Cresti, M., and Cai, G. (2013). Heat-shock protein 70 binds microtubules and interacts with kinesin in tobacco pollen tubes. *Cytoskeleton (Hoboken)* 70, 522–537. doi: 10.1002/cm.21134
- Pertl, H., Pöckl, M., Blaschke, C., and Obermeyer, G. (2010). Osmoregulation in *Lilium* pollen grains occurs via modulation of the plasma membrane H⁺ ATPase activity by 14-3-3 proteins. *Plant Physiol.* 154, 1921–1928. doi: 10.1104/pp.110.165696
- Phartiyal, P., Kim, W. S., Cahoon, R. E., Jez, J. M., and Krishnan, H. B. (2006). Soybean ATP sulfurylase, a homodimeric enzyme involved in sulfur assimilation, is abundantly expressed in roots and induced by cold treatment. *Arch. Biochem. Biophys.* 450, 20–29. doi: 10.1016/j.abb.2006.03.033
- Poghosyan, Z. P., Giannoulia, K., Katinakis, P., Murphy, D. J., and Hatzopoulos, P. (2005). Temporal and transient expression of olive enoyl-ACP reductase gene during flower and fruit development. *Plant Physiol. Biochem.* 43, 37–44. doi: 10.1016/j.plaphy.2004.12.002
- Prasinos, C., Krampis, K., Samakovli, D., and Hatzopoulos, P. (2005). Tight regulation of expression of two *Arabidopsis* cytosolic Hsp90 genes during embryo development. *J. Exp. Bot.* 56, 633–644. doi: 10.1093/jxb/eri035
- Qu, X., Jiang, Y., Chang, M., Liu, X., Zhang, R., and Huang, S. (2015). Organization and regulation of the actin cytoskeleton in the pollen tube. *Front. Plant Sci.* 5:786. doi: 10.3389/fpls.2014.00786
- Raghavan, V. (1977). Cell morphogenesis and macromolecule synthesis during phytochrome-controlled germination of spores of the fern, *Pteris vittata*. *J. Exp. Bot.* 28, 439–456. doi: 10.1093/jxb/28.2.439
- Raghavan, V. (1991). Gene activity during germination of spores of the fern, *Onoclea sensibilis*: RNA and protein synthesis and the role of stored mRNA. *J. Exp. Bot.* 42, 251–260. doi: 10.1093/jxb/42.2.251
- Raghavan, V. (1992). Germination of fern spores. *Am. Sci.* 80, 176–185.

- Richter, S., Müller, L. M., Stierhof, Y. D., Mayer, U., Takada, N., Kost, B., et al. (2011). Polarized cell growth in Arabidopsis requires endosomal recycling mediated by GBF1-related ARF exchange factors. *Nat. Cell Biol.* 14, 80–86. doi: 10.1038/ncb2389
- Salmi, M. L., Morris, K. E., Roux, S. J., and Porterfield, D. M. (2007). Nitric oxide and cGMP signaling in calcium-dependent development of cell polarity in *Ceratopteris richardii*. *Plant Physiol.* 144, 94–104. doi: 10.1104/pp.107.096131
- Sheoran, I. S., Pedersen, E. J., Ross, A. R., and Sawhney, V. K. (2009). Dynamics of protein expression during pollen germination in canola (*Brassica napus*). *Planta* 230, 779–793. doi: 10.1007/s00425-009-0983-6
- Sternlicht, H., Farr, G. W., Sternlicht, M. L., Driscoll, J. K., Willison, K., and Yaffe, M. B. (1993). The t-complex polypeptide 1 complex is a chaperonin for tubulin and actin *in vivo*. *Proc. Natl. Acad. Sci. U.S.A.* 90, 9422–9426. doi: 10.1073/pnas.90.20.9422
- Stirnemann, C. U., Petsalaki, E., Russell, R. B., and Muller, C. W. (2010). WD40 proteins propel cellular networks. *Trends Biochem. Sci.* 35, 565–574. doi: 10.1016/j.tibs.2010.04.003
- Suetsugu, N., and Wada, M. (2003). Cryptogam blue-light photoreceptors. *Curr. Opin. Plant Biol.* 6, 91–96. doi: 10.1016/S1369526602000067
- Suo, J., Chen, S., Zhao, Q., Shi, L., and Dai, S. (2015). Fern spore germination in response to environmental factors. *Front. Biol.* doi: 10.1007/s11515-015-1342-6
- Swanson, S., and Gilroy, S. (2010). ROS in plant development. *Physiol. Plantarum* 138, 384–392. doi: 10.1111/j.1399-3054.2009.01313.x
- Tan, L., Chen, S., Wang, T., and Dai, S. (2013). Proteomic insights into seed germination in response to environmental factors. *Proteomics* 13, 1850–1870. doi: 10.1002/pmic.201200394
- Taylor, L. P., and Hepler, P. K. (1997). Pollen germination and tube growth. *Ann. Rev. Plant Biol.* 48, 461–491. doi: 10.1146/annurev.arplant.48.1.461
- Timm, S., Florian, A., Arrivault, S., Stitt, M., Fernie, A. R., and Bauwe, H. (2012). Glycine decarboxylase controls photosynthesis and plant growth. *FEBS Lett.* 586, 3692–3697. doi: 10.1016/j.febslet.2012.08.027
- Tsuboi, H., Nakamura, S., Schäfer, E., and Wada, M. (2012). Red light-induced phytochrome relocation into the nucleus in *Adiantum capillus-veneris*. *Mol. Plant* 5, 85–92. doi: 10.1093/mp/ssr119
- Vizcaíno, J. A., Deutsch, E. W., Wang, R., Csordas, A., Reisinger, F., Ríos, D., et al. (2014). ProteomeXchange provides globally co-ordinated proteomics data submission and dissemination. *Nat. Biotechnol.* 30, 223–226. doi: 10.1038/nbt.2839
- Wang, X., Chen, S., Zhang, H., Shi, L., Cao, F., Guo, L., et al. (2010). Desiccation tolerance mechanism in resurrection fern-ally *Selaginella tamariscina* revealed by physiological and proteomic analysis. *J. Proteome Res.* 9, 6561–6577. doi: 10.1021/pr100767k
- Wang, X., Xu, Y., Han, Y., Bao, S., Du, J., Yuan, M., et al. (2006). Overexpression of *RAN1* in rice and Arabidopsis alters primordial meristem, mitotic progress, and sensitivity to auxin. *Plant Physiol.* 140, 91–101. doi: 10.1104/pp.105.071670
- Wilson, I. D., Neill, S. J., and Hancock, J. T. (2008). Nitric oxide synthesis and signaling in plants. *Plant Cell Environ.* 31, 622–631. doi: 10.1111/j.1365-3040.2007.01761.x
- Yorimitsu, T., Sato, K., and Takeuchi, M. (2014). Molecular mechanisms of Sar/Arf GTPases in vesicular trafficking in yeast and plants. *Front. Plant Sci.* 5:411. doi: 10.3389/fpls.2014.00411
- Yu, J., Chen, S., Zhao, Q., Wang, T., Yang, C., Diaz, C., et al. (2011). Physiological and proteomic analysis of salinity tolerance in *Puccinellia tenuiflora*. *J. Proteome Res.* 10, 3852–3870. doi: 10.1021/pr101102p
- Yue, X., Gao, X. Q., Wang, F., Dong, Y., Li, X., and Zhang, X. S. (2014). Transcriptional evidence for inferred pattern of pollen tube-stigma metabolic coupling during pollination. *PLoS ONE* 9:e107046. doi: 10.1371/journal.pone.0107046

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Zhao, Gao, Suo, Chen, Wang and Dai. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Proteasome targeting of proteins in Arabidopsis leaf mesophyll, epidermal and vascular tissues

Julia Svozil, Wilhelm Gruissem and Katja Baerenfaller*

Plant Biotechnology, Department of Biology, Swiss Federal Institute of Technology Zurich, Zurich, Switzerland

OPEN ACCESS

Edited by:

Marc Libault,
University of Oklahoma, USA

Reviewed by:

Kazuhiro Takemoto,
Kyushu Institute of Technology, Japan
Asdrubal Burgos,
Max Planck Institute for Molecular
Plant Physiology, Germany

*Correspondence:

Katja Baerenfaller,
Plant Biotechnology, Department of
Biology, Swiss Federal Institute of
Technology Zurich, Zurich
Universitaetstrasse 2, 8092 Zurich,
Switzerland
kbaerenfaller@ethz.ch

Specialty section:

This article was submitted to
Plant Systems and Synthetic Biology,
a section of the journal
Frontiers in Plant Science

Received: 24 March 2015

Accepted: 11 May 2015

Published: 28 May 2015

Citation:

Svozil J, Gruissem W and Baerenfaller
K (2015) Proteasome targeting of
proteins in Arabidopsis leaf mesophyll,
epidermal and vascular tissues.
Front. Plant Sci. 6:376.
doi: 10.3389/fpls.2015.00376

Protein and transcript levels are partly decoupled as a function of translation efficiency and protein degradation. Selective protein degradation via the Ubiquitin-26S proteasome system (UPS) ensures protein homeostasis and facilitates adjustment of protein abundance during changing environmental conditions. Since individual leaf tissues have specialized functions, their protein composition is different and hence also protein level regulation is expected to differ. To understand UPS function in a tissue-specific context we developed a method termed Meselect to effectively and rapidly separate *Arabidopsis thaliana* leaf epidermal, vascular and mesophyll tissues. Epidermal and vascular tissue cells are separated mechanically, while mesophyll cells are obtained after rapid protoplasting. The high yield of proteins was sufficient for tissue-specific proteome analyses after inhibition of the proteasome with the specific inhibitor Syringolin A (SylA) and affinity enrichment of ubiquitylated proteins. SylA treatment of leaves resulted in the accumulation of 225 proteins and identification of 519 ubiquitylated proteins. Proteins that were exclusively identified in the three different tissue types are consistent with specific cellular functions. Mesophyll cell proteins were enriched for plastid membrane translocation complexes as targets of the UPS. Epidermis enzymes of the TCA cycle and cell wall biosynthesis specifically accumulated after proteasome inhibition, and in the vascular tissue several enzymes involved in glucosinolate biosynthesis were found to be ubiquitylated. Our results demonstrate that protein level changes and UPS protein targets are characteristic of the individual leaf tissues and that the proteasome is relevant for tissue-specific functions.

Keywords: protein level regulation, ubiquitylation, proteasome inhibition, leaf tissues, epidermis, mesophyll, vasculature, *Arabidopsis thaliana*

Introduction

Plant organs are composed of different tissues that are specialized for particular biological processes and the functionality of the organ is the sum of the functions of each of its tissue types. Just as each Arabidopsis organ has its own functional proteome map (Baerenfaller et al., 2008) we also expect that individual tissues have specific protein compositions for their specific functions. In fully grown leaves, the main leaf tissue types are epidermis, mesophyll and vasculature. In Arabidopsis, the epidermis is composed of stomata and trichomes that are embedded in the single cell layer of pavement cells, which are covered with the waxy cuticle. The adaxial and abaxial epidermal tissues enclose the palisade and spongy mesophyll,

which represent the main photosynthetic capacity of the leaf. Embedded in the mesophyll tissue is the vascular tissue, which consists of phloem, xylem and cambial cells (Tsukaya, 2002).

Considering the specific functions of the different leaf tissues their protein composition would be expected to vary, however, leaf tissue-specific proteome information is currently not available. Since the correlation of absolute protein and transcript levels and their dynamic changes over time is limited (Baerenfaller et al., 2008, 2012; Walley et al., 2013), regulatory processes such as protein degradation likely influence the composition of tissue-specific proteomes. The activity of the ubiquitin-26S proteasome system (UPS) should therefore have specific signatures for individual specialized tissues. We previously identified changes in leaf protein composition after proteasome inhibition and a large number of direct UPS targets in leaves and roots (Svozil et al., 2014). However, the whole organ information does not have the necessary resolution to understand UPS function in a tissue-specific context because by homogenizing organs functional protein signatures cannot longer be attributed to individual tissues or cell types (Brandt, 2005).

To avoid this limitation and conserve specific information, different techniques for the separation of tissues or cell types have been developed. In plants, pollen and tissue culture cells can be collected effectively because they are not attached to other tissues. For example, it has been possible to obtain enough *Arabidopsis* pollen for a large-scale proteomics study (Grobei et al., 2009). Trichomes and root hairs are other specialized plant cells that can be easily collected for small-scale proteomics experiments (Wienkoop et al., 2004; Marks et al., 2008; Brechenmacher et al., 2009; Nestler et al., 2011; Van Cutsem et al., 2011). Microcapillaries have also been used to collect cell sap from epidermal cells (Wienkoop et al., 2004) and vascular S-cells (Koroleva and Cramer, 2011) for proteomic analyses. However, this method is restricted to accessible cell types and soluble proteins only. Alternatively, single cell or tissue types can be precisely excised from tissue sections using laser capture microdissection, for example epidermal and vascular cells of maize coleoptiles (Nakazono et al., 2003) and *Arabidopsis* vascular bundle cells (Schad et al., 2005). For any other approach of selecting specific cell types, the cells first need to be released from their tissue context, which in plants requires the degradation of cell walls to obtain protoplasts. Leaf mesophyll and guard cell protoplasts can be easily collected and analyzed (Zhao et al., 2008; Zhu et al., 2009). Enrichment of guard cell protoplasts expressing green fluorescent protein (GFP) has been achieved using fluorescent activating cell sorting (FACS) (Gardner et al., 2009). However, FACS of leaf protoplasts is challenging because of the chlorophyll autofluorescence that can interfere with the sorting process (Galbraith, 2007). Nevertheless, FACS in combination with enhancer trap lines expressing GFP specifically in the leaf epidermal, vasculature or guard cells has allowed enrichment of the respective cell type (Grønlund et al., 2012). FACS has also been employed for transcriptome and proteome analysis of specific cell types in *Arabidopsis* roots both under standard and stress conditions (Birnbaum et al., 2003, 2005; Brady et al., 2007; Dinneny et al., 2008; Petricka

et al., 2012). Recently, a method of protoplasting followed by sonication and manual separation of cotyledon epidermal and vascular cells was reported for tissue-specific transcriptome analyses (Endo et al., 2014). In general, however, single cell and tissue type-specific high-throughput proteomics experiments are challenging because of the availability of material and low protein content of plants cells (Wienkoop et al., 2004; Koroleva and Cramer, 2011).

To enable tissue-specific proteome analyses we developed a rapid and effective method for the separation of the different leaf tissue types (Meselect, mechanical separation of leaf compound tissues). Here we report the specificity of the method in separating leaf mesophyll, vasculature and epidermis tissues. Using Meselect we generated tissue type-specific functional protein maps and identified responses to inhibition of the proteasome that account for the diverse composition of the leaf tissue types. Because Meselect produces a high enough yield of tissue-specific protein extracts for affinity enrichment experiments, we found several tissue-specific UPS target proteins that provide new insights into the role of protein degradation in tissue functions.

Materials and Methods

Plant Growth and Inhibition of the Proteasome

Seeds of *Arabidopsis thaliana* ecotype Col-0 were stratified for 2 days in darkness at 4°C and afterwards grown in short day conditions with 8 h light and 16 h darkness for 55 days at 22°C, 70% humidity. The proteasome was inhibited by spraying the adaxial epidermis of leaves with 10 μM Syringolin A in 0.02% (v/v) Tween 20. As a mock control leaves were sprayed with 0.02% (v/v) Tween 20 only. The whole rosette was treated 30–60 min before the end of the night period and leaves were harvested 8 h after treatment. SylA was produced as described in Svozil et al. (2014) and was kindly provided by R. Dudler (University of Zurich, Switzerland). The experiment was performed in three biological replicates.

Separation of the different Leaf Tissues with the Meselect Method

After harvesting of the leaf, the TAPE sandwich method (Wu et al., 2009) was applied to remove the abaxial epidermis from the leaf. For this, the leaf was positioned between two tape stripes and the tape facing the abaxial epidermis was removed. The tape containing the abaxial epidermis was incubated for 15 min in protoplasting solution (1% cellulase Onozuka RS (Yakult, Japan), 0.25% macerozyme Onozuka R10 (Yakult, Japan), 0.4 M mannitol, 10 mM CaCl₂, 20 mM KCl, 0.1% BSA and 20 mM MES, pH 5.7) under constant agitation at 50 rpm. During this process spongy mesophyll cells adhering to the epidermis are released into the solution, but the epidermis remains attached to the tape. The tape including epidermis was washed twice in washing buffer (154 mM NaCl, 125 mM CaCl₂, 5 mM KCl, 5 mM glucose, and 2 mM MES, pH 5.7) and frozen in liquid nitrogen. After freezing, the epidermis could be easily removed from the tape with pre-cooled tweezers and scrapers. The epidermis was collected in a pre-cooled mortar, ground, and the tissue powder was stored

at -80°C until further processing. The remainder of the leaf without the abaxial epidermis that was attached to the other tape was incubated for about 1 h in protoplast solution, which releases the palisade and spongy mesophyll cells into solution. The mesophyll protoplasts were collected by centrifugation at $200 \times g$ for 2 min at 4°C and washed twice with washing buffer. After complete removal of the buffer, the protoplasts were frozen in liquid nitrogen and stored at -80°C until further processing. The tape with the remaining attached upper epidermal and vascular tissue was immersed in cold washing buffer. The vascular tissue network was removed with a tweezer, washed twice in washing buffer, frozen in liquid nitrogen, and stored at -80°C until further processing. The remainder of the tape with the attached adaxial epidermis was discarded because the specificity of the enrichment could not be confirmed for this tissue.

To verify the enrichment of different leaf tissues types the *Arabidopsis* GAL4-GFP enhancer trap lines with specific GFP expression in the epidermis (KC464), the vasculature (KC274) and the mesophyll (JR11-2) were used (Gardner et al., 2009) (kindly provided by Alex Webb, UK). The leaves of the JR11-2 line in the C24 ecotype background were too fragile for the Meselect method. The line was therefore backcrossed four times in the Col-0 ecotype background to yield sufficiently robust leaves for the generation of the tissue type-specific protein preparations.

Preparation of the Tissue Type-Specific Protein Extracts and Western Blot Analysis

For the preparation of the whole leaf and tissue type specific protein extracts the frozen plant material was ground with a mortar and pestle and proteins were extracted by incubation with SDS buffer [4% SDS, 40 mM Tris-base, 5 mM MgCl_2 , 2x protease inhibitor mix (Roche)] for 20 min at room temperature (RT). Non-soluble material was pelleted by centrifugation at RT for 10 min at $16,200 \times g$. The supernatant was subsequently cleared by ultracentrifugation at RT for 45 min at $100,000 \times g$. The SDS extract of epidermal cells was processed quickly and could not be stored, neither frozen nor at RT, since proteins will precipitate.

For mass spectrometry measurements, 50–150 μg of protein, depending on the tissue type, were subjected to SDS PAGE on 10% SDS gels. After electrophoretic separation of the proteins, the proteins were stained with Coomassie brilliant blue R250 and each lane was cut into 5 sections. Tryptic digest of the proteins and extraction of the peptides was carried out as described (Svozil et al., 2014).

For the Western Blots 15 μg of each extract were electrophoretically separated with SDS PAGE on 10% SDS gels. After blotting, the nitrocellulose membranes were cut horizontally in two halves. The lower parts were probed with α -GFP (1:1000 dilution, Roche) and α -mouse antibodies (1:5000, Roche), and the upper parts with α -HSP90 (1:3000, Agrisera) and α -rabbit (1:3000, Roche) antibodies. After detection of the immunofluorescence signal the membranes were stained with Coomassie R250.

Affinity Enrichment Experiment

Frozen plant material was ground with a mortar and pestle. Mesophyll tissue cells were collected from 8 to 10 leaves each

of three, vascular tissue of nine, and abaxial epidermal tissue of four plants. Proteins from mesophyll cells were extracted as described previously with a sequential extraction using native and urea buffer (Svozil et al., 2014). Proteins from vascular tissue or epidermal cells were extracted only with urea buffer. For the pre-clearing step 500 μg protein of mesophyll and vascular tissue in a total volume of 700–1200 μl and 400 μg protein of epidermal tissue in a total volume of 1800 μl were each washed with 100 μl sepharose CL4B (Sigma-Aldrich). The remaining affinity enrichment method was performed as described previously (Svozil et al., 2014).

Mass Spectrometry Measurements

Mass spectrometry measurements were performed using a LTQ OrbiTrap XL mass spectrometer (Thermo Fisher) coupled to a NanoLC-AS1 (Eksigent) using electrospray ionization. For LC separation a capillary column packed with 8 cm C18 beads with a diameter of 3 μm and a pore size of 100 \AA was used. Peptides were loaded on the column with a flow rate of 500 nl/min for 16 min and eluted by an increasing acetonitrile gradient from 3% acetonitrile to 50% acetonitrile for 60 min with a flow rate of 200 nl/min. One scan cycle was comprised of a survey full MS scan of spectra from m/z 300 to m/z 2000 acquired in the FT-Orbitrap with a resolution of $R = 60,000$ at m/z 400, followed by MS/MS scans of the five highest parent ions. CID was done with a target value of $1e4$ in the linear trap. Collision energy was set to 28 V, Q value to 0.25, and activation time to 30 ms. Dynamic exclusion was enabled at a duration of 120 s.

Interpretation of MS/MS Spectra

The acquired raw spectra were transformed to mgf data format and searched against the TAIR10 database (28, download on January 17th, 2011) (Lamesch et al., 2012) with concatenated decoy database and supplemented with common contaminants (71,032 entries) using the Mascot algorithm (version 2.3.02) (Mascot Science). The search parameters used were: mass = monoisotopic, requirement for tryptic ends, 2 missed cleavages allowed, precursor ion tolerance = ± 10 ppm, fragment ion tolerance = ± 0.8 Da, variable modifications of methionine (M, PSI-MOD name: oxidation, mono $\Delta = 15.995$) and static modifications of cysteine (C, PSI-MOD name: iodoacetamide derivatized residue, mono $\Delta = 57.0215$). The diglycine tag that remains attached to ubiquitylated peptides after tryptic digest was not included as variable modification because we had previously observed that the identification of peptide ubiquitylation sites in complex peptide mixtures in which only a small portion of the peptides carry the diglycine tag will result in unreliable identifications (Svozil et al., 2014). Peptide spectrum assignments with ionscore > 30 and expect value < 0.015 , except those of known contaminants, were filtered for ambiguity. Peptides matching to several proteins were excluded from further analyses. This does not apply to different splice variants of the same protein or to different loci sharing exactly the same amino acid sequence. All remaining spectrum assignments were inserted into the pep2pro database (Baerenfaller et al., 2011; Hirsch-Hoffmann et al., 2012). The false discovery rate (FDR) was calculated by dividing the number of reverse hits by the number

of true hits times 100%. It was 0.55% for the tissue-type specific extracts and 1.6% for the affinity enrichments. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<http://www.proteomexchange.org>) via the PRIDE partner repository (Vizcaíno et al., 2013) with the dataset identifiers PXD000941 and PXD000942. The data are also available in the pep2pro database at www.pep2pro.ethz.ch.

Statistical Analyses and Selection Criteria

Proteins were quantified by normalized spectral counting according to Svozil et al. (2014) by calculating the expected contribution of each individual protein to the samples total peptide pool correcting the values using a normalization factor that balances for the theoretical number of tryptic peptides per protein and sample depth. The statistical analyses and selection criteria were applied to each tissue type-specific dataset separately. For the total protein extracts, proteins with a minimum of 5 spectra within a tissue type were included in the quantification. After filtering, the normalized spectral counts for the proteins in each tissue type were re-normalized by scaling the average to a value of 1. Proteins were considered changing in abundance between SylA and mock treatment if the fold-change of the average relative abundance over three biological replicates was more than 1.5 and the corresponding *p*-value in a paired *t*-test was smaller than 0.05, or if the fold change was an outlier. As outliers we considered fold changes that were higher than the upper adjacent values in boxplot statistics, which correspond to the values of the largest observation that is less than or equal to the upper quartile plus 1.5 times the length of the interquartile range. The upper adjacent values were: 2.94 for the mesophyll, 2.55 for the vasculature and 6.05 for the epidermis. Those proteins, which were identified in at least two of the three biological replicates in one treatment and not at all in the other treatment were also considered as changing in abundance. Proteins considered to be exclusively identified in a specific tissue type were detected with a minimum of 5 spectra in the respective tissue type, but not at all in the other two tissue types. In the affinity enrichment experiments we required a minimum of 5 different spectra to call a protein identified. With these data an index was calculated, where each protein identification in the UBA-domain affinity enrichment counts +1 and each identification in the sepharose control counts -1 toward the final index. As an example, an index of 2 either indicates that the protein was identified in two of the three UBA affinity enrichment replicates and never in the sepharose control, or that it was identified in all three UBA affinity enrichment replicates and only once in the sepharose control. For accepting a protein to be ubiquitylated we required that the protein had a minimum index of 2 and was either not detected at all in the sepharose background control or was at least 5-fold enriched. Analyses were done using the R algorithm (R Core Team, 2012) (Supplemental Data Sheet 1).

Go Categorization

GO categorization was performed as described before (Svozil et al., 2014) using the Ontologizer software (<http://compbio.charite.de/ontologizer>) (Bauer et al., 2008) in combination with

the Arabidopsis annotation file considering aspect Biological Process. Annotations with GO evidence codes IEA (Inferred from Electronic Annotation) or RCA (Inferred from Reviewed Computational Analysis) were excluded from analyses. As background lists the organ-specific protein maps for leaves of the pep2pro TAIR10 was dataset were used (Baerenfaller et al., 2011). Over-representation was assessed with the Topology-weighted method and the *p*-values were corrected for multiple testing using the Bonferroni method. GO categories with a *p* < 0.01 were considered significant.

Results

Effective Separation of the Leaf Epidermal, Vascular, and Mesophyll Tissues

We developed Meselect (mechanical separation of leaf compound tissues) to specifically enrich the three main leaf tissue types. The method first utilizes the TAPE sandwich approach to separate the abaxial epidermis from the remainder of the leaf (Wu et al., 2009). Any attached mesophyll cells are released by rapid protoplasting, which leaves the epidermal cells intact. The tape is flash-frozen and the epidermal cells are collected from the tape with tweezers and scrapers. Mesophyll cells are released from the remainder of the leaf by rapid protoplasting. The vascular tissue embedded in the mesophyll tissue remains intact during protoplasting and is then isolated with a tweezer. To confirm the specificity of the enrichment we applied Meselect to leaves of Arabidopsis lines that express GFP specifically in epidermal, mesophyll and vascular tissues (Gardner et al., 2009) (Figure 1). In the tissue-type specific protein extracts of leaves from the KC464 line, which expresses GFP exclusively in the epidermis, the GFP protein was exclusively found in the epidermal protein extract but not in the protein extract of vasculature and mesophyll tissues (Figure 1A). Similarly, in the tissue type-specific protein extracts of leaves of the KC274 line, which expresses GFP only in the vasculature, GFP was specifically enriched in the vasculature protein extract (Figure 1C). We noticed small amount of GFP in the mesophyll but not the epidermal protein extract. In the JR11-2 line backcrossed with Col-0, GFP is expressed in the spongy mesophyll and GFP was detected only the mesophyll protein extract, confirming that both vasculature and epidermal protein extracts do not contain mesophyll proteins (Figure 1E). Together, the Meselect method effectively and efficiently separates the abaxial epidermal, mesophyll and vascular leaf tissues.

Experimental Design to Detect Tissue-Specific Proteins and differences in UPS Targeting in Different Leaf Tissues

We adopted the workflow in Figure 2 to investigate tissue-specific differences in UPS targeting. For this we extracted total proteins from the three tissues of mock-treated leaves or leaves that were treated with Syringolin A (SylA), which specifically and effectively inhibits the UPS (Groll et al., 2008; Svozil et al., 2014). Together, we identified a total of 1799 distinct proteins and 1114 proteins that were identified with at least 5 spectra in at least one tissue (Supplemental Table 1). Of these proteins only those were

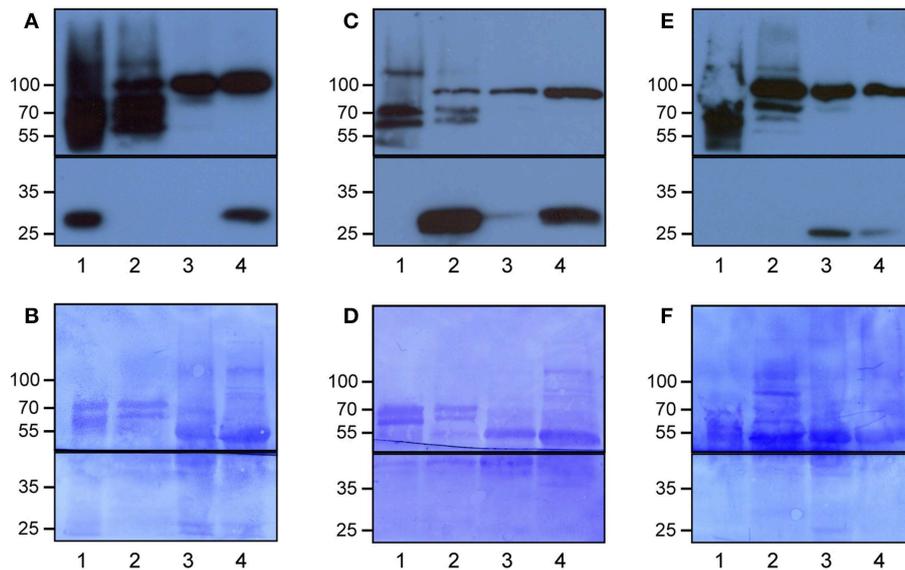


FIGURE 1 | Specific enrichment of the epidermal, vascular and mesophyll tissues using the Meselect method. The different tissue types were separated using leaves of GAL4-GFP enhancer trap lines with GFP expression specifically in the epidermal (A,B), vascular (C,D) or mesophyll (E,F) tissue. The protein blots were probed with antibodies against GFP (lower panels in A,C,E) and

HSP90 (upper panels in A,C,E) as loading control, and afterwards stained with Coomassie (B,D,F) to show protein loading. The black lines indicate where the membranes were cut. Lane 1, epidermal tissue proteins; 2, vascular tissue proteins; 3, mesophyll tissue proteins; and 4, total leaf proteins. At the left, the molecular weight is indicated.

considered to change in abundance if one of the following criteria was met: (i) the fold-change between SylA and mock-treated samples was >1.5 with a $p < 0.05$, (ii) the fold-change was large enough to be considered an outlier according to boxplot statistics, or (iii) the protein was identified in at least two of three biological replicates in one condition but not another condition (Table 1, Supplemental Table 2). Additionally, for tissue specificity we required that the protein was not identified in the other tissues (Table 1, Supplemental Table 3). In total, 225 distinct proteins were found that had increased and 30 that had decreased in the different leaf tissues after inhibition of the proteasome. As discussed previously, reduced protein levels could result from reduced transcription, translation or stabilization of the proteins after SylA treatment, or from proteasome-independent protein degradation (Svozil et al., 2014). In contrast, the proteins that increased after inhibition of the proteasome are likely targets of the UPS or participate in pathways that respond to SylA treatment.

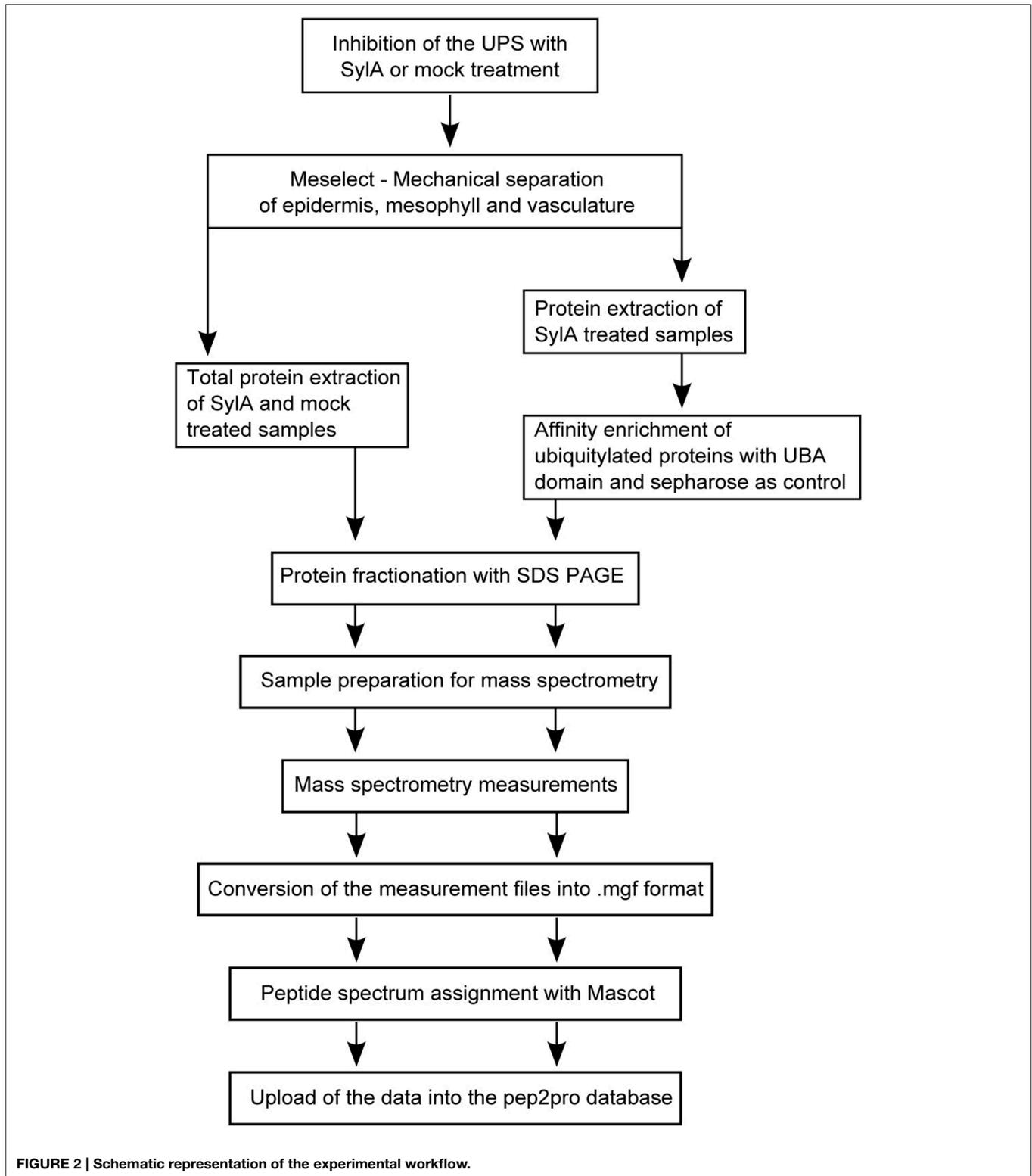
Using the Meselect method the protein yield in the tissue type-specific extracts was adequate for affinity enrichment of ubiquitylated proteins from SylA-treated leaves using the ubiquitin-binding UBA-domain (Sutovsky et al., 2005; Manzano et al., 2008; Svozil et al., 2014). To call a UBA affinity-enriched protein identified we required a minimum of 5 spectra for the protein in the respective tissue extract. We also calculated an index, in which each protein identification in the UBA-domain affinity enrichment counts +1 and in the sepharose control -1 for the final index. We accepted a protein to be ubiquitylated in a tissue extract if it had a minimum index of 2 and was at least

5-fold enriched over the sepharose background control. Using these criteria we identified a total of 519 ubiquitylated proteins in the three different tissues (Table 1, Supplemental Table 4). When comparing the tissue-specific ubiquitylated proteins with the reported ubiquitylated proteins identified in whole leaves (Svozil et al., 2014) we identified 161 proteins that are unique to the leaf tissue-specific dataset (Figure 3). The vascular tissue extract revealed the highest number of newly identified ubiquitylated proteins. Together, the separation of leaf tissues indeed allows the sampling of proteins that remain undetected in the whole leaf context where they are masked by proteins from abundant cell types. In the following, the datasets of proteins that are ubiquitylated and/or accumulate in the mesophyll, vasculature and epidermis after SylA treatment, as well as proteins identified in only one of the three tissue extracts, will be analyzed in detail for their tissue-type specific functions.

Proteins Identified in the Tissue-Specific Protein Extracts have Specific Roles in the Respective Tissues

Proteins Identified Exclusively in Mesophyll have Functions in Photosynthesis

Among the 85 proteins that were only identified in the mesophyll the GO categories that are most significantly over-represented are photosynthesis, starch metabolic process and cellular polysaccharide catabolic process. The proteins in these categories comprise chloroplast proteases, components of the photosystems, as well as proteins that facilitate the incorporation of proteins into the photosystems and a

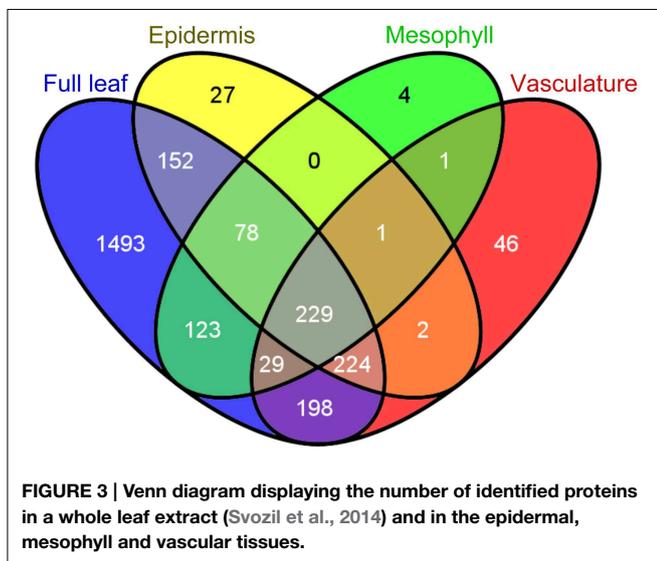


kinase that phosphorylates them (Table 2). The proteins involved in starch metabolic process can be classified into two groups, starch biosynthesis and starch breakdown (Streb

and Zeeman, 2012) (Table 2), and include proteins such as INOSITOL MONOPHOSPHATASE FAMILY PROTEIN (FBPase, AT1G43670) involved in cytosolic sucrose synthesis,

TABLE 1 | Summary of the number of identified proteins in total protein extracts of epidermal, vascular and mesophyll tissues, and in ubiquitin affinity enrichments of proteins from the respective tissue types.

Tissue	Total protein extract			Affinity purification			
	Proteins identified	Proteins identified with ≥ 5 spectra	Exclusive for each tissue type	SylA		Ubiquitylated proteins	Exclusive for each tissue type
				Accumulation	decrease		
Epidermis	1105	713	113	107	9	140	23
Vasculature	1161	730	150	61	12	353	132
Mesophyll	768	465	85	86	9	175	80
Total	1799	1114		225	30	519	



the starch biosynthesis enzyme PHOSPHOGLUCOMUTASE (PGM, AT5G51820) and the starch breakdown enzyme PHOSPHOGLUCAN, WATER DIKINASES (GWD3, AT5G26570). This indicates that starch metabolism is mostly confined to the mesophyll tissue, although starch granules can also be observed in guard cells (Stadler et al., 2003), which are specialized cells in the epidermis.

Epidermis-specific Proteins Function in Defense and Protection

The epidermis and cuticle form the outer barrier of the leaf to the environment. Therefore it can be expected that proteins, which are involved in cuticle formation and immunoprotection are enriched in the epidermal tissue (Barel and Ginzberg, 2008; Kaspar et al., 2010; Yeats et al., 2010). Indeed, the over-represented GO categories among the 113 proteins exclusively identified in the epidermis include *innate immune response*, *immune response* and *immune system process*, as well as *lignin*, *cutin*, and *cell wall pectin biosynthetic process* (Table 3). The proteins responsible for the over-representation of immune response categories also included HSP90.1, which in addition to its chaperone function is also required for the R-gene mediated

defense response (Takahashi et al., 2003). HSP90.1 was the only defense-related protein that accumulated after SylA treatment. This suggests that most of the identified defense related proteins are not targeted by the proteasome under these conditions and that the epidermal cells remain in a responsive state to counteract biotic stress. The categories *cutin biosynthetic process* and *cell wall pectin biosynthetic process* include proteins that were also identified to accumulate after inhibition of the proteasome (Table 3). These are the enzymes GLYCEROL-3-PHOSPHATE-SN-2-ACYLTRANSFERASE (GPAT4, AT1G01610) important for cutin biogenesis (Li et al., 2007; Yang et al., 2010), DEFICIENT IN CUTIN FERULATE (DCF, AT3G48720) that catalyzes the feruloylation of ω -hydroxy fatty acids, which are one type of cutin monomers (Rautengarten et al., 2012) as well as GALACTURONOSYLTRANSFERASE 1 (GAUT1, AT3G61130) and 8 (GAUT8/QUA1, AT3G25140). GAUT1 functions in an enzymatic complex that catalyzes the elongation step of homogalacturonan synthesis. GAUT8 may also function in homogalacturonan synthesis, and it was suggested to work in the hemicellulose pathway because *gaut8* mutants have a marked reduction in xylan synthase activity (Atmodjo et al., 2013). GAUT1 and GAUT7 are part of a complex that also includes KORRIGAN 1 (KOR1, AT5G49720) (Atmodjo et al., 2011) that is involved in cellulose biosynthesis and may form a link between cellulose and pectin biosynthesis and was also identified to accumulate after SylA treatment. The functions of the proteins identified exclusively in the epidermal protein extract are therefore important for establishing the special characteristics of the epidermal cell wall. Their accumulation after treatment with SylA might either point to their involvement in the general response to the treatment, or indicate that protein degradation by the UPS is one way to regulate epidermal cell wall composition.

The two cinnamyl alcohol dehydrogenases CAD7 and CAD9 that were identified with 94 and 168 spectra, respectively, are involved in lignin biosynthesis (Sibout et al., 2005; Eudes et al., 2006), but their enzymatic activities were reported to be either low or non-detectable (Kim et al., 2007). GUS staining in Arabidopsis showed that the expression or localization pattern of CAD7 and CAD9 changes during development. They are exclusively expressed in the vasculature of leaves in 2 week old seedlings, but at a later developmental stage that corresponds to the age of the leaves used in our experiments, they are localized in

TABLE 2 | Proteins that were exclusively identified in the mesophyll tissue and that were assigned to the significantly over-represented GO categories *photosynthesis* and *starch metabolic process*.

AGI	Name	Description	GO category	Function
AT1G50250	FtsH1	FTSH protease 1	Photosynthesis	Protease
AT5G35220	EGY1	Peptidase M50 family protein		
AT1G73060	LPA3	Low PSII accumulation 3	Photosynthesis, Starch metabolic process	Incorporation of proteins in photosystems
AT4G01800	AGY1	Albino or glassy yellow 1		
AT2G45770	cpFTSY	Signal recognition particle receptor protein, chloroplast (FTSY)		
AT1G08380	PSAO	Photosystem I subunit O		
ATCG01010	NDHF	NADH-Ubiquinone oxidoreductase (complex I), chain 5 protein	Photosystem components	
ATCG01110	NDHH	NAD(P)H dehydrogenase subunit H		
AT1G15980	NDH48	NDH-dependent cyclic electron flow 1		
AT5G01920	STN8	Protein kinase superfamily protein		
AT4G04020	FIB	Fibrillin		Phosphorylation of photosystems
AT1G43670	FBPase	Inositol monophosphatase family protein	Photosynthesis, Starch metabolic process	Photoinhibition
AT5G51820	PGM	Phosphoglucomutase	Starch metabolic process	Conversion triose-P to sucrose
AT5G19220	APL1	ADP glucose pyrophosphorylase large subunit 1		
AT2G41680	NTRC	NADPH-dependent thioredoxin reductase C		
AT5G03650	BE2	Starch branching enzyme 2.2		
AT5G26570	GWD3	Catalytics;carbohydrate kinases;phosphoglucan, water dikinases		Starch biosynthesis
AT3G52180	SEX4	Dual specificity protein phosphatase (DsPTP1) family protein	Starch breakdown	
AT4G00490	BAM2	Beta-amylase 2		
AT4G17090	BAM3	Chloroplast beta-amylase		

trichomes and hydrotodes (Kim et al., 2007). However, the analysis of CAD7 and CAD9 transcript accumulation over anatomy using Genevestigator (Hruz et al., 2008) revealed that their highest transcript levels were specifically detected in the translome of the leaf epidermis (Mustroph et al., 2009). Additionally, the mRNAs of CAD7 and CAD9 and 20 additional epidermis-specific proteins were classified into the epidermis-specific translome cluster (Mustroph et al., 2009) (Supplemental Table 3). The large overlap between exclusively identified epidermal proteins and the epidermis-specific translome emphasizes that information on translated transcripts or proteins is important to assess tissue specific functions.

Comparison of Leaf and Root Epidermal and Vascular Tissue Proteins Reveals Tissue- and Organ-specific Processes

When comparing the identified leaf vascular tissue protein set with the root vascular tissue proteins previously reported by Petricka et al. (2012) we found an overlap of 365 proteins. In the set of 920 proteins that were identified in the root vascular tissue, but not in the leaf vascular tissue, GO biological processes such as *protein targeting to mitochondrion* and *fatty acid β -oxidation* were over-represented, corresponding well with root as a heterotrophic organ. In contrast, the 365 proteins that were identified in the leaf vascular tissue, but not in the

root vascular tissue, were assigned to *glucosinolate biosynthesis*, processes related to photosynthesis, and *response to light stimulus*. The latter two were also over-represented in the 505 proteins that were identified in the leaf epidermis, but not in the root epidermis. In addition, *indole catabolic process* and *defense to fungus* were over-represented in the proteins identified in the leaf epidermis, but not in the root epidermis. In contrast, the 506 proteins that were identified in the root epidermis, but not in the leaf epidermis, were over-represented for *fatty acid β -oxidation*, *amino acid*, and *sterol biosynthesis*, and *thalianol metabolism*. Thalianol is a secondary metabolite that can be detected in Arabidopsis roots (Field and Osbourn, 2008). This demonstrates that the identified tissue type-specific proteins in different organs correspond to the specific functions of the respective tissue type and plant organ.

SylA-dependent Accumulating and Ubiquitylated Proteins in the Mesophyll Reveal that UPS-mediated Protein Degradation Modulates Plastid Protein Composition

Altogether we found that more than 50% of the ubiquitylated and more than 67% of the accumulating proteins detected in the mesophyll are predicted to be localized in the plastid according to SUBAcon in the SUBA3 database (Tanz et al., 2013). Furthermore, three of the four most significantly

TABLE 3 | Proteins that were exclusively identified in the epidermis and that were assigned to the significantly over-represented GO categories *innate immunity* and *immune response*, as well as the categories related to cell wall biosynthesis *lignin biosynthetic process*, *cutin biosynthetic process*, and *cell wall pectin biosynthetic process*.

AGI	Name	Accumulation after SylA	Description	GO category	Function
AT3G15356			Legume lectin family protein		Defense response, incompatible interaction
AT4G16260			Glycosyl hydrolase superfamily protein		
AT1G02920	GSTF7		Glutathione S-transferase 7		
AT3G04720	PR4	–	Pathogenesis-related 4		Systemic acquired resistance
AT2G44490	PEN2		Glycosyl hydrolase superfamily protein	Innate immunity, Immune response	Glucosinolate breakdown
AT5G44070	PCS1		Phytochelatin synthase 1		
AT2G26560	PLA2A		Phospholipase A 2A		Balance between cell death and defense signaling
AT5G52640	HSP90.1	+	Heat shock protein 90.1		R-gene mediated resistance
AT3G11820	SYLP121		Syntaxin of plants 121		Negative regulator of defense pathways
AT4G37980	CAD7, ELI3-1, CADB1		Elicitor-activated gene 3-1	Innate immunity, Immune response, Lignin biosynthetic process	Hypersensitive response Lignin biosynthesis
AT4G39330	CAD9		Cinnamyl alcohol dehydrogenase 9	Lignin biosynthetic process	Lignin biosynthesis
AT1G01610	GPAT4	+	Glycerol-3-phosphate acyltransferase 4	Cutin biosynthetic process	Modification of cutin monomers
AT3G48720	DCF	+	HXXXD-type acyl-transferase family protein		
AT3G25140	QUA1, GAUT8	+	Nucleotide-diphospho-sugar transferases superfamily protein	Cell wall pectin biosynthetic process	Homogalacturonate biosynthesis
AT3G61130	GAUT1	+	Galacturonosyltransferase 1		

In column "Accumulation after SylA" "+" indicates accumulation and "–" decreased protein amount after treatment with SylA.

over-represented GO categories of the proteins that accumulated in the mesophyll after SylA treatment are *protein targeting to chloroplast*, *establishment of protein localization to chloroplast*, and *protein localization to chloroplast*. These categories are also over-represented in the identified ubiquitylated mesophyll proteins. Proteins of the plastid import complexes that accumulated after SylA treatment are TRANSLOCON COMPLEXES AT THE OUTER CHLOROPLAST ENVELOPE MEMBRANE 75-III (TOC75-III, AT3G46740), CHLOROPLAST SIGNAL RECOGNITION PARTICLE 54 (cpSRP54, AT5G03940), the translocase ALBINO 3 (ALB3; AT2G28800), the chloroplast chaperone CHLOROPLAST HEAT SHOCK PROTEIN 70-2 (cpHsc70-2, AT5G03940) and TRANSLOCON COMPLEXES AT THE INNER CHLOROPLAST ENVELOPE MEMBRANE 55 (TIC55, AT2G24820). TOC75-III is the channel protein in the TOC complex and TIC55 is a component of the redox regulon of the TIC complex (Kovács-Bogdán et al., 2010). Stromal Hsp70 is important for the import of precursor proteins (Su and Li, 2010), while the cpSRP consisting of cpSRP54 and cpSRP43 facilitates the passage of imported proteins through the stroma to the thylakoids (Falk and Sinning, 2010). The identified ubiquitylated mesophyll proteins included

the TOC complex protein TOC159 (AT4G02510), cpSRP43 (AT2G47450) and HEAT SHOCK PROTEIN 90.5 (HSP90.5; AT2G04030). Since many of the accumulating and ubiquitylated mesophyll proteins are associated with plastid import, we compared the accumulating proteins with the set of proteins that were decreased in a TOC159-deficient mutant and are therefore putative TOC159-dependent substrates (Bischof et al., 2011). The 10 proteins in the overlap are likely proteins that are transported via the plastid import machinery and whose pre-proteins are targets of the UPS (Supplemental Table 2). Additional plastid-localized proteins that accumulated after SylA treatment are part of the plastid proteolysis system and include the proteases ORGANELLAR OLIGOPEPTIDASE (OOP; AT5G65620) and PRESEQUENCE PROTEASE 1 (PREP1; AT3G19170), as well as proteins involved in chloroplast protein degradation, namely the FtsH1 (AT1G50250), FtsH2 (AT2G30950), and FtsH8 (AT1G06430) subunits of the FtsH membrane-bound metalloprotease complex, the DEG2 (AT2G47940) subunit of the Degradation of periplasmic proteins (Deg)/high temperature requirement A (HtrA) protease, and the ClpB3 (AT5G15450) subunit of the CLP proteases. Another subunit of the CLP proteases, ClpP4 (AT5G45390) was found to be ubiquitylated.

Together, the identification of the above proteins as accumulating or ubiquitylated proteins suggests that these proteins are subject to proteasome degradation and indicates that the UPS is involved in modulating the protein composition in the plastid.

Proteins that Accumulate in the Epidermis After SylA Treatment Function in the Non-cyclic Flux Mode of the TCA Cycle

In the list of proteins that accumulated in the epidermis after SylA treatment, GO categories related to tricarboxylic acid (TCA) cycle and cell wall biosynthesis were over-represented. The plant TCA cycle produces not only reducing equivalents and ATP, but also carbon skeletons for the synthesis of amino acids (Millar et al., 2011; Nunes-Nesi et al., 2013). The non-cyclic flux mode of the TCA cycle produces 2-oxoglutarate as precursor for the amino acids glutamate and glutamine, which are the products of nitrate assimilation (Sweetlove et al., 2010). We found a remarkably high overlap in the epidermis between proteins involved in the non-cyclic flux mode of the TCA cycle and proteins that accumulated after SylA treatment (Figure 4). In addition to mitochondrial proteins that are directly involved in the TCA cycle, cytosolic proteins from connected pathways also accumulated after SylA treatment. This suggests that inhibition of the proteasome might deplete the pool of free amino acids, thus leading to increased *de novo* synthesis of glutamine and glutamate. Consistent with

this hypothesis, NITRITE REDUCTASE 1 (NIR1, AT2G15620), which produces the ammonium required for the biosynthesis of glutamine and glutamate in the plastid, and DICARBOXYLATE TRANSPORT 2.1 (DiT2.1, AT5G64290), which transports glutamate from the plastid into the cytosol in exchange for malate (Renné et al., 2003), also accumulated after SylA treatment. The two mitochondrial proteins NADPH:QUINONE OXIDOREDUCTASE (AT3G27890) and a subunit of ATP synthase (AT4G29480) are involved in metabolizing NADH and FADH₂ and generating ATP, respectively, and are thus part of the full TCA cycle. Their decrease after SylA treatment therefore supports the hypothesis of the TCA cycle operating in a non-flux mode in the epidermis. On the other hand, ATP:citrate lyase (ACL), which utilizes citrate to produce acetyl-CoA and thereby removes substrate from the non-cyclic flux mode TCA cycle, accumulates after SylA treatment. Cytosolic acetyl-CoA is required for the biosynthesis of a range of metabolites including elongated fatty acids (Fatland et al., 2005). Increased levels of ACL might therefore suggest an increased requirement for acetyl-CoA after SylA treatment.

None of the described enzymes of the core TCA cycle was identified to be ubiquitylated. However, CYTOSOLIC NADP⁺-DEPENDENT ISOCITRATE DEHYDROGENASE (CICDH, AT1G65930) and the enzymes of glutamate metabolism, GLUTAMATE DEHYDROGENASE 1 and 2 (GDH1, AT5G18170; GDH2, AT5G07440), GLUTAMATE

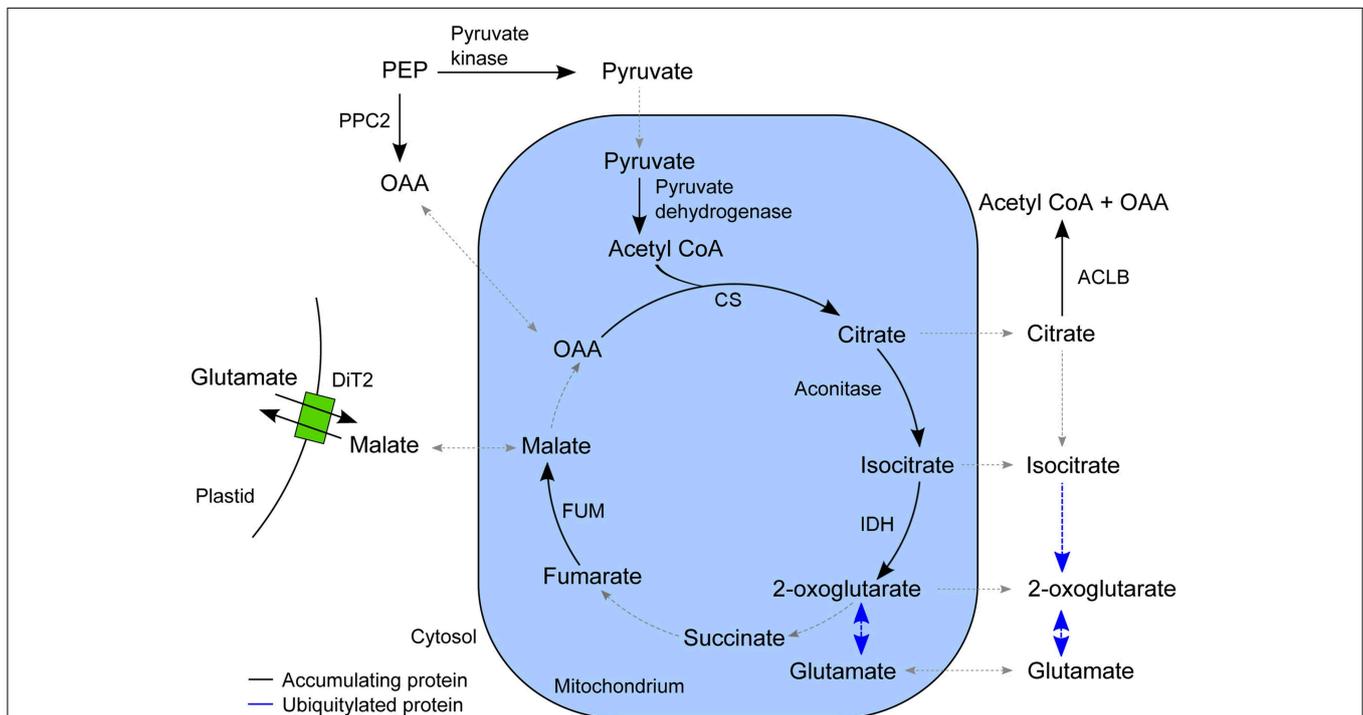


FIGURE 4 | Schematic overview of the TCA cycle and associated pathways. The enzymes that accumulated in the epidermis after SylA treatment are indicated in their proposed subcellular compartment. Reactions for which the corresponding enzyme did not accumulate after SylA treatment are indicated by a dotted line, and those for which the corresponding enzyme was ubiquitylated by a blue line. PPC2, PEP

carboxylase (AT2G42600); pyruvate kinase (AT5G56350); pyruvate dehydrogenase, E2 subunit (AT3G13930); CS, citrate synthase (AT2G44350); aconitase (AT4G35830); IDH, isocitrate dehydrogenase (AT4G35260); FUM, fumarase (AT2G47510); ACLB-2, ATP:citrate lyase, subunit B (AT5G49460) and ACLB-1 (AT3G06650); DiT2, dicarboxylate transport 2.1 (AT5G64290); OAA, oxaloacetate; PEP, phosphoenolpyruvate.

DECARBOXYLASE 2 (AT1G65960) and GLUTAMINE SYNTHETASE 2 (AT5G35630), were ubiquitylated and are therefore putative targets of the UPS.

The expression of GDH subunits was reported to be limited to companion cells of roots and shoots for GDH1 and GDH2, or only roots for GDH3 (Fontaine et al., 2012). However, we identified GDH1 and GDH2 both in the affinity enrichments and in total extracts of epidermis and vasculature, and GDH3 was identified in the total vasculature extract. These localization data are supported by the cell-type specific leaf transcriptome data (Mustroph and Bailey-Serres, 2010), which indicates that the GDH subunit proteins are present in more tissues than previously reported and that their levels and the subunit composition of the GDH enzyme might be regulated by the UPS.

Glucosinolate Biosynthesis Occurs Primarily in the Vascular Tissue and is Regulated by UPS Targeting

In vascular tissue protein extracts we identified an exceptionally high number of 353 ubiquitylated proteins (Table 1). Apart from various response pathways, the GO category *glucosinolate biosynthetic process* was over-represented in this set of proteins. This category was also over-represented in the list of proteins that were exclusively identified in the vascular but not in the epidermal or mesophyll tissue protein extracts (Table 4). Glucosinolates are derived from different amino acids, including methionine and phenylalanine. Depending on the amino acid precursor, they are assigned to different groups. Here, we focus on the biosynthesis of aliphatic glucosinolates, since the identified ubiquitylated proteins mainly catalyze reactions in this pathway.

TABLE 4 | Proteins involved in glucosinolate metabolism identified in the tissue protein extracts and their ubiquitin affinity enrichments.

AGI	Name	Description	Process	Tissue protein extracts		Affinity enrichments		
				Identified	Only in vasculature	Ubiquitylated	Only in vasculature	
AT3G19710	BCAT4	Branched-chain aminotransferase4	Methionine chain elongation	230	x	62	x	
AT5G23010	MAM1	Methylthioalkylmalate synthase 1		156	x	69	x	
AT4G13430	ATLEUC1, IPMI LSU1	Isopropyl malate isomerase large subunit 1		140	x	39	x	
AT2G43100	IPMI SSU2	Isopropylmalate isomerase 2		11	x			
AT3G58990	IPMI SSU3	Isopropylmalate isomerase 1		17	x			
AT5G14200	IMD1,IPMDH1	Isopropylmalate dehydrogenase 1		257	x	80	x	
AT3G49680	BCAT3	Branched-chain aminotransferase3		14	x	14	x	
AT1G16410	CYP79F1	Cytochrome p450 79f1	Glucone core structure formation	102	x			
AT4G13770	CYP83A1	Cytochrome P450, family 83, subfamily A, polypeptide 1		146	x			
AT1G78370	GSTU20	Glutathione S-transferase TAU 20		179	x	48	x	
AT4G30530	GGP1	Gamma glutamyl peptidase 1		47		33		
AT2G20610	SUR1	Tyrosine transaminase family protein, superroot 1		182	x	60	x	
AT2G31790	UGT74C1	UDP-Glycosyltransferase superfamily protein		38	x	11	x	
AT1G74090	ST5b, SOT18	Desulfo-glucosinolate sulfotransferase 18		37	x	20	x	
AT1G18590	ST5c, SOT17	Sulfotransferase 17		61	x	14	x	
AT1G65860	FMO GSOX1	Flavin-monooxygenase glucosinolate S-oxygenase 1		Secondary side-chain modification			6	x
AT1G62560	FMO GSOX3	Flavin-monooxygenase glucosinolate S-oxygenase 3			13	x	15	x
AT2G25450	GS-OH	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	57			27	x	
AT2G43910	ATHOL1	Harmless to ozone layer 1	Glucosinolate degradation	16				
AT5G26000	TGG1	Thioglucoside glucohydrolase 1		1815		225		
AT5G25980	TGG2	Glucoside glucohydrolase 2		515		99		

Proteins were assigned to the different processes in glucosinolate biosynthesis and breakdown. In columns "Identified" and "Ubiquitylated" the numbers of spectra are indicated with which the protein was identified. "x" indicates exclusive identification of the respective protein in the vascular protein extract.

The biosynthesis of aliphatic glucosinolates involves three major steps (Grubb and Abel, 2006; Sawada et al., 2009; Sønderby et al., 2010). The first step is methionine chain elongation, which includes methionine deamination, repeated condensation, isomerization, oxidative decarboxylation and transamination. The proteins involved in this step were exclusively identified in vascular tissue protein extract and in affinity enriched ubiquitylated proteins from vascular tissue (Table 4, Figure 5). In the second step the glucone core structure is formed, which involves the incorporation of sulfur. All of the identified enzymes of the core biosynthetic process except for GGP1 were also identified exclusively in the vascular tissue protein extract or after affinity purification (Table 4, Figure 5). The third and last step in glucosinolate biosynthesis is secondary side chain modification. The identified enzymes in this process were found to be ubiquitylated exclusively in the vasculature tissue (Figure 5; Table 4). In summary, we identified most of the enzymes that catalyze the core biosynthetic steps of glucosinolate biosynthesis only in vascular tissue extracts and in affinity-enriched ubiquitylated proteins from the vascular tissue.

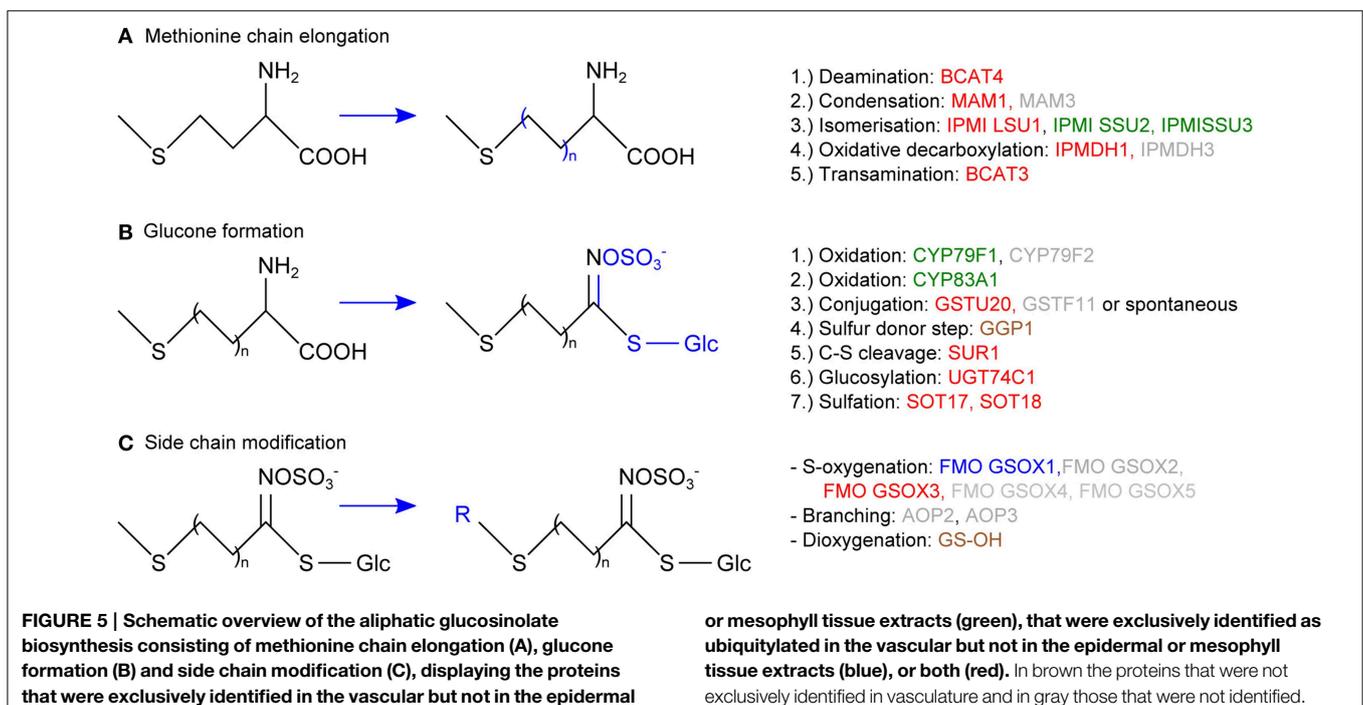
Glucosinolate breakdown is catalyzed by myrosinases and produces toxic isothiocyanates and nitriles and other reactive products that are important in plant defense. Due to the toxicity of the degradation products glucosinolates and myrosinases are located in different cell types and only brought together after tissue damage or transport of glucosinolates (Grubb and Abel, 2006; Halkier and Gershenzon, 2006; Wittstock and Burow, 2010). While we found a strong tissue-type specificity in glucosinolate biosynthesis, the localization of the two Arabidopsis myrosinases TGG1 and TGG2 is more widespread, as they were identified in all tissue-specific protein extracts with highest levels in the epidermis, and also as ubiquitylated

proteins (Table 4). This broad expression pattern of the TGG1 and TGG2 proteins corresponds very well with their transcript translation data in the Arabidopsis translome atlas (Mustroph and Bailey-Serres, 2010), although previous studies had found their expression to be limited to guard cells and phloem myrosin cell idioblasts (Koroleva et al., 2000; Andréasson et al., 2001; Husebye et al., 2002; Thangstad et al., 2004; Zhao et al., 2008).

Following myrosinase action, the S-adenosyl-L-methionine-dependent methyltransferase HARMLESS TO OZONE LAYER 1 (ATHOL1) converts thiocyanate to methylthiocyanate (Nagatoshi and Nakamura, 2009). ATHOL1 was the only protein in the glucosinolate pathway that accumulated after SylA treatment. Together, the large number of identified ubiquitylated enzymes suggests that glucosinolate synthesis and metabolism are tightly controlled by the UPS.

Discussion

Cellular processes and responses that are analyzed in the total leaf context cannot reveal the contribution of specific tissues. We therefore developed a method to effectively separate different leaf tissues and to investigate tissue-specific processes and the response to treatment of the leaf with SylA, which inhibits the proteasome. The Meselect method described here is an effective, high-yielding method to separate mesophyll, epidermal and vascular tissues. We demonstrated that the resulting tissue type-specific protein extracts had essentially no contaminations from other tissues. Since the separation of the three tissues takes only approximately 1 h, it provides a clear advantage over other methods such as FACS. In addition, Meselect does not require the use of specific fluorescent reporter lines and is therefore compatible with experimental workflows using wild type or



any transgenic plants. Meselect furthermore allows the isolation of the three different tissues from the same leaves, permitting comparisons between the tissue types and the elucidation of tissue type-specific processes. Employing the Meselect method reduced the complexity of the total leaf protein samples and therefore increased the sensitivity of the analysis. Our approach resulted in the identification of proteins that were exclusively found in the respective tissue types, of proteins that accumulated after SylA treatment and of proteins that are ubiquitylated. Based on the resulting protein lists we could assign specific processes and UPS targets to each of the three tissues.

The proteins identified in the tissue-type specific extracts are indicative of specific functions of the respective tissue. For example, proteins exclusively identified in mesophyll are related to photosynthesis as well as starch and sugar metabolic processes, and proteins in the epidermis function in defense and leaf protection. The tissue localization of other proteins that we detected in several tissue types, e.g., GDH subunits, CAD7/CAD9 and TGG1/TGG2 corresponded with the localization information in the cell-type specific leaf transcriptome atlas (Mustroph and Bailey-Serres, 2010). These proteins are therefore more ubiquitously expressed than previously reported. This supports our view that tissue-specific protein expression data enhance our understanding of tissue-specific processes and that tissue-type specific protein localization cannot be inferred from transcript expression data only.

Interestingly, many of the identified ubiquitylated and SylA-dependent accumulating proteins in the mesophyll are localized to the plastid. Cytosolic precursors of plastid-localized proteins are known to be degraded by the proteasome following ubiquitylation (Shen et al., 2007a,b; Lee et al., 2009, 2013). A direct interaction between the transit peptide and the 26S proteasome has been reported for certain plastid-localized proteins that may target plastid protein precursors for degradation if they are not imported (Sako et al., 2014). We therefore suggest that the ubiquitylated and SylA-dependent accumulating proteins we identified are pre-proteins that would normally be ubiquitylated and degraded in the cytosol. Furthermore, plastid import complexes themselves seem to be substrates of the UPS because E3 ligase SP1 associates with TOC complexes and mediates the ubiquitylation and degradation of TOC subunits such as TOC159 (Ling et al., 2012). We identified TOC159 and additional proteins associated with plastid transport to be ubiquitylated or accumulating after SylA treatment. UPS-mediated protein degradation might therefore modulate plastid protein composition not only by the degradation of cytosolic precursor proteins, but also by reorganization of the protein import machinery.

The UPS may also affect plastid protein homeostasis by degrading enzymes of the plastid proteolysis system involved in transit peptide cleavage and recycling, protein maturation and degradation (Adam et al., 2006; Van Wijk, 2015). After import, transit peptides of plastid-localized proteins are cleaved off in the stroma and are subsequently degraded by proteases including PREP1 and OOP (Richter and Lamppa, 1998; Kmiec and Glaser, 2012), which accumulated after inhibition of the proteasome. Protein homeostasis in the chloroplast depends on protein

degradation systems of the FtsH complex, the Clp protease system, and additional proteases (Adam et al., 2006; Van Wijk, 2015). For example, the soluble protease DEG2 in cooperation with FtsH subunits is involved in repair of photosystem II by degrading the D1 reaction center protein that is damaged by light and needs to be replaced (Zaltsman et al., 2005; Kato et al., 2012). ClpB3 was shown to be essential for normal chloroplast development (Lee et al., 2007; Singh and Grover, 2010). It has also been reported that the UPS targets the precursors of the plastid proteases FtsH1 and ClpP4 for degradation (Shen et al., 2007a,b). Our finding that several proteins of the chloroplast protein degradation systems accumulated after SylA treatment suggests that the UPS is involved in plastid protein homeostasis by controlling the cytosolic accumulation of precursor proteins that themselves have proteolytic activities.

Glucosinolate biosynthesis and metabolism is an excellent example for tissue type-specific localization of proteins. Glucosinolates are sulfur-containing secondary compounds that participate in plant defense. Upon mechanical wounding of leaves, glucosinolates are hydrolysed by myrosinases that produce various toxic breakdown products, especially isothiocyanates and nitriles (Grubb and Abel, 2006; Halkier and Gershenzon, 2006). Because of the toxicity of glucosinolate breakdown products glucosinolate biosynthetic enzymes and myrosinases should not co-localize. While glucosinolates are mainly stored in vascular S-cells, myrosinases were reported to be expressed in guard cells and in idioblasts, which are located in the phloem parenchyma (Koroleva et al., 2000; Andréasson et al., 2001; Husebye et al., 2002; Thangstad et al., 2004; Grubb and Abel, 2006; Zhao et al., 2008; Wittstock and Burow, 2010). Correspondingly, we identified most of the aliphatic glucosinolate biosynthesis enzymes specifically in vascular tissue extracts. In contrast to previous reports of results from promoter:GUS expression constructs (Husebye et al., 2002; Thangstad et al., 2004), we detected the myrosinases TGG1 and TGG2 in all tissue types with highest levels in the epidermis. In Arabidopsis, vegetative rosette leaves are the major site of aliphatic, methionine-derived glucosinolate biosynthesis and storage (Andersen et al., 2013). Accordingly, we found the GO category *glucosinolate biosynthetic process* over-represented for proteins that were identified in the leaf, but not in the root vasculature (Petricka et al., 2012). As many of the glucosinolate biosynthetic enzymes were found to be ubiquitylated, the levels of these enzymes seem to be controlled by the UPS. Further experiments are required to clarify if targeted degradation regulates the tissue-type specific protein expression of the glucosinolate biosynthetic enzymes and their increased accumulation after the plant defense response has been triggered.

In the epidermis we found that many proteins of the non-cyclic flux mode of the TCA cycle accumulated after inhibition of the proteasome. In normal and carbon starvation conditions the TCA cycle drives ATP synthesis by the oxidation of respiratory substrates. Also, energy and stress signaling are thought to converge because under conditions of carbon deprivation, protein degradation is a key process for recycling cellular molecules (Baena-González and Sheen, 2008). Furthermore,

UPS-mediated protein degradation has an important role in adaptation to carbon and nitrogen availability (Kang and Turano, 2003; Sato et al., 2011). However, in the non-cyclic flux mode, the TCA cycle can provide C skeletons for the biosynthesis of the primary amino acids glutamine and glutamate as well as other metabolites (Sweetlove et al., 2010). One of the functions of the proteasome system is to maintain the free amino acid pools that are needed for protein synthesis through continuous degradation of proteins (Vierstra, 1996; Zhang et al., 2014). Complete recycling of amino acids first involves the partial cleavage of proteins by the UPS followed by further degradation to free amino acids by various endo- and exopeptidases (Book et al., 2005). Inhibition of the proteasome therefore may deplete the pools of free amino acids. It was suggested that the supply of free amino acids is linked to the rate of proteolysis and that low amino acid supplies are monitored in the cell by the increasing levels of uncharged tRNAs (Vierstra, 1996). Depleted pools of free amino acids after proteasome inhibition may therefore lead to enhanced *de novo* synthesis of the primary amino acids glutamine and glutamate through activation of the non-cyclic flux mode of the TCA cycle. The mode of the TCA cycle might therefore switch between providing 2-oxoglutarate for amino acid synthesis or consuming 2-oxoglutarate for the provision of pyruvate, depending on the metabolic status of the cell and the availability of carbon, nitrogen, and amino acids. The coordination of carbon and nitrogen metabolism involves the inter-conversion of keto acids and amino acids, which is catalyzed by the enzyme glutamate dehydrogenase (GDH) (Melo-Oliveira et al., 1996; Forde and Lea, 2007; Miyashita and Good, 2008). We found that the GDH subunit proteins are present in all three leaf tissues and that they are possible targets of the UPS. In addition, CICDH that converts isocitrate to 2-oxoglutarate in the cytosol was also ubiquitinated. For darkened leaves, citrate is known to be catabolized through the TCA cycle, but in illuminated leaves, some of the produced citrate is exported from the mitochondrion to the cytosol, bypassing the mitochondrial aconitase and isocitrate dehydrogenase steps, to produce 2-oxoglutarate via cytosolic reactions (Hanning and Heldt, 1993; Lee et al., 2010). The ubiquitylation of the cytosolic enzymes for glutamate and glutamine synthesis indicates that the

cytosolic part of this pathway is a direct target of the UPS. It will be interesting to determine which E3 ligases are responsible for the ubiquitylation of these proteins and what role targeted protein degradation has for regulating the mode of the TCA cycle.

In summary, the Meselect method effectively and rapidly separates epidermis, vasculature, and mesophyll tissues from the same leaf in approximately 1 h and yields high tissue amounts for proteomics. Analysis of the tissue type-specific protein localization revealed novel insights into tissue-specific processes. Moreover, the quantitative information on tissue-specific protein level changes and the types of ubiquitylated proteins that accumulated after inhibition of the proteasome by SylA has expanded our understanding of UPS-mediated control of protein accumulation in leaves. The data also support the view that protein degradation is an important mechanism for optimizing functional tissue-specific proteomes.

Author Contributions

JS and KB designed research; JS performed research; JS and KB analyzed data; JS, WG, and KB wrote the paper and approved the final version to be published.

Acknowledgments

We thank the Functional Genomics Center Zurich for infrastructure and technical support, Dr. Alex Webb (University of Cambridge, UK) for providing seeds of GAL4-GFP enhancer trap lines JR11-2, KC274 and KC464, and R. Dudler (University of Zurich, Switzerland) for providing Syringolin A. We also thank Johannes Fütterer for critical reading of the manuscript and helpful discussions. This work was supported by ETH Zurich and SNF grant 31003A awarded to KB.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2015.00376/abstract>

References

- Adam, Z., Rudella, A., and van Wijk, K. J. (2006). Recent advances in the study of Clp, FtsH and other proteases located in chloroplasts. *Curr. Opin. Plant Biol.* 9, 234–240. doi: 10.1016/j.pbi.2006.03.010
- Andersen, T. G., Nour-Eldin, H. H., Fuller, V. L., Olsen, C. E., Burrow, M., and Halkier, B. A. (2013). Integration of biosynthesis and long-distance transport establish organ-specific glucosinolate profiles in vegetative *Arabidopsis*. *Plant Cell* 25, 3133–3145. doi: 10.1105/tpc.113.110890
- Andréasson, E., Bolt Jørgensen, L., Höglund, A. S., Rask, L., and Meijer, J. (2001). Different myrosinase and idioblast distribution in *Arabidopsis* and *Brassica napus*. *Plant Physiol.* 127, 1750–1763. doi: 10.1104/pp.010334
- Atmodjo, M. A., Hao, Z., and Mohnen, D. (2013). Evolving views of pectin biosynthesis. *Annu. Rev. Plant Biol.* 64, 747–779. doi: 10.1146/annurev-arplant-042811-105534
- Atmodjo, M. A., Sakuragi, Y., Zhu, X., Burrell, A. J., Mohanty, S. S., Atwood J. A. III, et al. (2011). Galacturonosyltransferase (GAUT1) and GAUT7 are the core of a plant cell wall pectin biosynthetic homogalacturonan:galacturonosyltransferase complex. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20225–20230. doi: 10.1073/pnas.1112816108
- Baena-González, E., and Sheen, J. (2008). Convergent energy and stress signaling. *Trends Plant Sci.* 13, 474–482. doi: 10.1016/j.tplants.2008.06.006
- Baerenfaller, K., Grossmann, J., Grobei, M. A., Hull, R., Hirsch-Hoffmann, M., Yalovsky, S., et al. (2008). Genome-scale proteomics reveals *Arabidopsis thaliana* gene models and proteome dynamics. *Science* 320, 938–941. doi: 10.1126/science.1157956
- Baerenfaller, K., Hirsch-Hoffmann, M., Svozil, J., Hull, R., Russenberger, D., Bischof, S., et al. (2011). pep2pro: a new tool for comprehensive proteome data analysis to reveal information about organ-specific proteomes in *Arabidopsis thaliana*. *Integr. Biol. (Camb)*. 3, 225–237. doi: 10.1039/c0ib00078g
- Baerenfaller, K., Massonnet, C., Walsh, S., Baginsky, S., Bühlmann, P., Hennig, L., et al. (2012). Systems-based analysis of *Arabidopsis* leaf growth reveals adaptation to water deficit. *Mol. Syst. Biol.* 8:606. doi: 10.1038/msb.2012.39

- Barel, G., and Ginzberg, I. (2008). Potato skin proteome is enriched with plant defence components. *J. Exp. Bot.* 59, 3347–3357. doi: 10.1093/jxb/ern184
- Bauer, S., Grossmann, S., Vingron, M., and Robinson, P. N. (2008). Ontologizer 2.0—a multifunctional tool for GO term enrichment analysis and data exploration. *Bioinformatics* 24, 1650–1651. doi: 10.1093/bioinformatics/btn250
- Birnbaum, K., Jung, J. W., Wang, J. Y., Lambert, G. M., Hirst, J. A., Galbraith, D. W., et al. (2005). Cell type-specific expression profiling in plants via cell sorting of protoplasts from fluorescent reporter lines. *Nat. Methods* 2, 615–619. doi: 10.1038/nmeth0805-615
- Birnbaum, K., Shasha, D. E., Wang, J. Y., Jung, J. W., Lambert, G. M., Galbraith, D. W., et al. (2003). A gene expression map of the Arabidopsis root. *Science* 302, 1956–1960. doi: 10.1126/science.1090022
- Bischof, S., Baerenfaller, K., Wildhaber, T., Troesch, R., Vidi, P.-A., Roschitzki, B., et al. (2011). Plastid Proteome Assembly without Toc159: Photosynthetic Protein Import and Accumulation of N-Acetylated Plastid Precursor Proteins. *Plant Cell* 23, 3911–3928. doi: 10.1105/tpc.111.092882
- Book, A. J., Yang, P., Scaff, M., Smith, L. M., and Vierstra, R. D. (2005). Tripeptidyl peptidase II. An oligomeric protease complex from Arabidopsis. *Plant Physiol.* 138, 1046–1057. doi: 10.1104/pp.104.057406
- Brady, S. M., Orlando, D. A., Lee, J.-Y., Wang, J. Y., Koch, J., Dinneny, J. R., et al. (2007). A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science* 318, 801–806. doi: 10.1126/science.1146265
- Brandt, S. P. (2005). Microgenomics: gene expression analysis at the tissue-specific and single-cell levels. *J. Exp. Bot.* 56, 495–505. doi: 10.1093/jxb/eri066
- Brechenmacher, L., Lee, J., Sachdev, S., Song, Z., Nguyen, T. H. N., Joshi, T., et al. (2009). Establishment of a protein reference map for soybean root hair cells. *Plant Physiol.* 149, 670–682. doi: 10.1104/pp.108.131649
- Dinneny, J. R., Long, T. A., Wang, J. Y., Jung, J. W., Mace, D., Pointer, S., et al. (2008). Cell identity mediates the response of Arabidopsis roots to abiotic stress. *Science* 320, 942–945. doi: 10.1126/science.1153795
- Endo, M., Shimizu, H., Nohales, M. A., Araki, T., and Kay, S. A. (2014). Tissue-specific clocks in Arabidopsis show asymmetric coupling. *Nature* 515, 419–422. doi: 10.1038/nature13919
- Eudes, A., Pollet, B., Sibout, R., Do, C.-T., Séguin, A., Lapiere, C., et al. (2006). Evidence for a role of AtCAD 1 in lignification of elongating stems of *Arabidopsis thaliana*. *Planta* 225, 23–39. doi: 10.1007/s00425-006-0326-9
- Falk, S., and Sinning, I. (2010). cpSRP43 is a novel chaperone specific for light-harvesting chlorophyll a,b-binding proteins. *J. Biol. Chem.* 285, 21655–21661. doi: 10.1074/jbc.C110.132746
- Fatland, B. L., Nikolau, B. J., and Wurtele, E. S. (2005). Reverse genetic characterization of cytosolic acetyl-CoA generation by ATP-citrate lyase in Arabidopsis. *Plant Cell* 17, 182–203. doi: 10.1105/tpc.104.026211
- Field, B., and Osbourn, A. E. (2008). Metabolic diversification-independent assembly of operon-like gene clusters in different plants. *Science* 320, 543–547. doi: 10.1126/science.1154990
- Fontaine, J.-X., Terce-Laforgue, T., Armengaud, P., Clement, G., Renou, J.-P., Pelletier, S., et al. (2012). Characterization of a NADH-dependent glutamate dehydrogenase mutant of Arabidopsis demonstrates the key role of this enzyme in root carbon and nitrogen metabolism. *Plant Cell* 24, 4044–4065. doi: 10.1105/tpc.112.103689
- Forde, B. G., and Lea, P. J. (2007). Glutamate in plants: metabolism, regulation, and signalling. *J. Exp. Bot.* 58, 2339–2358. doi: 10.1093/jxb/ern121
- Galbraith, D. W. (2007). Protoplast analysis using flow cytometry and sorting. *Flow Cytom. Plant Cells Anal. Genes Chromosom. Genomes* 231–250. doi: 10.1002/9783527610921.ch10
- Gardner, M. J., Baker, A. J., Assie, J.-M., Poethig, R. S., Haseloff, J. P., and Webb, A. A. R. (2009). GAL4 GFP enhancer trap lines for analysis of stomatal guard cell development and gene expression. *J. Exp. Bot.* 60, 213–226. doi: 10.1093/jxb/ern292
- Grobei, M. A., Qeli, E., Brunner, E., Rehrauer, H., Zhang, R., Roschitzki, B., et al. (2009). Deterministic protein inference for shotgun proteomics data provides new insights into Arabidopsis pollen development and function. *Genome Res.* 19, 1786–1800. doi: 10.1101/gr.089060.108
- Groll, M., Schellenberg, B., Bachmann, A. S., Archer, C. R., Huber, R., Powell, T. K., et al. (2008). A plant pathogen virulence factor inhibits the eukaryotic proteasome by a novel mechanism. *Nature* 452, 755–758. doi: 10.1038/nature06782
- Gronlund, J. T., Eyres, A., Kumar, S., Buchanan-Wollaston, V., and Gifford, M. L. (2012). Cell specific analysis of Arabidopsis leaves using fluorescence activated cell sorting. *J. Vis. Exp.* 68:4214. doi: 10.3791/4214
- Grubb, C. D., and Abel, S. (2006). Glucosinolate metabolism and its control. *Trends Plant Sci.* 11, 89–100. doi: 10.1016/j.tplants.2005.12.006
- Halkier, B. A., and Gershenzon, J. (2006). Biology and biochemistry of glucosinolates. *Annu. Rev. Plant Biol.* 57, 303–333. doi: 10.1146/annurev.arplant.57.032905.105228
- Hanning, I., and Heldt, H. W. (1993). On the function of mitochondrial metabolism during photosynthesis in spinach (*Spinacia oleracea* L.) leaves (partitioning between respiration and export of redox equivalents and precursors for nitrate assimilation products). *Plant Physiol.* 103, 1147–1154. doi: 10.1104/pp.103.4.1147
- Hirsch-Hoffmann, M., Gruissem, W., and Baerenfaller, K. (2012). pep2pro: the high-throughput proteomics data processing, analysis, and visualization tool. *Front. Plant Sci.* 3:123. doi: 10.3389/fpls.2012.00123
- Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., et al. (2008). Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv. Bioinforma.* 2008:420747. doi: 10.1155/2008/420747
- Husebye, H., Chadchawan, S., Winge, P., Thangstad, O. P., and Bones, A. M. (2002). Guard cell- and phloem idioblast-specific expression of thioglucoside glucohydrolase 1 (myrosinase) in Arabidopsis. *Plant Physiol.* 128, 1180–1188. doi: 10.1104/pp.010925
- Kang, J., and Turano, F. J. (2003). The putative glutamate receptor 1.1 (AtGLR1.1) functions as a regulator of carbon and nitrogen metabolism in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 100, 6872–6877. doi: 10.1073/pnas.1030961100
- Kaspar, S., Matros, A., and Mock, H.-P. (2010). Proteome and flavonoid analysis reveals distinct responses of epidermal tissue and whole leaves upon UV-B radiation of barley (*Hordeum vulgare* L.) seedlings. *J. Proteome Res.* 9, 2402–2411. doi: 10.1021/pr901113z
- Kato, Y., Sun, X., Zhang, L., and Sakamoto, W. (2012). Cooperative D1 degradation in the photosystem II repair mediated by chloroplastic proteases in Arabidopsis. *Plant Physiol.* 159, 1428–1439. doi: 10.1104/pp.112.199042
- Kim, S.-J., Kim, K.-W., Cho, M.-H., Franceschi, V. R., Davin, L. B., and Lewis, N. G. (2007). Expression of cinnamyl alcohol dehydrogenases and their putative homologues during *Arabidopsis thaliana* growth and development: lessons for database annotations? *Phytochemistry* 68, 1957–1974. doi: 10.1016/j.phytochem.2007.02.032
- Kmiec, B., and Glaser, E. (2012). A novel mitochondrial and chloroplast peptidase, PreP. *Physiol. Plant.* 145, 180–186. doi: 10.1111/j.1399-3054.2011.01531.x
- Koroleva, O. A., and Cramer, R. (2011). Single-cell proteomic analysis of glucosinolate-rich S-cells in *Arabidopsis thaliana*. *Methods* 54, 413–423. doi: 10.1016/j.ymeth.2011.06.005
- Koroleva, O. A., Davies, A., Deeken, R., Thorpe, M. R., Tomos, A. D., and Hedrich, R. (2000). Identification of a new glucosinolate-rich cell type in Arabidopsis flower stalk. *Plant Physiol.* 124, 599–608. doi: 10.1104/pp.124.2.599
- Kovács-Bogdán, E., Soll, J., and Bölter, B. (2010). Protein import into chloroplasts: the Tic complex and its regulation. *Biochim. Biophys. Acta* 1803, 740–747. doi: 10.1016/j.bbamcr.2010.01.015
- Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., et al. (2012). The Arabidopsis information resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.* 40, D1202–D1210. doi: 10.1093/nar/gkr1090
- Lee, C. P., Eubel, H., and Millar, A. H. (2010). Diurnal changes in mitochondrial function reveal daily optimization of light and dark respiratory metabolism in Arabidopsis. *Mol. Cell. Proteomics* 9, 2125–2139. doi: 10.1074/mcp.M110.001214
- Lee, D. W., Jung, C., and Hwang, I. (2013). Cytosolic events involved in chloroplast protein targeting. *Biochim. Biophys. Acta* 1833, 245–252. doi: 10.1016/j.bbamcr.2012.03.006
- Lee, J., He, K., Stolz, V., Lee, H., Figueroa, P., Gao, Y., et al. (2007). Analysis of transcription factor HY5 genomic binding sites revealed its hierarchical role in light regulation of development. *Plant Cell* 19, 731–749. doi: 10.1105/tpc.106.047688

- Lee, S., Lee, D. W., Lee, Y., Mayer, U., Stierhof, Y.-D., Lee, S., et al. (2009). Heat shock protein cognate 70-4 and an E3 ubiquitin ligase, CHIP, mediate plastid-destined precursor degradation through the ubiquitin-26S proteasome system in Arabidopsis. *Plant Cell* 21, 3984–4001. doi: 10.1105/tpc.109.071548
- Li, Y., Beisson, F., Koo, A. J. K., Molina, I., Pollard, M., and Ohlrogge, J. (2007). Identification of acyltransferases required for cutin biosynthesis and production of cutin with suberin-like monomers. *Proc. Natl. Acad. Sci. U.S.A.* 104, 18339–18344. doi: 10.1073/pnas.0706984104
- Ling, Q., Huang, W., Baldwin, A., and Jarvis, P. (2012). Chloroplast biogenesis is regulated by direct action of the ubiquitin-proteasome system. *Science* 338, 655–659. doi: 10.1126/science.1225053
- Manzano, C., Abraham, Z., López-Torrejón, G., and Pozo, J. C., Del (2008). Identification of ubiquitinated proteins in Arabidopsis. *Plant Mol. Biol.* 68, 145–158. doi: 10.1007/s11103-008-9358-9
- Marks, M. D., Betancur, L., Gilding, E., Chen, F., Bauer, S., Wenger, J. P., et al. (2008). A new method for isolating large quantities of Arabidopsis trichomes for transcriptome, cell wall and other types of analyses. *Plant J.* 56, 483–492. doi: 10.1111/j.1365-313X.2008.03611.x
- Melo-Oliveira, R., Oliveira, I. C., and Coruzzi, G. M. (1996). Arabidopsis mutant analysis and gene regulation define a nonredundant role for glutamate dehydrogenase in nitrogen assimilation. *Proc. Natl. Acad. Sci. U.S.A.* 93, 4718–4723. doi: 10.1073/pnas.93.10.4718
- Millar, A. H., Whelan, J., Soole, K. L., and Day, D. A. (2011). Organization and regulation of mitochondrial respiration in plants. *Annu. Rev. Plant Biol.* 62, 79–104. doi: 10.1146/annurev-arplant-042110-103857
- Miyashita, Y., and Good, A. G. (2008). NAD(H)-dependent glutamate dehydrogenase is essential for the survival of *Arabidopsis thaliana* during dark-induced carbon starvation. *J. Exp. Bot.* 59, 667–680. doi: 10.1093/jxb/erm340
- Mustroph, A., and Bailey-Serres, J. (2010). The Arabidopsis transcriptome cell-specific mRNA atlas: Mining suberin and cutin lipid monomer biosynthesis genes as an example for data application. *Plant Signal. Behav.* 5, 320–324. doi: 10.4161/psb.5.3.11187
- Mustroph, A., Zanetti, M. E., Jang, C. J. H., Holtan, H. E., Repetti, P. P., Galbraith, D. W., et al. (2009). Profiling transcriptomes of discrete cell populations resolves altered cellular priorities during hypoxia in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* 106, 18843–18848. doi: 10.1073/pnas.0906131106
- Nagatohi, Y., and Nakamura, T. (2009). Arabidopsis HARMLESS TO OZONE LAYER protein methylates a glucosinolate breakdown product and functions in resistance to *Pseudomonas syringae* pv. *maculicola*. *J. Biol. Chem.* 284, 19301–19309. doi: 10.1074/jbc.M109.001032
- Nakazono, M., Qiu, F., Borsuk, L. A., and Schnable, P. S. (2003). Laser-capture microdissection, a tool for the global analysis of gene expression in specific plant cell types: identification of genes expressed differentially in epidermal cells or vascular tissues of maize. *Plant Cell* 15, 583–596. doi: 10.1105/tpc.008102
- Nestler, J., Schütz, W., and Hochholdinger, F. (2011). Conserved and unique features of the maize (*Zea mays* L.) root hair proteome. *J. Proteome Res.* 10, 2525–2537. doi: 10.1021/pr200003k
- Nunes-Nesi, A., Araújo, W. L., Obata, T., and Fernie, A. R. (2013). Regulation of the mitochondrial tricarboxylic acid cycle. *Curr. Opin. Plant Biol.* 16, 335–343. doi: 10.1016/j.pbi.2013.01.004
- Petricka, J. J., Schauer, M. A., Megraw, M., Breakfield, N. W., Thompson, J. W., Georgiev, S., et al. (2012). The protein expression landscape of the Arabidopsis root. *Proc. Natl. Acad. Sci. U.S.A.* 109, 6811–6818. doi: 10.1073/pnas.1202546109
- R Core Team (2012). *R: A Language and Environment for Statistical Computing*. Available online at: <http://www.r-project.org>
- Rautengarten, C., Ebert, B., Ouellet, M., Nafisi, M., Baidoo, E. E. K., Benke, P., et al. (2012). Arabidopsis deficient in cutin ferulate encodes a transferase required for feruloylation of ω -hydroxy fatty acids in cutin polyester. *Plant Physiol.* 158, 654–665. doi: 10.1104/pp.111.187187
- Renné, P., Dressen, U., Hebbeker, U., Hille, D., Flügge, U.-I., Westhoff, P., et al. (2003). The Arabidopsis mutant *dct* is deficient in the plastidic glutamate/malate translocator DIT2. *Plant J.* 35, 316–331. doi: 10.1046/j.1365-313X.2003.01806.x
- Richter, S., and Lamppa, G. K. (1998). A chloroplast processing enzyme functions as the general stromal processing peptidase. *Proc. Natl. Acad. Sci. U.S.A.* 95, 7463–7468. doi: 10.1073/pnas.95.13.7463
- Sako, K., Yanagawa, Y., Kanai, T., Sato, T., Seki, M., Fujiwara, M., et al. (2014). Proteomic analysis of the 26S proteasome reveals its direct interaction with transit peptides of plastid protein precursors for their degradation. *J. Proteome Res.* 13, 3223–3230. doi: 10.1021/pr401245g
- Sato, T., Maekawa, S., Yasuda, S., and Yamaguchi, J. (2011). Carbon and nitrogen metabolism regulated by the ubiquitin-proteasome system. *Plant Signal. Behav.* 6, 1465–1468. doi: 10.4161/psb.6.10.17343
- Sawada, Y., Kuwahara, A., Nagano, M., Narisawa, T., Sakata, A., Saito, K., et al. (2009). Omics-based approaches to methionine side chain elongation in Arabidopsis: characterization of the genes encoding methylthioalkylmalate isomerase and methylthioalkylmalate dehydrogenase. *Plant Cell Physiol.* 50, 1181–1190. doi: 10.1093/pcp/pcp079
- Schad, M., Lipton, M. S., Giavalisco, P., Smith, R. D., and Kehr, J. (2005). Evaluation of two-dimensional electrophoresis and liquid chromatography–tandem mass spectrometry for tissue-specific protein profiling of laser-microdissected plant samples. *Electrophoresis* 26, 2729–2738. doi: 10.1002/elps.200410399
- Shen, G., Adam, Z., and Zhang, H. (2007a). The E3 ligase AtCHIP ubiquitylates FtsH1, a component of the chloroplast FtsH protease, and affects protein degradation in chloroplasts. *Plant J.* 52, 309–321. doi: 10.1111/j.1365-313X.2007.03239.x
- Shen, G., Yan, J., Pasapula, V., Luo, J., He, C., Clarke, A. K., et al. (2007b). The chloroplast protease subunit ClpP4 is a substrate of the E3 ligase AtCHIP and plays an important role in chloroplast function. *Plant J.* 49, 228–237. doi: 10.1111/j.1365-313X.2006.02963.x
- Sibout, R., Eudes, A., Mouille, G., Pollet, B., Lapierre, C., Jouanin, L., et al. (2005). CINNAMYL ALCOHOL DEHYDROGENASE-C and -D are the primary genes involved in lignin biosynthesis in the floral stem of Arabidopsis. *Plant Cell* 17, 2059–2076. doi: 10.1105/tpc.105.030767
- Singh, A., and Grover, A. (2010). Plant Hsp100/ClpB-like proteins: poorly-analyzed cousins of yeast ClpB machine. *Plant Mol. Biol.* 74, 395–404. doi: 10.1007/s11103-010-9682-8
- Sønderby, I. E., Geu-Flores, F., and Halkier, B. A. (2010). Biosynthesis of glucosinolates—gene discovery and beyond. *Trends Plant Sci.* 15, 283–290. doi: 10.1016/j.tplants.2010.02.005
- Stadler, R., Büttner, M., Ache, P., Hedrich, R., Ivashikina, N., Melzer, M., et al. (2003). Diurnal and light-regulated expression of AtSTP1 in guard cells of Arabidopsis. *Plant Physiol.* 133, 528–537. doi: 10.1104/pp.103.024240
- Streb, S., and Zeeman, S. C. (2012). Starch metabolism in Arabidopsis. *Arab. B.* 10:e0160. doi: 10.1199/tab.0160
- Su, P.-H., and Li, H. (2010). Stromal Hsp70 is important for protein translocation into pea and Arabidopsis chloroplasts. *Plant Cell* 22, 1516–1531. doi: 10.1105/tpc.109.071415
- Sutovsky, P., Manandhar, G., Laurincik, J., Letko, J., Caamaño, J. N., Day, B. N., et al. (2005). Expression and proteasomal degradation of the major vault protein (MVP) in mammalian oocytes and zygotes. *Reproduction* 129, 269–282. doi: 10.1530/rep.1.00291
- Svozil, J., Hirsch-Hoffmann, M., Dudler, R., Gruissem, W., and Baerenfaller, K. (2014). Protein abundance changes and ubiquitylation targets identified after inhibition of the proteasome with Syringolin A. *Mol. Cell. Proteomics* 13, 1523–1536. doi: 10.1074/mcp.M113.036269
- Sweetlove, L. J., Beard, K. F. M., Nunes-Nesi, A., Fernie, A. R., and Ratcliffe, R. G. (2010). Not just a circle: flux modes in the plant TCA cycle. *Trends Plant Sci.* 15, 462–470. doi: 10.1016/j.tplants.2010.05.006
- Takahashi, A., Casais, C., Ichimura, K., and Shirasu, K. (2003). HSP90 interacts with RAR1 and SGT1 and is essential for RPS2-mediated disease resistance in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11777–11782. doi: 10.1073/pnas.2033934100
- Tanz, S. K., Castleden, I., Hooper, C. M., Vacher, M., Small, I., and Millar, H. A. (2013). SUBA3: a database for integrating experimentation and prediction to define the SUBcellular location of proteins in Arabidopsis. *Nucleic Acids Res.* 41, D1185–D1191. doi: 10.1093/nar/gkx1151
- Thangstad, O. P., Gilde, B., Chadchawan, S., Seem, M., Husebye, H., Bradley, D., et al. (2004). Cell specific, cross-species expression of myrosinases in Brassica napus, *Arabidopsis thaliana* and Nicotiana tabacum. *Plant Mol. Biol.* 54, 597–611. doi: 10.1023/B:PLAN.0000038272.99590.10
- Tsukaya, H. (2002). Leaf development. *Arab. B.* 1:e0072. doi: 10.1199/tab.0072
- Van Cutsem, E., Simonart, G., Degand, H., Faber, A.-M., Morsomme, P., and Boutry, M. (2011). Gel-based and gel-free proteomic analysis of

- Nicotiana tabacum trichomes identifies proteins involved in secondary metabolism and in the (a)biotic stress response. *Proteomics* 11, 440–454. doi: 10.1002/pmic.201000356
- Van Wijk, K. J. (2015). Protein maturation and proteolysis in plant plastids, mitochondria, and peroxisomes. *Annu. Rev. Plant Biol.* 66, 75–111. doi: 10.1146/annurev-arplant-043014-115547
- Vierstra, R. D. (1996). Proteolysis in plants: mechanisms and functions. *Plant Mol. Biol.* 32, 275–302. doi: 10.1007/BF00039386
- Vizcaino, J. A., Côté, R. G., Csordas, A., Dianes, J. A., Fabregat, A., Foster, J. M., et al. (2013). The PRoteomics IDentifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.* 41, D1063–D1069. doi: 10.1093/nar/gks1262
- Walley, J. W., Shen, Z., Sartor, R., Wu, K. J., Osborn, J., Smith, L. G., et al. (2013). Reconstruction of protein networks from an atlas of maize seed proteotypes. *Proc. Natl. Acad. Sci. U.S.A.* 110, E4808–E4817. doi: 10.1073/pnas.1319113110
- Wienkoop, S., Zoeller, D., Ebert, B., Simon-Rosin, U., Fisahn, J., Glinski, M., et al. (2004). Cell-specific protein profiling in *Arabidopsis thaliana* trichomes: identification of trichome-located proteins involved in sulfur metabolism and detoxification. *Phytochemistry* 65, 1641–1649. doi: 10.1016/j.phytochem.2004.03.026
- Wittstock, U., and Burow, M. (2010). Glucosinolate breakdown in arabidopsis: mechanism, regulation and biological significance. *Arab. B.* 8:e0134. doi: 10.1199/tab.0134
- Wu, F.-H., Shen, S.-C., Lee, L.-Y., Lee, S.-H., Chan, M.-T., and Lin, C.-S. (2009). Tape-arabidopsis sandwich - a simpler arabidopsis protoplast isolation method. *Plant Methods* 5:16. doi: 10.1186/1746-4811-5-16
- Yang, W., Pollard, M., Li-Beisson, Y., Beisson, F., Feig, M., and Ohlrogge, J. (2010). A distinct type of glycerol-3-phosphate acyltransferase with sn-2 preference and phosphatase activity producing 2-monoacylglycerol. *Proc. Natl. Acad. Sci. U.S.A.* 107, 12040–12045. doi: 10.1073/pnas.0914149107
- Yeats, T. H., Howe, K. J., Matas, A. J., Buda, G. J., Thannhauser, T. W., and Rose, J. K. C. (2010). Mining the surface proteome of tomato (*Solanum lycopersicum*) fruit for proteins associated with cuticle biogenesis. *J. Exp. Bot.* 61, 3759–3771. doi: 10.1093/jxb/erq194
- Zaltsman, A., Ori, N., and Adam, Z. (2005). Two types of FtsH protease subunits are required for chloroplast biogenesis and Photosystem II repair in Arabidopsis. *Plant Cell* 17, 2782–2790. doi: 10.1105/tpc.105.035071
- Zhang, Y., Nicholatos, J., Dreier, J. R., Ricoult, S. J. H., Widenmaier, S. B., Hotamisligil, G. S., et al. (2014). Coordinated regulation of protein synthesis and degradation by mTORC1. *Nature* 513, 440–443. doi: 10.1038/nature13492
- Zhao, Z., Zhang, W., Stanley, B. A., and Assmann, S. M. (2008). Functional proteomics of *Arabidopsis thaliana* guard cells uncovers new stomatal signaling pathways. *Plant Cell* 20, 3210–3226. doi: 10.1105/tpc.108.063263
- Zhu, M., Dai, S., McClung, S., Yan, X., and Chen, S. (2009). Functional differentiation of Brassica napus guard cells and mesophyll cells revealed by comparative proteomics. *Mol. Cell Proteomics* 8, 752–766. doi: 10.1074/mcp.M800343-MCP200

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Svozil, Gruissem and Baerenfaller. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The guard cell metabolome: functions in stomatal movement and global food security

Biswapriya B. Misra¹, Biswa R. Acharya², David Granot³, Sarah M. Assmann² and Sixue Chen^{1,4*}

¹ Department of Biology, Genetics Institute, Plant Molecular and Cellular Biology Program, University of Florida, Gainesville, FL, USA, ² Department of Biology, Pennsylvania State University, PA, USA, ³ Department of Vegetable Research, Institute of Plant Sciences, Agricultural Research Organization, Bet-Dagan, Israel, ⁴ Interdisciplinary Center for Biotechnology Research, University of Florida, Gainesville, FL, USA

OPEN ACCESS

Edited by:

Yariv Brotman,

Max Planck Institute of Molecular Plant Physiology, Germany

Reviewed by:

Norihito Nakamichi,

Nagoya University, Japan

Maria F. Drincovich,

Rosario National University, Argentina

*Correspondence:

Sixue Chen,

Department of Biology, Genetics Institute, Plant Molecular and Cellular Biology Program, University of Florida, Cancer and Genetics Research Complex, Room 438, 2033 Mowry Road, Gainesville, FL 32610, USA
schen@ufl.edu

Specialty section:

This article was submitted to Plant Systems and Synthetic Biology, a section of the journal *Frontiers in Plant Science*

Received: 21 March 2015

Accepted: 28 April 2015

Published: 19 May 2015

Citation:

Misra BB, Acharya BR, Granot D, Assmann SM and Chen S (2015) The guard cell metabolome: functions in stomatal movement and global food security. *Front. Plant Sci.* 6:334. doi: 10.3389/fpls.2015.00334

Guard cells represent a unique single cell-type system for the study of cellular responses to abiotic and biotic perturbations that affect stomatal movement. Decades of effort through both classical physiological and functional genomics approaches have generated an enormous amount of information on the roles of individual metabolites in stomatal guard cell function and physiology. Recent application of metabolomics methods has produced a substantial amount of new information on metabolome control of stomatal movement. In conjunction with other “omics” approaches, the knowledge-base is growing to reach a systems-level description of this single cell-type. Here we summarize current knowledge of the guard cell metabolome and highlight critical metabolites that bear significant impact on future engineering and breeding efforts to generate plants/crops that are resistant to environmental challenges and produce high yield and quality products for food and energy security.

Keywords: stomata, primary metabolites, abscisic acid, phytohormones, lipids, specialized metabolites, food security

Introduction

Guard cells as a unique plant single cell-type perform many functions essential to plant growth and survival. Each pair of guard cells and the regulated pore they enclose, known as a stoma or stomate, provides a conduit for atmospheric photosynthetic gas exchange (CO₂ uptake and O₂ release) and transpirational release of water (H₂O) in terrestrial plants, in addition to defense against pathogenic invasion. Stomatal opening and closing, in which the guard cells actively increase and decrease their volume via turgor changes to regulate the pore size in response to environmental stimuli, are vital processes in maintaining the balance of H₂O loss and CO₂ fixation. While drought stress induces stomatal closure, pathogens exploit stomatal opening to facilitate entry into the leaf (Zeng and He, 2010). Abscisic acid (ABA), CO₂ and blue light mediated stomatal movements have generated tremendous interest in their signaling mechanisms. Each pathway/network has unique components such as distinct receptors and early signaling elements. They also have common components, for example, actual stomatal movement is caused by water influx/efflux mostly driven by solute fluxes through plasma membrane anion channels and K⁺ channels. When the concentrations of solutes decrease in guard cells in the cases of ABA and elevated CO₂, water potential increases in the cells and water flows out, causing a decrease of turgor pressure and closure of the pores. Blue light activates H⁺-ATPases and resultant membrane hyperpolarization drives K⁺ influx, leading to decreased

water potential, increased turgor pressure, and stomatal opening. Please refer to excellent articles published over the years on these signaling mechanisms (e.g., Zeiger and Zhu, 1998; Schroeder et al., 2001; Yu and Assmann, 2014; Zhang et al., 2014; Tian et al., 2015). Although stomatal movements in response to ABA, CO₂, and blue light are well studied, the metabolome of the guard cell is far from catalogued. Endeavors in metabolomic approaches have led to deeper understanding of the biology inherent to several specialized important single-cell types including guard cells (Misra et al., 2014). Recent efforts to comprehend the guard cell metabolome (Jin et al., 2013) and systems biology approaches to identify the critical regulators in stomatal movement (Sun et al., 2014) have provided interesting leads into the intricate regulation of stomatal movement in response to environmental stimuli. Ongoing systems biology approaches, combining modeling and high-throughput experiments, will help to elucidate the mechanisms underlying stomatal control and unravel targets for modulation of stomatal responses to environment (Medeiros et al., 2015).

Many factors pose immense challenges to global food and bioenergy security, including population growth, climate, and environmental changes coupled to land degradation and changes in hydrological resources, essential ecosystem services, and agricultural production systems. Urgent efforts are needed to enhance the resilience of crops to the adverse effects of climate change. Stomata are highly responsive to hormonal and environmental cues, including those associated with climate change: water availability, temperature, and CO₂ concentrations. Thus, understanding the basic biology, the concealed information content, and the connection to functional output of guard cells through multiple -omics approaches such as transcriptomics (Wang et al., 2011), proteomics (Zhao et al., 2008; Zhu et al., 2014) and metabolomics (Jin et al., 2013) is highly relevant to the goal of improving crop productivity and yield in ever changing climatic regimens. Here we briefly review collective efforts to unravel the functional guard cell metabolome (Figure 1), to discover metabolites of convergence and divergence among various environmental cues, to examine the molecular mechanisms of guard cell metabolic regulation, and ultimately to highlight the potential of guard cell biology in harnessing possible solutions for global food and bioenergy security.

Primary Metabolites of Carbon Metabolism in Regulatory Roles of Stomatal Function

Early hypotheses regarding guard cell osmoregulation suggested that sugar generated from starch degradation at dawn is the primary osmolyte that opens stomata (Lloyd, 1908). Upon the discovery that potassium (K⁺) ions, with chloride (Cl⁻) and malate²⁻ as counter anions, are osmolytes that open stomata, a role for sugar in guard cell osmoregulation and stomatal movement was abandoned for several decades (Imamura, 1943; Yamashita, 1952; Fischer, 1968; Humble and Raschke, 1971; Allaway, 1973; Outlaw and Lowry, 1977; Asai et al., 2000). A later study reporting that blue and red (photosynthetic) light can open stomata and are followed by sucrose accumulation in guard cells revived the hypothesis of sucrose as an osmolyte that opens stomata

(Talbot and Zeiger, 1993). A correlation between the decline of K⁺ content in guard cells in the middle of the day concomitantly with an increase in sucrose content further suggested that sucrose is an osmolyte that replaces K⁺ and maintains stomatal opening (Amodeo et al., 1996; Talbot and Zeiger, 1996).

The origin of sucrose in guard cells is not yet clear. Potentially, sucrose could be obtained from guard cell starch degradation, guard cell photosynthesis, or import from mesophyll cells (Gotow et al., 1988; Talbot and Zeiger, 1988; Lawson et al., 2014). It is generally accepted though, that the contribution of sucrose produced from guard cell photosynthesis to the osmotic requirement for stomatal opening is minimal and that most of the sugar or the organic compounds from which sugar can be synthesized is obtained from the mesophyll cells (Outlaw, 1989; Reckmann et al., 1990). When exported out of the mesophyll cells for phloem loading, some sucrose accumulates in the guard cell apoplast (Lu et al., 1995, 1997; Outlaw and De Vlieghere-He, 2001). As a result, the concentration of sucrose in the guard cell apoplast increases as photosynthesis proceeds. This sucrose may be imported by guard cells and contribute to guard cell osmolarity and stomatal opening. But it also has been proposed that as sucrose accumulates in the apoplast, its osmotic effect drives water efflux from guard cells, resulting in a decrease in stomatal apertures in a mechanism that thus inversely coordinates photosynthesis and transpiration rates (Lu et al., 1997; Outlaw and De Vlieghere-He, 2001).

Apoplastic sucrose may enter the guard cells either via sucrose transporters, or via guard cell hexose transporters following sucrose cleavage by apoplastic invertase to yield the hexoses glucose and fructose (Stadler et al., 2003; Weise et al., 2008; Bates et al., 2012). Regardless of its origin, mesophyll or guard cells, sucrose must be cleaved to be metabolized, and the hexoses obtained from sucrose cleavage, glucose and fructose, must be phosphorylated by intracellular hexose phosphorylating enzymes, hexokinases (HXK) and fructokinases (FRK; Dennis and Blakeley, 2000). Glucose can be phosphorylated only by HXK, an enzyme demonstrated to exist in guard cells and to participate in sugar sensing (Moore et al., 2003; Rolland et al., 2006; Granot, 2008). A recent study has shown that sugars such as sucrose, glucose, and fructose, do not exert an apoplastic osmotic effect on guard cells, but rather are sensed within guard cells by HXK to stimulate stomatal closure, thus coordinating photosynthesis and sugar levels with transpiration (Kelly et al., 2013).

The phosphorylated hexoses (hexose-P) within guard cells may be converted to starch or enter glycolysis and the tricarboxylic acid (TCA) cycle to yield energy (ATP) and various metabolites including pyruvate and malate that regulate stomatal movement (Allaway, 1973; Pearson, 1973; Outlaw and Lowry, 1977). Glycolysis is a central metabolic pathway for cellular respiration and generation of energy in the form of ATP. In the glycolytic pathway, 2,3-bisphosphoglycerate-independent phosphoglycerate mutase (iPGAM) catalyzes the interconversion of 3-phosphoglycerate to 2-phosphoglycerate. *Arabidopsis thaliana* double mutants of *iPGAM* genes show hyposensitivity in blue light, ABA, and low CO₂ regulated stomatal movements, confirming a role of glycolysis in guard cell function (Zhao and Assmann, 2011). ABA inhibition of stomatal opening in *Commelina benghalensis* is reversed by exogenous ATP and pyruvate

from guard cell starch degradation or from triose-phosphates produced in guard cell chloroplasts and exported to the cytoplasm where triose-P metabolism yields malate among other metabolites. ABA-stimulated stomatal closure is accompanied by malate disposal through release, gluconeogenesis, or consumption in the TCA cycle, supporting the role of malate as an osmolyte that opens stomata (Dittrich and Raschke, 1977). In the guard cell cytosol, malate can be metabolized into oxaloacetate (OAA) by malate dehydrogenase. Subsequently, phosphoenolpyruvate carboxykinase (PEPCK) can catalyze the production of PEP from OAA that in turn would enter into gluconeogenesis. An isoform of PEPCK, PCK1, is expressed in *A. thaliana* guard cells according to three experimental approaches: PCK1 gene promoter analysis and analyses of the proteome, and transcriptome of guard cell protoplasts (Leonhardt et al., 2004; Penfield et al., 2012; and Zhao et al., 2008). Loss-of-function PCK1 plants (*pck1-2*) show hyposensitivity in response to dark-induced (but not ABA-induced) stomatal closure, indicating the importance of malate metabolism for some stomatal responses (Penfield et al., 2012). Malate produced in photosynthetic tissues may also arrive at and enter the guard cells through malate transporters (Lee et al., 2008). Mesophyll-produced malate also coordinates stomatal behavior with mesophyll photosynthesis, as increasing apoplastic malate activates anion channels that reduce stomatal aperture (Hedrich and Marten, 1993; Fernie and Martinoia, 2009; Araujo et al., 2011). In addition, methylglyoxal, an oxygenated short aldehydic glycolytic intermediate, can induce stomatal closure in *A. thaliana* accompanied by extracellular reactive oxygen species (ROS) production mediated by SHAM-sensitive peroxidases, intracellular ROS accumulation, and suppression of free cytosolic (Ca^{2+}) oscillations (Hoque et al., 2012). These results indicate a strong interconnectivity between central carbon metabolism and ABA signaling in guard cells.

Reactive Oxygen Species Related Metabolites in Guard Cell Signaling

Reactive oxygen species and nitric oxide (NO) are central components of the signaling network regulating stomatal movement in response to ABA, jasmonic acid (JA), darkness, UV, pathogen, and high CO_2 concentrations (Zhang et al., 2001; Desikan et al., 2004, 2006; Zhu et al., 2012; Akter et al., 2013; He et al., 2013; Joudoi et al., 2013; Ou et al., 2014). Upon application of NO-releasing compounds, NO induces dose-dependent stomatal closure. In contrast, NO has also been implicated as a key component in negative feedback regulation of ABA guard cell signaling through S-nitrosylation of OST1 at cysteine 137 and subsequent inactivation of kinase activity that in turn blocks the positive regulatory role of OST1 in ABA signaling (Wang et al., 2015). NO-mediated negative feedback regulation may prevent complete stomatal closure, allowing some basal level of CO_2 uptake and photosynthesis. Hydrogen peroxide (H_2O_2) may also elicit stomatal movement in a similar manner through redox modification of guard cell signaling components. However, experimental data are lacking for this hypothesis. In addition, ascorbic acid (Asc) and glutathione (GSH) are critical in maintaining cellular ROS levels and redox homeostasis (Noctor and Foyer,

1998). Asc is a key antioxidant that scavenges ROS including H_2O_2 . Dehydroascorbate reductase (DHAR) is the key regulatory enzyme that catalyzes the generation of Asc (reduced form) from dehydroascorbate (DAsc, oxidized form) in a reaction that requires GSH. Tobacco DHAR overexpression lines that have elevated levels of reduced Asc in guard cells show hyposensitivity in stomatal response to ABA and H_2O_2 and these plants are drought susceptible. In contrast, DHAR antisense tobacco lines show drought tolerance (Chen and Gallie, 2004). These findings indicate that Asc redox state plays an important regulatory role in ABA and H_2O_2 mediated stomatal responses. Altered redox state and stomatal aperture in mutants defective in GSH synthesis are well established (Okuma et al., 2011; Munemasa et al., 2013). Negative regulation of methyl jasmonate (MeJA)-induced stomatal closure by GSH in *A. thaliana* has been demonstrated (Akter et al., 2013). In addition, GSH peroxidases are known to function as redox transducers as well as scavengers in ABA-mediated stress responses (Miao et al., 2006). Thus, understanding redox changes and their regulation and coordination with stomatal functions would provide new insights into guard cell signaling networks.

Stomatal guard cells have a thick cuticular layer containing high concentrations of wax-bound phenolics that provide protection against UV radiation (Karabourniotis et al., 2001; He et al., 2013) and form a constitutive defense barrier against pathogens and insects. Intracellular phenolics and flavonoids synthesized from the phenylpropanoid pathway are also responsible for cellular defense and pigmentation among other functions. Flavonoids protect plants from UV-B irradiation (Li et al., 1993) and also function as stress-induced antioxidants (Dixon and Paiva, 1995). Flavonols accumulate in guard cells of *A. thaliana*, but not in the surrounding pavement cells (Watkins et al., 2014). Enhanced flavonol content and decreased ROS levels upon ethylene (ET) treatment in guard cells were correlated with a reduction in the rate of stomatal closure in response to ABA. The results suggest that flavonols may quench the ABA-dependent ROS burst (Watkins et al., 2014). Moreover, some flavonoids, such as quercetin, apigenin, and kaempferol, have functions similar to synthetic auxin transport inhibitors, so changes in the synthesis or deposition of specific flavonoids within cells may act to change the rate or direction of auxin transport (Winkel-Shirley, 2002). Given that ABA reduces guard cell auxin concentrations (Jin et al., 2013), it would be interesting to further investigate the interrelationships between flavonoids, ROS, ABA, and guard cell auxin transport.

Role of Lipid Signaling in Stomatal Movement and Development

Lipids are essential for membrane formation and energy storage. In addition, lipids and their metabolites are also important cellular signaling molecules, including in stomatal regulation. For instance, lipid-based secondary messengers that positively regulate guard cell ABA signaling and stomatal closure include phosphatidic acid (PhA), phosphatidylinositol-3-phosphate (PI3P), inositol-1,4,5-trisphosphate (IP3), inositol-6-phosphate (IP6), and sphingolipids (Kim et al., 2010; **Figure 1**).

Phosphoinositides play important roles in guard cell signaling. Phospholipase C (PLC) hydrolyses phosphatidylinositol 4,5-bisphosphate (PIP₂) to produce 1,2-diacylglycerol (DAG) and IP₃. ABA induced the production of IP₃ in *Vicia faba* guard cells (Lee et al., 1996), and cytosolic Ca²⁺ elevation and subsequent stomatal closure occurred upon experimental elevation of cytosolic IP₃ in *Commelina communis* guard cells (Gilroy et al., 1990). Increases in guard cell PI3P and PI 4-phosphate (PI4P; the products of PI 3-kinase (PI3K) and PI 4-kinase (PI4K) activities, respectively) induce stomatal closure mediated by ABA-induced ROS generation (Jung et al., 2002; Park et al., 2003). IP₆ is generated in guard cells in response to ABA. IP₆ is an endomembrane-acting Ca²⁺-release signal that inhibits the inwardly rectifying K⁺ channel, which would then inhibit stomatal opening (Lemtiri-Chlieh et al., 2003).

PhA, a product of phospholipase D α 1 (PLD α 1) activity, is a positive regulator in ABA-induced ROS and NO production that promotes stomatal closure (Zhang et al., 2009). In *A. thaliana* guard cells, NO synthesis is positively regulated by both ABA and ROS, and interaction of PhA with the two NADPH oxidases, AtrbohD and AtrbohF (Kwak et al., 2003), is necessary for ABA-induced ROS production (Zhang et al., 2009). The NADPH oxidase-deficient double mutant *atrbohD/F* shows impaired ABA induction of NO production and stomatal closure, indicating that ROS production is necessary for NO production. Application of NO scavengers can inhibit ROS-mediated stomatal closure, indicating that NO is required for ROS-promoted stomatal closure. In contrast, application of NO cannot induce ROS production in *A. thaliana* guard cells (Bright et al., 2006). These findings indicate that PhA functions upstream of ROS production and ROS function upstream of NO production.

The lipid metabolite sphingosine-1-phosphate (S1P) is a product of sphingosine kinase (SPHK) activity, which uses the long-chain amine alcohol sphingosine as a substrate. S1P induced increases of cytosolic (Ca²⁺) (Ng et al., 2001) and stimulated stomatal closure in *C. communis* and *A. thaliana* (Ng et al., 2001, Coursol et al., 2003). The *A. thaliana* genome encodes two functional SPHK genes, *SPHK1* and *SPHK2* (Worrall et al., 2008; Guo et al., 2011). Both SPHKs can use sphingosine and phyto-sphingosine as substrates to produce S1P and phyto-S1P, respectively. Both S1P and phyto-S1P induce stomatal closure in *A. thaliana* (Coursol et al., 2005). S1P inhibits inward K⁺ channels and promotes slow anion channel activity in *A. thaliana* guard cell protoplasts, which in turn cause inhibition of stomatal opening and promotion of stomatal closure, respectively (Coursol et al., 2003). In *A. thaliana*, a functional G-protein α -subunit (GPA1) is required for S1P regulation of ion channels (Coursol et al., 2003). In *A. thaliana*, PhA interacts with SPHKs, promoting substrate binding, which in turn increases SPHK activity. Phyto-S1P induces PhA production in wild type (WT) *A. thaliana*, but not in the *pld α* mutant, indicating a positive regulatory role of phyto-S1P in PLD α -mediated PhA production. It has been suggested that phyto-S1P promotes PLD α activity by increasing cytoplasmic Ca²⁺ concentration (Guo and Wang, 2012). These findings indicate that phyto-S1P and PhA are dependent on each other via positive feedback regulation.

A guard cell-specific and ABA-independent oxylipin pathway was recently reported (Montillet and Hirt, 2013). Derived from complex membrane lipids, unesterified fatty acids are catalyzed by lipoxygenase (LOX) into various oxylipin products, such as JA, fatty acid hydroperoxides, and reactive electrophile species (RES) oxylipins, and these can induce stomatal closure at nanomolar concentrations (Montillet et al., 2013). *A. thaliana lox1* mutants were as sensitive to exogenously applied ABA as WT plants, suggesting that LOX1 activity is not involved in ABA-induced stomatal closure. In addition, a transgenic SA-deficient NahG line, and the two SA biosynthesis mutant lines, *sid1-1* and *sid2-1*, responded normally to ABA, but were non-responsive to RES oxylipins. In addition, *lox1* mutant lines were as sensitive to SA (100 μ M) as WT, demonstrating that exogenously applied SA compensated for the LOX1 deficiency. The results indicate that SA is required to convey the RES oxylipin signal, but not the ABA-mediated signal, leading to stomatal closure.

Naturally occurring saturated short, straight chain fatty acids, such as decanoic and undecanoic acids, can inhibit stomatal opening and cause stomatal closure in epidermal strips of *C. communis* (Willmer et al., 1978). In contrast, some polyunsaturated fatty acids, such as linolenic and arachidonic acid enhance stomatal opening and inhibit stomatal closing, consistent with their promotion of inward K⁺ channel activity and inhibition of outward K⁺ channel activity (Lee et al., 1994). Very-long-chain polyunsaturated fatty acids (VLCPUFAs), such as eicosapentaenoic acid (20:5 $\delta^{5,8,11,14,17}$) are abundant lipids in several key plant pathogens (Sun et al., 2013), and may elicit plant defense responses, including stomatal closure. Interestingly, it was shown that exogenous application of eicosadienoic and eicosatrienoic acids to WT plants or endogenous production in the transgenic plants could reduce water loss from excised leaves and confer ABA hypersensitivity to stomatal responses (Yuan et al., 2014). Some fatty acids have been shown to regulate stomatal development, thus affecting the overall plant response to the environment. The *A. thaliana* gene *HIC* (high carbon dioxide) encodes a putative 3-keto acyl coenzyme A synthase (KCS), an enzyme involved in the synthesis of very-long-chain fatty acids (VLCFA) and is a negative regulator of stomatal development in response to CO₂ (Gray et al., 2000). Mutant *hic* plants exhibit up to a 42% increase in stomatal density in response to a doubling of CO₂, possibly by preventing the synthesis of component(s) of the extracellular matrix found at the guard cell surface, such as waxes, glycerolipids, sphingolipids, and cutin (Gray et al., 2000). FATTY-ACID DESATURASE4 (FAD4) is required to desaturate palmitic acid (16:0), and the *fad4* mutant is unable to change stomatal index (defined as the percentage of stomata as compared to all the epidermal cells (including stomata) in a unit area of leaf) in response to elevated CO₂ (Lake et al., 2002). Metabolic profiling of *sdd1* (*STOMATAL DENSITY AND DISTRIBUTION1*) plants, which have three to fourfold higher stomatal density than WT plants, showed a fivefold reduction of unsaturated C16 fatty acids compared to WT, and a concomitant rise in saturated fatty acid 16:0 species (i.e., palmitic acid; Fiehn et al., 2000). The fates of these fatty acids are scarcely known, although it is assumed that some are incorporated into the cutin layers. In *A. thaliana*, mutations of the VLCFA-producing enzymes CER6, CER1, and HIC that are involved in cuticle

biosynthesis result in increased stomatal index (Gray et al., 2000). Whether stomatal index/density affects stomatal movement is not clear. Nevertheless, the aforementioned roles of fatty acid metabolites and their metabolic enzymes offer new avenues to elucidate lipid signaling networks in guard cells, which will facilitate engineering of fatty acid metabolism in crops for enhanced stress tolerance and productivity.

Phytohormone Cross-talk in Stomatal Function

The phytohormone ABA, first reported in plants in the 1960s (Eagles and Wareing, 1963; Ohkuma et al., 1963), is the single most studied metabolite in guard cell physiology owing to its distinct stress (e.g., drought) responsiveness and strong effect on stomatal closure. ABA causes stomatal closure, prevents opening of closed stomata, and reduces transpiration in the leaves of a wide range of species. Stomata accumulate (Cornish and Zeevaart, 1986), catabolize (Grantz et al., 1985), and conjugate exogenously supplied ABA (Grantz et al., 1985; Lee et al., 2006), but to date it is unclear if stomatal opening initially includes or requires depletion of endogenous guard cell ABA (Tallman, 2004). The biosynthesis of ABA from carotenoids in plastids and its catabolism and storage in the cytosol and endoplasmic reticulum in plant cells is well characterized (Nambara and Marion-Poll, 2005). The regulatory network of ABA sensing involve three major components, PYRABACTIN RESISTANCE1 (PYR1)/PYR1-LIKE (PYL)/REGULATORY COMPONENTS OF ABA RECEPTORS (RCAR; i.e., PYR/PYL/RCAR; an ABA receptor; Ma et al., 2009; Park et al., 2009; Joshi-Saha et al., 2011), type 2C protein phosphatase (PP2C; a negative regulator) and SNF1-related protein kinase 2 (SnRK2; a positive regulator), and they offer a double negative regulatory system, (PYR/PYL/RCAR—|PP2C—|SnRK2), which has been well studied (Klingler et al., 2010; Umezawa et al., 2010). PP2Cs inactivate SnRK2s kinases by physical interaction and direct dephosphorylation. Upon ABA binding, PYLs change their conformations and then physically interact and inhibit PP2Cs. However, PYLs inhibit PP2Cs in both the presence and absence of ABA and activate SnRK2s (Zhang et al., 2015). Several natural and artificial compounds interacting with the ABA receptor PYR/PYL/RCAR family are now known (Hitomi et al., 2013). Evolutionary insights obtained from studies on components of the ABA signaling network indicate that PYR/RCAR ABA receptor and ABF-type (ABA-responsive element binding factors) transcription factor families arose during land colonization by plants, while the ABA biosynthesis enzymes have evolved in different plant and fungal specific pathways (Hauser et al., 2011). The structural insights provided from the three-dimensional structures of module PYR/PYL/RCAR-ABA-PP2C pave the way to the design of ABA agonists able to modulate the plant stress response (Santiago et al., 2012).

ABA is transported over short and long distances in plants. Plasma membrane-localized ABA transporters belonging to ATP-binding cassette (ABC; Kang et al., 2010) and nitrate transporter 1/peptide transporter (NRT1/PTR) families are established (Boursiac et al., 2013) and ABA-perception sites were visualized on the plasma membrane of stomatal guard cells (Yamazaki et al.,

2003), in addition to internal sites of perception. For instance, application of ABA into the cytosol of *V. faba* guard-cell protoplasts via patch-clamp techniques inhibited inward K^+ currents thus inhibiting stomatal opening (Schwartz et al., 1994). Although ABA synthesis in guard cells and vascular tissues has been shown (Seo and Koshiba, 2011; Bauer et al., 2013; Boursiac et al., 2013), the relative extent to which guard cells and vascular tissues contribute to the ABA dynamics in guard cells is a topic of ongoing interest. For instance, the recent design, engineering and use of ABAleons with ABA affinities in the range of 100–600 nM to map ABA concentration changes in plant tissues with spatial and temporal resolution in distinct cell types, and in response to low humidity and NaCl in guard cells (Waadt et al., 2014) has promising future applications.

ABA causes alkalization of the guard cell cytosol (Blatt and Armstrong, 1993), which directly enhances outward K^+ channel activity (Blatt and Armstrong, 1993; Ilan et al., 1994; Miedema and Assmann, 1996), and a sustained efflux of both anions and K^+ from guard cells contributes to loss of guard cell turgor, thus facilitating stomatal closing. In addition, ABA-induced stomatal closing can be Ca^{2+} -dependent or -independent (Schroeder et al., 2001). ABA mediated inhibition of stomatal opening is a process distinct from ABA-induced stomatal closure, and it is unclear if H_2O_2 and NO are involved in the ABA inhibition of stomatal opening (Desikan et al., 2004). Even after half a century of research, the role of ABA in guard cell signaling continues to be elucidated (Kim et al., 2010; Yu and Assmann, 2014). ABA content can be decreased via catabolism to phaseic acid (PA), sequestration in the form of an ABA-glucose ester (ABA-GE), which is thought to be physiologically inactive, or deposition in vacuoles (Nambara and Marion-Poll, 2005). Studies on sugar-response mutants indicate that ABA and sugar-response pathways overlap extensively (León and Sheen, 2003). It is known that the sugar sensing effects mediated by HXK are dependent on production of and signaling by ABA (Rolland et al., 2006; Rognoni et al., 2007; Ramon et al., 2008); for example, these interactions take place in mesophyll cells where sugar and HXK inhibit expression of photosynthesis genes (Rolland et al., 2006). Recently, it has been shown that sugar and HXK stimulate the ABA signaling pathway within guard cells, promoting stomatal closure (Kelly et al., 2013). These effects were also observed in epidermal peels, suggesting that sugar and HXK stimulate production of ABA, release of biologically active ABA from inactive ABA pools, and/or inhibition of ABA degradation within guard cells (Koiwai et al., 2004; Christmann et al., 2005; Melhorn et al., 2008; Wasilewska et al., 2008; Zhu et al., 2011). These observations also imply that ABA is probably essential for daily regulation of stomatal aperture even in the absence of water stress (Kelly et al., 2013).

A comprehensive and comparative metabolomics study undertaken in guard and mesophyll cells of *A. thaliana* revealed that following ABA treatment, metabolites are clustered into different temporal modules in guard cells and mesophyll cells (Jin et al., 2013). Guard cell modules differ in WT plants as compared to the modules in the heterotrimeric G-protein α subunit null mutant (*gpa1*), with fewer metabolites showing ABA-altered profiles in *gpa1*, consistent with hyposensitivity of *gpa1* K^+ , anion, and Ca^{2+} channels to ABA (Wang et al., 2001; Fan et al., 2008;

Zhang et al., 2011). For instance, the Ca^{2+} -mobilizing metabolites SIP and cyclic adenosine 5'-diphosphoribose (cADPR) exhibited weaker ABA-stimulated increases in *gpa1* than in WT guard cells. Phytohormones such as ABA catabolites, i.e., ABA glucose-ester, PA, and dihydrophaseic acid (DiHPA), and indole-3-acetic acid (IAA), JA, MeJA, and methyl salicylate were responsive to ABA, with greater responsiveness in WT than in the *gpa1* guard cells. In particular, IAA concentrations in guard cells declined following ABA treatment in WT guard cells but not in *gpa1* guard cells. These findings are consistent with the observation that exogenous application of IAA activates the guard cell H^+ -ATPase and impairs ABA-inhibition of stomatal opening, and suggest that endogenous ABA in guard cells functions upstream to regulate other endogenous hormones, particularly IAA, consistent with G proteins modulating multiple hormonal signaling pathways. Most phytohormones also showed differential ABA responses in guard cells as compared to mesophyll cells (Jin et al., 2013). In support of the idea that multiple hormones regulate guard cell responses, in *V. faba*, cytokinin and auxin induced stomatal opening (Levitt et al., 1987; Song et al., 2006) in conjunction with decreased H_2O_2 production (Song et al., 2006). Salicylic acid (SA) is a ubiquitous phenolic phytohormone involved in stomatal movement. Addition of 1 mM SA to fully opened stomata resulted in a significant reduction (75%) in stomatal aperture (Lee, 1998) in *C. communis*. SA is known to induce stomatal closure accompanied by extracellular ROS production, intracellular ROS accumulation and inward K^+ channel inactivation (Khokon et al., 2011a). Although both ABA and SA were reported to be needed for stomatal closure in response to pathogens, with SA action upstream of ABA (Melotto et al., 2006), a recent study using the ABA biosynthesis mutant *aba2* and a mutant of JA biosynthesis reported no differences in SA induced stomatal closure in the mutants as compared to WT. The authors concluded that neither ABA nor JA is involved in SA, yeast elicitor, or chitosan-induced stomatal closure in *A. thaliana* (Issak et al., 2013). These results appear to indicate the presence of an ABA independent SA signaling pathway in guard cells, but more research is needed to fully resolve the contradictory conclusions in the literature.

In *V. faba* (Zhang et al., 2001) and *Pisum sativum* (Suhita et al., 2004), ABA-mediated stomatal closure is preceded by cytoplasmic alkalization and H_2O_2 production, events that also occur during MeJA-mediated stomatal closure. In fact, ABA and MeJA-mediated stomatal closure share several characteristic signaling components, such as Ca^{2+} involvement, protein phosphorylation, cytoplasmic alkalization, ROS production, and modulation of plasma membrane K^+ channels in the guard cells (Suhita et al., 2004). Extremely low levels of the phytotoxin coronatine (COR), secreted by virulent strains of *Pseudomonas syringae* p.v. *tomato* (*Pst*) act as a JA mimic, activate the JA signaling pathway, and enable the strains to reopen stomata, thereby circumventing host stomatal defense (Montillet and Hirt, 2013). However, unlike COR, exogenous MeJA does not appear to antagonize ABA-induced stomatal closure (Melotto et al., 2006). In fact, the ability of MeJA to regulate stomatal apertures remains controversial (Montillet et al., 2013). Allene oxide synthase (AOS) is a key enzyme in the oxylipin pathway and plays a vital role in production of 12-oxo-phytodienoic acid (12-OPDA, a JA precursor) and

JA. Recently, it has been proposed that 12-OPDA, rather than MeJA, acts in promotion of stomatal closure (Savchenko et al., 2014).

The role of brassinosteroids (BRs) in stomatal movements is less established. Brassinolide (BL), the most bioactive BR form, has been shown to promote stomatal closure in *V. faba* (Haubrick et al., 2006), where BL-induced stomatal opening was not observed. Interestingly, low concentrations of epibrassinolide (eBL) promoted stomatal opening in epidermal peels of *Solanum lycopersicum* in the dark, whereas high concentrations of eBL promoted stomatal closure in the light (Xia et al., 2014). Exogenous (apoplastic) and endogenous (cytosolic) BR may act differently, and guard cells of different species may respond differently to BL application. In *S. lycopersicum*, transient H_2O_2 production was deemed essential for poising the cellular redox status, which played an important role in BR-induced stomatal opening (Xia et al., 2014). BR promoted stomatal closure through apparent biosynthesis of ABA, while stomatal opening was dependent on the GSH redox status of the guard cells. It was proposed that GSH regeneration and/or biosynthesis, leading to a reduced redox status, strictly controls the ROS level and negatively regulates the ABA response pathway, and that BR can directly induce ROS production independently of ABA via NADPH oxidase.

ET and its precursor 1-aminocyclopropane-1-carboxylic acid (ACC) activate the production of H_2O_2 in guard cells and induce stomatal closure in *V. faba* (Song et al., 2014), and the closure was preceded by elevated ROS generated by NADPH oxidases (Desikan et al., 2006). However, the ET effect varies depending on species and conditions. For example, ET promotes stomatal closure in *Arachis hypogaea* (Pallas and Kays, 1982) and *A. thaliana* (using intact leaves; Desikan et al., 2006), but evokes stomatal opening in *Dianthus caryophyllus* and *S. lycopersicum* (Madhavan et al., 1983), *V. faba* (Levitt et al., 1987), and *A. thaliana* (using epidermal peels; Tanaka et al., 2005). The ET effect on stomatal opening was attributed to its impairment of ABA regulation of stomatal closure (Tanaka et al., 2005), but recently, it was shown in *A. thaliana* that ET mediated BR-induced stomatal closure via $\text{G}\alpha$ protein-activated AtrbohF-dependent H_2O_2 production and subsequent Nia1-catalyzed NO production (Ge et al., 2015; Shi et al., 2015). Nonetheless, the exact mechanisms underlying the different ET effects are unknown.

Nitrogen and Sulfur Rich Metabolites in Guard Cell Signaling

Nitrogenous bases in the form of purines and pyrimidines form an essential pool of nitrogen in plant cells. Nitrogenous metabolites have been extensively studied as metabolic intermediates and signaling molecules in stomatal movement and guard cell function. Important nitrogenous signaling molecules, such as cADPR, a metabolite derived from nicotinamide adenine dinucleotide (NAD; Wu et al., 1997), play important roles in guard cell ABA signaling. Injection of cADPR into guard cells resulted in $[(\text{Ca}^{2+})_{\text{cyt}}]$ increases and turgor reduction. When guard cells were preloaded with the cADPR antagonist 8NH₂-cADPR, a slowing of stomatal closure was observed in response to ABA (Leckie et al., 1998). Recently, it was established that inhibition of the poly (ADP-R)

polymerase activity correlated with increased number of stomata in *A. thaliana* leaves (Schulz et al., 2014), highlighting the role of poly (ADP-R) metabolism in stomatal development. Another nucleotide-related metabolite, cyclic guanosine monophosphate (cGMP), has been implicated in ABA-induced stomatal closure by acting downstream of H₂O₂ and NO in the signaling pathway by which ABA induces stomatal closure (Dubovskaya et al., 2011). H₂O₂ and NO-induced cytosolic calcium increases [(Ca²⁺)_{cyt}] were cGMP-dependent, positioning cGMP upstream of (Ca²⁺)_{cyt}, and involved the action of the type 2C protein phosphatase, ABI1. Increases in cGMP were mediated through the stimulation of guanylyl cyclase by H₂O₂ and NO (Dubovskaya et al., 2011). The nitrated form of cGMP (8-nitro-cGMP) is a positive regulator in promotion of stomatal closure (Joudoi et al., 2013). NO and cGMP induce the synthesis of 8-nitro-cGMP in guard cells in the presence of ROS leading to the hypothesis that NO-dependent guanine nitration of cGMP may occur in plants and the resulting 8-nitro-cGMP acts as a signaling molecule that activates cADPR production in guard cells. By contrast, a positive role for cGMP in kinetin- and natriuretic peptide-induced stomatal opening in *Tradescantia albiflora* (Pharmawati et al., 1998) and in auxin-induced stomatal opening in *C. communis* and *A. thaliana* (Cousson and Vavasseur, 1998; Cousson, 2003) also has been recognized. Furthermore, application of 8-bromo-cGMP, a membrane-permeant cGMP analog, causes stomata to open in the dark (Cousson, 2003; Joudoi et al., 2013), but 8-nitro-cGMP does not. These results lead to the conclusion that cGMP and its nitrated derivative play different roles in guard cell signaling, wherein cGMP promotes stomatal opening in the dark, while 8-nitro-cGMP promotes stomatal closure in the light (Joudoi et al., 2013).

Another important group of N-containing specialized metabolites in plants are polyamines. Exogenous application of polyamines, such as 1 mM spermine, inhibit stomatal opening by inhibiting inwardly rectifying K⁺ channels (Liu et al., 2000). Application of spermidine also promotes stomatal closure but the mechanism is unknown as outward K⁺ channels and anion channels are not affected (Liu et al., 2000). On the other hand, (acetyl)-1,3-diaminopropane (DAP), a product of oxidative deamination of spermidine and spermine, suppresses anionic currents, and increases those of inwardly rectifying K⁺ channels, and may induce membrane hyperpolarization and extracellular acidification by activating the H⁺ ATPase, thus restraining stomatal closing (Jammes et al., 2014). These mechanisms act antagonistically to ABA. It is thought that during acclimation to low soil-water availability, acetyl-DAP prevents complete stomatal closure (Jammes et al., 2014). Moreover, DAP and such amine oxidase reaction products are precursors of γ -amino butyric acid, alkaloids, β -alanine, and other uncommon polyamines that play significant roles in stress tolerance and defense (Bouchereau et al., 1999). In fact, based on proteome analysis in *Brassica napus* guard cells, ABA-responsive proteins that decrease in abundance include those involved in spermidine synthesis, purine metabolism, and alkaloid biosynthesis pathways (Zhu et al., 2010).

Glucosinolates are N- and S-containing specialized metabolites in plants that have been shown to be present in guard

cells. Glucosinolate-myrosinase systems in Brassicales, especially *A. thaliana*, are well understood in plant-herbivore interactions and defense against pathogens (Yan and Chen, 2007; Andersson et al., 2015). However, recent evidence from proteomic investigations has indicated that glucosinolates are required for ABA responses of guard cells (Zhao et al., 2008, Zhu et al., 2010, 2014). THIOGLUCOSIDE GLUCOHYDROLASE1 (TGG1), encoding a myrosinase that catalyzes the production of isothiocyanates (ITC) from glucosinolates, is highly abundant in guard cell proteomes. In fact, myrosinases are proposed to redundantly function downstream of ROS production and upstream of cytosolic Ca²⁺ elevation in ABA and MeJA signaling in guard cells (Islam et al., 2009). *A. thaliana* *tgg1* mutants are hyposensitive to ABA inhibition of guard cell inward K⁺ channels and stomatal opening. In addition, thiol-reagents such as ITCs have been shown to be potent inducers of stomatal closure, possibly via covalent reactions with RES oxylipin targets (Montillet et al., 2013). Some of the glucosinolate-producing plant species, such as *Brassica juncea*, produce 2-propenylglucosinolate, which can be hydrolyzed to allylisothiocyanate (allylITC). Exogenous application of allylITC was found to induce stomatal closure (Khokon et al., 2011b). The stomatal closure by allylITC was induced via production of ROS and NO, and elevation of cytosolic Ca²⁺. In addition, other ITCs, nitriles, and thiocyanates (e.g., 3-butenenitrile and ethyl thiocyanate) have also been shown to induce foliar ROS generation and stomatal closure (Hossain et al., 2013). Manipulation of glucosinolate metabolic pathways by plant metabolic engineering and breeding approaches may lead to development of crop varieties with combined disease and drought resistance. Recently, another sulfur containing compound, hydrogen sulfide (H₂S), generated by L-cysteine desulfhydrase was shown to act upstream of NO to modulate ABA-dependent stomatal closure (Scuffi et al., 2014).

Conclusion

The plant leaf metabolome can boast as many as 5,000 different metabolites (Bino et al., 2004). Considering the roles of established metabolites in guard cell functions, we have begun the heydays of functional genomics, fluxomics, and systems biology toward understanding of this highly sophisticated single cell type model system. Although studies on guard cell metabolism are highly biased toward ABA and osmolytes owing to their primary importance in stomatal movement, the identification of additional critical metabolites (as shown in **Figure 1**) underlying or correlated with stomatal movement will form a solid foundation toward a broader understanding of optimal plant adaptation to environmental changes. For example, although progress in the study of stomatal movement in plant immunity has been made (Zhang et al., 2008), a deeper mechanistic understanding is required to harness the potential for generation of disease resistant crops. Information currently available has revealed universal and diverse metabolites and pathways leading to stomatal responses.

Many years of traditional breeding has unknowingly selected varieties with cool leaf temperature in some species, i.e., larger stomatal opening for higher yield. For instance, in Pima cotton

(*Gossypium barbadense* L.) and bread wheat, increased stomatal conductance led to lint and grain yield increases respectively (Lu et al., 1998). Furthermore, in cotton, stomatal conductance and leaf cooling were significantly correlated with fruiting prolificacy and yield during the hottest period of the year (Radin et al., 1994). Understanding the functions and molecular networks of the regulatory metabolites of stomatal functions would open avenues for development of “smart” crops, providing a unique platform for endeavors at the genetic level to favor food security and human nutrition. Although we did not focus on the roles of stomatal ontogeny, shape, size, and distribution, which can also significantly affect plant water balance, growth and biomass, the engineering of stomatal development and response as a means to improve water use efficiency is an attractive approach

to improve drought tolerance in crops (Schroeder et al., 2001). Guard cell metabolomics and systems biology hold the potential to unravel key molecular networks that control plant productivity and defense in a changing climate.

Acknowledgments

Work on guard cell signaling in the Assmann laboratory is supported by BARD grant IS-4541-12 and by NSF grants MCB-1121612, MCB-1157921, and MCB-1412644 to SA. Research on guard cell signaling in the Chen laboratory is supported by NSF grants MCB-0818051, MCB-1158000, and MCB-1412547 to SC. Work on guard cell signaling in the Granot laboratory is supported by BARD grant IS-4541-12.

References

- Akter, N., Okuma, E., Sobahan, M. A., Uraji, M., Munemasa, S., Nakamura, Y., et al. (2013). Negative regulation of methyl jasmonate-induced stomatal closure by glutathione in *Arabidopsis*. *J. Plant Growth Regul.* 32, 208–215. doi: 10.1007/s00344-012-9291-7
- Allaway, W. G. (1973). Accumulation of malate in guard cells of *Vicia faba* during stomatal opening. *Planta* 110, 63–70. doi: 10.1007/BF00386923
- Amodeo, G., Talbott, L. D., and Zeiger, E. (1996). Use of potassium and sucrose by onion guard cells during a daily cycle of osmoregulation. *Plant Cell Physiol.* 37, 575–579. doi: 10.1093/oxfordjournals.pcp.a028983
- Andersson, M. X., Nilsson, A. K., Johansson, O. N., Boztaş, G., Adolfsson, L. E., Pinosa, F., et al. (2015). Involvement of the electrophilic isothiocyanate sulforaphane in *Arabidopsis* local defense responses. *Plant Physiol.* 167, 251–261. doi: 10.1104/pp.114.251892
- Araujo, W. L., Nunes-Nesi, A., Osorio, S., Usadel, B., Fuentes, D., Nagy, R., et al. (2011). Antisense inhibition of the iron-sulphur subunit of succinate dehydrogenase enhances photosynthesis and growth in tomato via an organic acid-mediated effect on stomatal aperture. *Plant Cell* 23, 600–627. doi: 10.1105/tpc.110.081224
- Asai, N., Nakajima, N., Tamaoki, M., Kamada, H., and Kondo, N. (2000). Role of malate synthesis mediated by phosphoenolpyruvate carboxylase in guard cells in the regulation of stomatal movement. *Plant Cell Physiol.* 41, 10–15. doi: 10.1093/pcp/41.1.10
- Bates, G. W., Rosenthal, D. M., Sun, J., Chattopadhyay, M., Peffer, E., Yang, J., et al. (2012). A comparative study of the *Arabidopsis thaliana* guard-cell transcriptome and its modulation by sucrose. *PLoS ONE* 7:e49641. doi: 10.1371/journal.pone.0049641
- Bauer, H., Ache, P., Lautner, S., Fromm, J., Hartung, W., Al-Rasheid, K. A., et al. (2013). The stomatal response to reduced relative humidity requires guard cell-autonomous ABA synthesis. *Curr. Biol.* 23, 53–57. doi: 10.1016/j.cub.2012.11.022
- Bino, R. J., Hall, R. D., Fiehn, O., Kopka, J., Saito, K., Draper, J., et al. (2004). Potential of metabolomics as a functional genomics tool. *Trends Plant Sci.* 9, 418–425. doi: 10.1016/j.tplants.2004.07.004
- Blatt, M. R., and Armstrong, F. (1993). K⁺ channels of stomatal guard cells: abscisic-acid-evoked control of the outward-rectifier mediated by cytoplasmic pH. *Planta* 191, 330–341. doi: 10.1007/BF00195690
- Bouchereau, A., Aziz, A., Larher, F., and Martin-Tanguy, J. (1999). Polyamines and environmental challenges: recent development. *Plant Sci.* 140, 103–125. doi: 10.1016/S0168-9452(98)00218-0
- Boursiac, Y., Lérans, S., Corratgé-Faillie, C., Gojon, A., Krouk, G., and Lacombe, B. (2013). ABA transport and transporters. *Trends Plant Sci.* 18, 325–333. doi: 10.1016/j.tplants.2013.01.007
- Bright, J., Desikan, R., Hancock, J. T., Weir, I. S., and Neill, S. J. (2006). ABA-induced NO generation and stomatal closure in *Arabidopsis* are dependent on H₂O₂ synthesis. *Plant J.* 45, 113–122. doi: 10.1111/j.1365-313X.2005.02615.x
- Chen, Z., and Gallie, D. R. (2004). The ascorbic acid redox state controls guard cell signaling and stomatal movement. *Plant Cell* 16, 1143–1162. doi: 10.1105/tpc.021584
- Christmann, A., Hoffmann, T., Teplova, I., Grill, E., and Muller, A. (2005). Generation of active pools of abscisic acid revealed by *in vivo* imaging of water-stressed *Arabidopsis*. *Plant Physiol.* 137, 209–219. doi: 10.1104/pp.104.053082
- Cornish, K., and Zeevaert, J. A. (1986). Abscisic acid accumulation by *in situ* and isolated guard cells of *Pisum sativum* L. and *Vicia faba* L. in relation to water stress. *Plant Physiol.* 81, 1017–1021. doi: 10.1104/pp.81.4.1017
- Coursol, S., Fan, L. M., Le Stunff, H., Spiegel, S., Gilroy, S., and Assmann, S. M. (2003). Sphingolipid signaling in *Arabidopsis* guard cells involves heterotrimeric G proteins. *Nature* 423, 651–654. doi: 10.1038/nature01643
- Coursol, S., Le Stunff, H., Lynch, D. V., Gilroy, S., Assmann, S. M., and Spiegel, S. (2005). *Arabidopsis* sphingosine kinase and the effects of phytosphingosine-1-phosphate on stomatal aperture. *Plant Physiol.* 137, 724–737. doi: 10.1104/pp.104.055806
- Cousson, A. (2003). Pharmacological evidence for a positive influence of the cyclic GMP-independent transduction on the cyclic GMP-mediated Ca²⁺-dependent pathway within the *Arabidopsis* stomatal opening in response to auxin. *Plant Sci.* 164, 759–767. doi: 10.1016/S0168-9452(03)00062-1
- Cousson, A., and Vavasseur, A. (1998). Putative involvement of cytosolic Ca²⁺ and GTP-binding proteins in cyclic-GMP-mediated induction of stomatal opening by auxin in *Commelina communis* L. *Planta* 206, 308–314. doi: 10.1007/s004250050405
- Dennis, D. T., and Blakeley, S. D. (2000). “Carbohydrate metabolism,” in *Biochemistry and Molecular Biology of Plants*, eds B. B. Buchanan, W. Gruissem, and R. L. Jones (Rockville, MD: American Society of Plant Physiologists), 676–728.
- Desikan, R., Cheung, M. K., Bright, J., Henson, D., Hancock, J. T., and Neill, S. J. (2004). ABA, hydrogen peroxide and nitric oxide signalling in stomatal guard cells. *J. Exp. Bot.* 55, 205–212. doi: 10.1093/jxb/erh033
- Desikan, R., Last, K., Harrett-Williams, R., Tagliavia, C., Harter, K., Hooley, R., et al. (2006). Ethylene-induced stomatal closure in *Arabidopsis* occurs via AtrbohF-mediated hydrogen peroxide synthesis. *Plant J.* 47, 907–916. doi: 10.1111/j.1365-313X.2006.02842.x
- Dittrich, P., and K. Raschke, K. (1977). Malate metabolism in isolated epidermis of *Commelina communis* L. in relation to stomatal functioning. *Planta* 134, 77–81. doi: 10.1007/BF00390098
- Dixon, R. A., and Paiva, N. L. (1995). Stress-induced phenylpropanoid metabolism. *Plant Cell* 7, 1085–1097. doi: 10.1105/tpc.7.7.1085
- Dubovskaya, L. V., Bakakina, Y. S., Kolesneva, E. V., Sodel, D. L., McAinsh, M. R., and Hetherington, A. M. (2011). cGMP-dependent ABA-induced stomatal closure in the ABA-insensitive *Arabidopsis* mutant *abi1-1*. *New Phytol.* 191, 57–69. doi: 10.1111/j.1469-8137.2011.03661.x
- Eagles, C. F., and Wareing, P. F. (1963). Dormancy regulators in woody plants: experimental induction of dormancy in *Betula pubescens*. *Nature* 199, 874–875. doi: 10.1038/199874a0
- Fan, L. M., Zhang, W., Chen, J. G., Taylor, J. P., Jones, A. M., and Assmann, S. M. (2008). Abscisic acid regulation of guard-cell K⁺ and anion channels in Gβ- and RGS-deficient *Arabidopsis* lines. *Proc. Natl. Acad. Sci. U.S.A.* 105, 8476–8481. doi: 10.1073/pnas.0800980105
- Fernie, A. R., and Martinoia, E. (2009). Malate. Jack of all trades or master of a few? *Phytochemistry* 70, 828–832. doi: 10.1016/j.phytochem.2009.04.023

- Fiehn, O., Kopka, J., Dörmann, P., Altmann, T., Trethewey, R. N., and Willmitzer, L. (2000). Metabolite profiling for plant functional genomics. *Nat. Biotechnol.* 18, 1157–1161. doi: 10.1038/81137
- Fischer, R. A. (1968). Stomatal opening: role of potassium uptake by guard cells. *Science* 160, 784–785. doi: 10.1126/science.160.3829.784
- Ge, X. M., Cai, H. L., Lei, X., Zhou, X., Yue, M., and He, J. M. (2015). Heterotrimeric G protein mediates ethylene-induced stomatal closure via hydrogen peroxide synthesis in *Arabidopsis*. *Plant J.* 82, 138–150. doi: 10.1111/tj.12799
- Gilroy, S., Read, N., and Trewavas, A. J. (1990). Elevation of cytoplasmic calcium by caged calcium or caged inositol trisphosphate initiates stomatal closure. *Nature* 346, 769–771. doi: 10.1038/346769a0
- Gotow, K., Taylor, S., and Zeiger, E. (1988). Photosynthetic carbon fixation in guard cell protoplasts of *Vicia faba* L.: evidence from radiolabel experiments. *Plant Physiol.* 86, 700–705. doi: 10.1104/pp.86.3.700
- Granot, D. (2008). Putting plant hexokinases in their proper place. *Phytochemistry* 69, 2649–2654. doi: 10.1016/j.phytochem.2008.08.026
- Grantz, D. A., Ho, T. H. D., Uknes, S. J., Cheeseman, J. M., and Boyer, J. S. (1985). Metabolism of abscisic acid in guard cells of *Vicia faba* L. and *Commelina communis* L. *Plant Physiol.* 78, 51–56. doi: 10.1104/pp.78.1.51
- Gray, J. E., Holroyd, G. H., van der Lee, F. M., Bahrami, A. R., Sijmons, P. C., Woodward, F. I., et al. (2000). The HIC signaling pathway links CO₂ perception to stomatal development. *Nature* 408, 713–716. doi: 10.1038/35047071
- Guo, L., and Wang, X. (2012). Crosstalk between phospholipase D and sphingosine kinase in plant stress signaling. *Front. Plant Sci.* 3:51. doi: 10.3389/fpls.2012.00051
- Guo, L., Mishra, G., Taylor, K., and Wang, X. (2011). Phosphatidic acid binds and stimulates *Arabidopsis* sphingosine kinases. *J. Biol. Chem.* 286, 13336–13345. doi: 10.1074/jbc.M110.190892
- Haubrick, L. L., Torsethaugen, G., and Assmann, S. M. (2006). Effect of brassinolide, alone and in concert with abscisic acid, on control of stomatal aperture and potassium currents of *Vicia faba* guard cell protoplasts. *Physiol. Plant.* 128, 134–143. doi: 10.1111/j.1399-3054.2006.00708.x
- Hauser, F., Waadt, R., and Schroeder, J. I. (2011). Evolution of abscisic acid synthesis and signaling mechanisms. *Curr. Biol.* 21, R346–R355. doi: 10.1016/j.cub.2011.03.015
- He, J. M., Ma, X. G., Zhang, Y., Sun, T. F., Xu, F. F., Chen, Y. P., et al. (2013). Role and interrelationship of G α protein, hydrogen peroxide, and nitric oxide in ultraviolet B-induced stomatal closure in *Arabidopsis* leaves. *Plant Physiol.* 161, 1570–1583. doi: 10.1104/pp.112.21.1623
- Hedrich, R., and Marten, I. (1993). Malate-induced feedback regulation of plasma membrane anion channels could provide a CO₂ sensor to guard cells. *EMBO J.* 12, 897–901.
- Hitomi, K., Getzoff, E. D., and Schroeder, J. I. (2013). Going green: phytohormone mimetics for drought rescue. *Plant Physiol.* 163, 1087–1088. doi: 10.1104/pp.113.227660
- Hoque, T. S., Uraji, M., Ye, W., Hossain, M. A., Nakamura, Y., and Murata, Y. (2012). Methylglyoxal-induced stomatal closure accompanied by peroxidase-mediated ROS production in *Arabidopsis*. *J. Plant Physiol.* 169, 979–986. doi: 10.1016/j.jplph.2012.02.007
- Hossain, M. S., Ye, W., Hossain, M. A., Okuma, E., Uraji, M., Nakamura, Y., et al. (2013). Glucosinolate degradation products, isothiocyanates, nitriles, and thiocyanates, induce stomatal closure accompanied by peroxidase-mediated reactive oxygen species production in *Arabidopsis thaliana*. *Biosci. Biotechnol. Biochem.* 77, 977–983. doi: 10.1271/bbb.120928
- Humble, G. D., and Raschke, K. (1971). Stomatal opening quantitatively related to potassium transport: evidence from electron probe analysis. *Plant Physiol.* 48, 447–453. doi: 10.1104/pp.48.4.447
- Ilan, N., Schwartz, A., and Moran, N. (1994). External pH effects on the depolarization-activated K channels in guard cell protoplast of *Vicia faba*. *J. Gen. Physiol.* 103, 807–831. doi: 10.1085/jgp.103.5.807
- Imamura, S. (1943). Untersuchungen über den mechanismus der turgorschwundung der spaltöffnungszellen. *Jpn. J. Bot.* 12, 82–88.
- Islam, M. M., Tani, C., Watanabe-Sugimoto, M., Uraji, M., Jahan, M. S., Masuda, C., et al. (2009). Myrosinases, TGG1 and TGG2, redundantly function in ABA and MeJA signaling in *Arabidopsis* guard cells. *Plant Cell Physiol.* 50, 1171–1175. doi: 10.1093/pcp/pcp066
- Issak, M., Okuma, E., Munemasa, S., Nakamura, Y., Mori, I. C., and Murata, Y. (2013). Neither endogenous abscisic acid nor endogenous jasmonate is involved in salicylic acid-, yeast elicitor-, or chitosan-induced stomatal closure in *Arabidopsis thaliana*. *Biosci. Biotechnol. Biochem.* 77, 1111–1113. doi: 10.1271/bbb.120980
- Jammes, F., Leonhardt, N., Tran, D., Bousserouel, H., Véry, A.-A., Renou, J.-P., et al. (2014). Acetylated 1,3-diaminopropane antagonizes abscisic acid-mediated stomatal closing in *Arabidopsis*. *Plant J.* 79, 322–333. doi: 10.1111/tj.12564
- Jin, X., Wang, R. S., Zhu, M., Jeon, B. W., Albert, R., Chen, S., et al. (2013). Abscisic acid-responsive guard cell metabolomes of *Arabidopsis* wild-type and gpa1 G-protein mutants. *Plant Cell* 25, 4789–4811. doi: 10.1105/tpc.113.119800
- Joshi-Saha, A., Valon, C., and Leung, J. (2011). A brand new START: abscisic acid perception and transduction in the guard cell. *Sci. Signal.* 4, re4–re4. doi: 10.1126/scisignal.2002164
- Joudoi, T., Shichiri, Y., Kamizono, N., Akaike, T., Sawa, T., Yoshitake, J., et al. (2013). Nitrated cyclic GMP modulates guard cell signaling in *Arabidopsis*. *Plant Cell* 25, 558–571. doi: 10.1105/tpc.112.105049
- Jung, J. Y., Kim, Y. W., Kwak, J. M., Hwang, J. U., Young, J., Schroeder, J. I., et al. (2002). Phosphatidylinositol 3- and 4-phosphate are required for normal stomatal movements. *Plant Cell* 14, 2399–2412. doi: 10.1105/tpc.004143
- Kang, J., Hwang, J. U., Lee, M., Kim, Y. Y., Assmann, S. M., Martinoia, E., et al. (2010). PDR-type ABC transporter mediates cellular uptake of the phytohormone abscisic acid. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2355–2360. doi: 10.1073/pnas.0909222107
- Karabourniotis, G., Tzobanoglou, D., Nikolopoulos, D., and Liakopoulos, G. (2011). Epicuticular phenolics over guard cells: exploitation for *in situ* stomatal counting by fluorescence microscopy and combined image analysis. *Ann. Bot.* 87, 631–639. doi: 10.1006/anbo.2001.1386
- Kelly, G., Moshelion, M., David-Schwartz, R., Halperin, O., Wallach, R., Attia, Z., et al. (2013). Hexokinase mediates stomatal closure. *Plant J.* 75, 977–988. doi: 10.1111/tj.12258
- Khokon, A. R., Okuma, E., Hossain, M. A., Munemasa, S., Uraji, M., Nakamura, Y., et al. (2011a). Involvement of extracellular oxidative burst in salicylic acid-induced stomatal closure in *Arabidopsis*. *Plant Cell Environ.* 34, 434–443. doi: 10.1111/j.1365-3040.2010.02253.x
- Khokon, M., Jahan, M. S., Rahman, T., Hossain, M. A., Muroyama, D., Minami, I., et al. (2011b). Allyl isothiocyanate (AITC) induces stomatal closure in *Arabidopsis*. *Plant Cell Environ.* 34, 1900–1906. doi: 10.1111/j.1365-3040.2011.02385.x
- Kim, T. H., Böhmer, B., Hu, H., Nishimura, N., and Schroeder, J. I. (2010). Guard cell signal transduction network: advances in understanding abscisic acid, CO₂, and Ca²⁺ signaling. *Annu. Rev. Plant Biol.* 61, 561–591. doi: 10.1146/annurev-arplant-042809-112226
- Klingler, J. P., Batelli, G., and Zhu, J. K. (2010). ABA receptors: the START of a new paradigm in phytohormone signaling. *J. Exp. Bot.* 61, 3199–3210. doi: 10.1093/jxb/erq151
- Koiwai, H., Nakaminami, K., Seo, M., Mitsuhashi, W., Toyomasu, T., and Koshihara, T. (2004). Tissue-specific localization of an abscisic acid biosynthetic enzyme, AAO3, in *Arabidopsis*. *Plant Physiol.* 134, 1697–1707. doi: 10.1104/pp.103.036970
- Kwak, J. M., Mori, I. C., Pei, Z. M., Leonhardt, N., Torres, M. A., Dangl, J. L., et al. (2003). NADPH oxidase AtrbohD and AtrbohF genes function in ROS-dependent ABA signaling in *Arabidopsis*. *EMBO J.* 22, 2623–2633. doi: 10.1093/emboj/cdg277
- Lake, J. A., Woodward, F. I., and Quick, W. P. (2002). Long-distance CO₂ signaling in plants. *J. Exp. Bot.* 53, 183–193. doi: 10.1093/jxb/53.3.183
- Lawson, T., Simkin, A. J., Kelly, G., and Granot, D. (2014). Mesophyll photosynthesis and guard cell metabolism impacts on stomatal behavior. *New Phytol.* 203, 1064–1081. doi: 10.1111/nph.12945
- Leckie, C. P., McAinsh, M. R., Allen, G. J., Sanders, D., and Hetherington, A. M. (1998). Abscisic acid-induced stomatal closure mediated by cyclic ADP-ribose. *Proc. Natl. Acad. Sci. U.S.A.* 95, 15837–15842. doi: 10.1073/pnas.95.26.15837
- Lee, J.-S. (1998). The mechanism of stomatal closing by salicylic acid in *Commelina communis* L. *J. Plant Biol.* 41, 97–102. doi: 10.1007/BF03030395
- Lee, K. H., Piao, H. L., Kim, H. Y., Choi, S. M., Jiang, F., Hartung, W., et al. (2006). Activation of glucosidase via stress-induced polymerization rapidly increases active pools of abscisic acid. *Cell* 126, 1109–1120. doi: 10.1016/j.cell.2006.07.034
- Lee, M., Choi, Y., Burla, B., Kim, Y. Y., Jeon, B., Maeshima, M., et al. (2008). The ABC transporter AtABC14 is a malate importer and modulates stomatal response to CO₂. *Nat. Cell Biol.* 10, 1217–1223. doi: 10.1038/ncb1782
- Lee, Y., Choi, Y. B., Suh, S., Lee, J., Assmann, S. M., Joe, C. O., et al. (1996). Abscisic acid-induced phosphoinositide turnover in guard cell protoplasts of *Vicia faba*. *Plant Physiol.* 110, 987–996.

- Lee, Y., Lee, H. J., Crain, R. C., Lee, A., and Korn, S. J. (1994). Polyunsaturated fatty acid modulates stomatal aperture and two distinct K^+ channel currents in guard cells. *Cell. Signal.* 6, 181–186. doi: 10.1016/0898-6568(94)90075-2
- Lemtiri-Chlieh, F., MacRobbie, E. A., Webb, A. A., Manison, N. F., Brownlee, C., Skepper, J. N., et al. (2003). Inositol hexakisphosphate mobilizes an endomembrane store of calcium in guard cells. *Proc. Natl. Acad. Sci. U.S.A.* 100, 10091–10095. doi: 10.1073/pnas.1133289100
- León, P., and Sheen, J. (2003). Sugar and hormone connections. *Trends Plant Sci.* 8, 110–116. doi: 10.1016/s1360-1385(03)00011-6
- Leonhardt, N., Kwak, J. M., Robert, N., Waner, D., Leonhardt, G., and Schroeder, J. I. (2004). Microarray expression analyses of *Arabidopsis* guard cells and isolation of a recessive abscisic acid hypersensitive protein phosphatase 2C mutant. *Plant Cell* 16, 596–615. doi: 10.1105/tpc.019000
- Levitt, L. K., Stein, D. B., and Rubinstein, B. (1987). Promotion of stomatal opening by indole acetic acid and ethrel in epidermal strips of *Vicia faba* L. *Plant Physiol.* 85, 318–323. doi: 10.1104/pp.85.2.318
- Li, C. L., Wang, M., Ma, X. Y., and Zhang, W. (2014). NRG1A, a putative mitochondrial pyruvate carrier, mediates abscisic acid regulation of guard cell ion channels and drought stress responses in *Arabidopsis*. *Mol. Plant* 7, 1508–1521. doi: 10.1093/mp/ssu061
- Li, J., Ou-Lee, T. M., Raba, R., Amundson, R. G., and Last, R. L. (1993). *Arabidopsis* flavonoid mutants are hypersensitive to UV-B irradiation. *Plant Cell* 5, 171–179. doi: 10.1105/tpc.5.2.171
- Liu, K., Fu, H., Bei, Q., and Luan, S. (2000). Inward potassium channel in guard cells as a target for polyamine regulation of stomatal movements. *Plant Physiol.* 124, 1315–1326. doi: 10.1104/pp.124.3.1315
- Lloyd, F. E. (1908). *The Physiology of Stomata*. Washington: Carnegie Institution for Science 82, 1–142.
- Lu, P., Outlaw, W. H. Jr, Smith, B. G., and Freed, G. A. (1997). A new mechanism for the regulation of stomatal aperture size in intact leaves. Accumulation of mesophyll-derived sucrose in the guard-cell wall of *Vicia faba*. *Plant Physiol.* 114, 109–118.
- Lu, P., Zhang, S. Q., Outlaw, W. H. Jr, and Riddle, K. A. (1995). Sucrose: a solute that accumulates in the guard-cell apoplast and guard-cell symplast of open stomata. *FEBS Lett.* 362, 180–184. doi: 10.1016/0014-5793(95)00239-6
- Lu, Z., Percy, R. G., Qualset, C. O., and Zeiger, E. (1998). Stomatal conductance predicts yields in irrigated Pima cotton and bread wheat grown at high temperatures. *J. Exp. Bot.* 49, 453–460. doi: 10.1093/jxb/49.Special_Issue.453
- Ma, Y., Szostkiewicz, I., Korte, A., Moes, D., Yang, Y., Christmann, A., et al. (2009). Regulators of PP2C phosphatase activity function as abscisic acid sensors. *Science* 324, 1064–1068. doi: 10.1126/science.1172408
- Madhavan, S., Chrominski, A., and Smith, B. N. (1983). Effect of ethylene on stomatal opening in tomato and carnation leaves. *Plant Cell Physiol.* 24, 569–572.
- Medeiros, D. B., Daloso, D. M., Fernie, A. R., Nikoloski, Z., and Araújo, W. L. (2015). Utilizing systems biology to unravel stomatal function and the hierarchies underpinning its control. *Plant Cell Environ.* doi: 10.1111/pce.12517 [Epub ahead of print].
- Melhorn, V., Matsumi, K., Koiwai, H., Ikegami, K., Okamoto, M., and Nambara, E. (2008). Transient expression of *AtNCE3* and *AAO3* genes in guard cells causes stomatal closure in *Vicia faba*. *J. Plant Res.* 121, 125–131. doi: 10.1007/s10265-007-0127-7
- Melotto, M., Underwood, W., Koczan, J., Nomura, K., and He, S. Y. (2006). Plant stomata function in innate immunity against bacterial invasion. *Cell* 126, 969–980. doi: 10.1016/j.cell.2006.06.054
- Miao, Y., Lv, D., Wang, P., Wang, X.-C., Chen, J., Miao, C., et al. (2006). An *Arabidopsis* glutathione peroxidase functions as both a redox transducer and a scavenger in abscisic acid and drought stress responses. *Plant Cell* 18, 2749–2766. doi: 10.1105/tpc.106.044230
- Miedema, H., and Assmann, S. M. (1996). A membrane-delimited effect of internal pH on the K^+ outward rectifier of *Vicia faba* guard cells. *J. Mem. Biol.* 154, 227–237. doi: 10.1007/s002329900147
- Misra, B. B., Assmann, S. M., and Chen, S. (2014). Plant single-cell and single-cell-type metabolomics. *Trends Plant Sci.* 19, 637–646. doi: 10.1016/j.tplants.2014.05.005
- Montillet, J. L., and Hirt, H. (2013). New checkpoints in stomatal defense. *Trends Plant Sci.* 18, 295–297. doi: 10.1016/j.tplants.2013.03.007
- Montillet, J. L., Leonhardt, N., Mondy, S., Tranchimand, S., Rumeau, D., Boudsocq, M., et al. (2013). An abscisic acid-independent oxylipin pathway controls stomatal closure and immune defense in *Arabidopsis*. *PLoS Biol.* 11:e1001513. doi: 10.3410/f.717991704.793474995
- Moore, B., Zhou, L., Rolland, F., Hall, Q., Cheng, W. H., Liu, Y. X., et al. (2003). Role of the *Arabidopsis* glucose sensor HXK1 in nutrient, light, and hormonal signaling. *Science* 300, 332–336. doi: 10.1126/science.1080585
- Munemasa, S., Muroyama, D., Nagahashi, H., Nakamura, Y., Mori, I. C., and Murata, Y. (2013). Regulation of reactive oxygen species-mediated abscisic acid signaling in guard cells and drought tolerance by glutathione. *Front. Plant Sci.* 4:472. doi: 10.3389/fpls.2013.00472
- Nambara, E., and Marion-Poll, A. (2005). Abscisic acid biosynthesis and catabolism. *Annu. Rev. Plant Biol.* 56, 165–185. doi: 10.1146/annurev.arplant.56.032604.144046
- Ng, C. K. Y., Carr, K., McAnish, M. R., Powell, B., and Hetherington, A. M. (2001). Drought-induced guard cell signal transduction involves sphingosine-1-phosphate. *Nature* 410, 596–599. doi: 10.1038/35069092
- Noctor, G., and Foyer, C. H. (1998). Ascorbate and glutathione: keeping active oxygen under control. *Annu. Rev. Plant Biol.* 49, 249–279. doi: 10.1146/annurev.arplant.49.1.249
- Ohkuma, K., Lyon, J. L., Addicott, F. T., and Smith, O. E. (1963). Abscission-accelerating substance from young cotton fruit. *Science* 142, 1592–1593.
- Okuma, E., Jahan, M. S., Munemasa, S., Hossain, M. A., Muroyama, D., Islam, M. M., et al. (2011). Negative regulation of abscisic acid-induced stomatal closure by glutathione in *Arabidopsis*. *J. Plant Physiol.* 168, 2048–2055. doi: 10.1016/j.jplph.2011.06.002
- Ou, X., Gan, Y., Chen, P., Qiu, M., Jiang, K., and Wang, G. (2014). Stomata prioritize their responses to multiple biotic and abiotic signal inputs. *PLoS ONE* 9:e101587. doi: 10.1371/journal.pone.0101587
- Outlaw, W. H. (1989). Critical examination of the quantitative evidence for and against photosynthetic CO_2 fixation by guard cells. *Physiol. Plant.* 77, 275–281. doi: 10.1111/j.1399-3054.1989.tb04981.x
- Outlaw, W. H., and Lowry, O. H. (1977). Organic acid and potassium accumulation in guard cells during stomatal opening. *Proc. Natl. Acad. Sci. U.S.A.* 74, 4434–4438.
- Outlaw, W. H. Jr, and De Vlieghere-He, X. (2001). Transpiration rate. An important factor controlling the sucrose content of the guard cell apoplast of broad bean. *Plant Physiol.* 126, 1716–1724. doi: 10.1104/pp.126.4.1716
- Pallas, J. E., and Kays, S. J. (1982). Inhibition of photosynthesis by ethylene—a stomatal effect. *Plant Physiol.* 70, 598–601. doi: 10.1104/pp.70.2.598
- Park, K. Y., Jung, J. Y., Park, J., Hwang, J. U., Kim, Y. W., and Hwang, I. (2003). A role for phosphatidylinositol 3-phosphate in abscisic acid-induced reactive oxygen species generation in guard cells. *Plant Physiol.* 132, 92–98. doi: 10.1104/pp.102.016964
- Park, S. Y., Fung, P., Nishimura, N., Jensen, D. R., Fujii, H., Zhao, Y., et al. (2009). Abscisic acid inhibits type 2C protein phosphatases via the PYR/PYL family of START proteins. *Science* 324, 1068–1071. doi: 10.1126/science.1173041
- Pearson, C. J. (1973). Daily changes in stomatal aperture and in carbohydrates and malate within epidermis and mesophyll of leaves of *Commelina cyanea* and *Vicia faba*. *Aust. J. Biol. Sci.* 26, 1035–1044.
- Penfield, S., Clements, S., Bailey, K. J., Gilday, A. D., Leegood, R. C., Gray, J. E., et al. (2012). Expression and manipulation of PHOSPHOENOLPYRUVATE CARBOXYKINASE 1 identifies a role for malate metabolism in stomatal closure. *Plant J.* 69, 679–688. doi: 10.1111/j.1365-3113.2011.04822.x
- Pharmawati, M., Billington, T., and Gehring, C. A. (1998). Stomatal guard cell responses to kinetin and natriuretic peptides are cGMP-dependent. *Cell. Mol. Life Sci.* 54, 272–276. doi: 10.1007/s000180050149
- Radin, J. W., Lu, Z., Percy, R. G., and Zeiger, E. (1994). Genetic variability for stomatal conductance in Pima cotton and its relation to improvements of heat adaptation. *Proc. Natl. Acad. Sci. U.S.A.* 91, 7217–7221.
- Raghavendra, A. S., Rao, I. M., and Das, V. S. R. (1976). Characterization of abscisic acid inhibition of stomatal opening in isolated epidermal strips. *Plant Sci. Lett.* 6, 111–115. doi: 10.1016/0304-4211(76)90144-9
- Ramon, M., Rolland, F., and Sheen, J. (2008). Sugar sensing and signaling. *Arabidopsis Book* 6:e0117. doi: 10.1199/tab.0117
- Reckmann, U., Scheibe, R., and Raschke, K. (1990). Rubisco activity in guard cells compared with the solute requirement for stomatal opening. *Plant Physiol.* 92, 246–253. doi: 10.1104/pp.92.1.246
- Rognoni, S., Teng, S., Arru, L., Smeeckens, S. C. M., and Perata, P. (2007). Sugar effects on early seedling development in *Arabidopsis*. *Plant Growth Regul.* 52, 217–228. doi: 10.1007/s10725-007-9193-z

- Rolland, F., Baena-Gonzalez, E., and Sheen, J. (2006). Sugar sensing and signaling in plants: conserved and novel mechanisms. *Annu Rev. Plant Biol.* 57, 675–709. doi: 10.1146/annurev.arplant.57.032905.105441
- Santiago, J., Dupeux, F., Betz, K., Antoni, R., Gonzalez-Guzman, M., Rodriguez, L., et al. (2012). Structural insights into PYR/PYL/RCAR ABA receptors and PP2Cs. *Plant Sci.* 182, 3–11. doi: 10.1016/j.plantsci.2010.11.014
- Savchenko, T., Kolla, V. A., Wang, C. Q., Nasafi, Z., Hicks, D. R., Phadungchob, B., et al. (2014). Functional convergence of oxylipin and abscisic acid pathways controls stomatal closure in response to drought. *Plant Physiol.* 164, 1151–1160. doi: 10.1104/pp.113.234310
- Schroeder, J. I., Kwak, J. M., and Allen, G. J. (2001). Guard cell abscisic acid signaling and engineering drought hardiness in plants. *Nature* 410, 327–330. doi: 10.1038/35066500
- Schulz, P., Jansseune, K., Degenkolbe, T., Méret, M., Claeys, H., Skirycz, A., et al. (2014). Poly (ADP-Ribose) polymerase activity controls plant growth by promoting leaf cell number. *PLoS ONE* 9:e90322. doi: 10.1371/journal.pone.0090322
- Schwartz, A., Wu, W. H., Tucker, E. B., and Assmann, S. M. (1994). Inhibition of inward K⁺ channels and stomatal response by abscisic acid: an intracellular locus of phytohormone action. *Proc. Natl. Acad. Sci. U.S.A.* 91, 4019–4023.
- Scuffi, D., Núñez, Á., Lasplina, N., Gotor, C., Lamattina, L., and Garcia-Mata, C. (2014). Hydrogen sulfide generated by L-cysteine desulhydrase acts upstream of nitric oxide to modulate ABA-dependent stomatal closure. *Plant Physiol.* 245373, 114. doi: 10.1104/pp.114.245373
- Seo, M., and Koshiba, T. (2011). Transport of ABA from the site of biosynthesis to the site of action. *J. Plant Res.* 124, 501–507. doi: 10.1007/s10265-011-0411-4
- Shi, C., Qi, C., Ren, H., Huang, A., Hei, S., and She, X. (2015). Ethylene mediates brassinosteroid-induced stomatal closure via Gα protein-activated hydrogen peroxide and nitric oxide production in *Arabidopsis*. *Plant J.* 82, 280–301. doi: 10.1111/tpj.12815
- Song, X., She, X., He, J., Huang, C., and Song, T. (2006). Cytokinin- and auxin-induced stomatal opening involves a decrease in levels of hydrogen peroxide in guard cells of *Vicia faba*. *Funct. Plant Biol.* 33, 573–583. doi: 10.1071/FP05232
- Song, X. G., She, X. P., Yue, M., Liu, Y. E., Wang, Y. X., Zhu, X., et al. (2014). Involvement of copper amine oxidase (CuAO)-dependent hydrogen peroxide synthesis in ethylene-induced stomatal closure in *Vicia faba*. *Russ. J. Plant Physiol.* 61, 390–396. doi: 10.1134/S1021443714020150
- Stadler, R., Buttner, M., Ache, P., Hedrich, R., Ivashikina, N., Melzer, M., et al. (2003). Diurnal and light-regulated expression of AtSTP1 in guard cells of *Arabidopsis*. *Plant Physiol.* 133, 528–537. doi: 10.1104/pp.103.024240
- Suhita, D., Raghavendra, A. S., Kwak, J. M., and Vasseur, A. (2004). Cytoplasmic alkalization precedes reactive oxygen species production during methyl jasmonate- and abscisic acid-induced stomatal closure. *Plant Physiol.* 134, 1536–1545. doi: 10.1104/pp.103.032250
- Sun, Q., Liu, J., Zhang, Q., Qing, X., Dobson, G., Li, X., et al. (2013). Characterization of three novel desaturases involved in the delta-6 desaturation pathways for polyunsaturated fatty acid biosynthesis from *Phytophthora infestans*. *Appl. Microbiol. Biotechnol.* 97, 7689–7697. doi: 10.1007/s00253-012-4613-z
- Sun, Z., Jin, X., Albert, R., and Assmann, S. M. (2014). Multi-level modeling of light-induced stomatal opening offers new insights into its regulation by drought. *PLoS Comp. Biol.* 10:e1003930. doi: 10.1371/journal.pcbi.1003930
- Talbott, L. D., and Zeiger, E. (1988). Light quality and osmoregulation in *Vicia* guard cells: evidence for involvement of three metabolic pathways. *Plant Physiol.* 88, 887–895. doi: 10.1104/pp.88.3.887
- Talbott, L. D., and Zeiger, E. (1993). Sugar and organic acid accumulation in guard cells of *Vicia faba* in response to red and blue light. *Plant Physiol.* 102, 1163–1169.
- Talbott, L. D., and Zeiger, E. (1996). Central roles for potassium and sucrose in guard-cell osmoregulation. *Plant Physiol.* 111, 1051–1057.
- Tallman, G. (2004). Are diurnal patterns of stomatal movement the result of alternating metabolism of endogenous guard cell ABA and accumulation of ABA delivered to the apoplast around guard cells by transpiration? *J. Exp. Bot.* 55, 1963–1976. doi: 10.1093/jxb/erh212
- Tanaka, Y., Sano, T., Tamaoki, M., Nakajima, N., Kondo, N., and Hasezawa, S. (2005). Ethylene inhibits abscisic acid-induced stomatal closure in *Arabidopsis*. *Plant Physiol.* 138, 2337–2343. doi: 10.1104/pp.105.063503
- Tian, W., Hou, C., Ren, Z., Pan, Y., Jia, J., Zhang, H., et al. (2015). A molecular pathway for CO₂ response in *Arabidopsis* guard cells. *Nat. Commun.* 6:6057. doi: 10.1038/ncomms7057
- Umezawa, T., Nakashima, K., Miyakawa, T., Kuromori, T., Tanokura, M., Shinozaki, K., et al. (2010). Molecular basis of the core regulatory network in ABA responses: sensing, signaling and transport. *Plant Cell Physiol.* 51, 1821–1839. doi: 10.1093/pcp/pcq156
- Waadt, R., Hitomi, K., Nishimura, N., Hitomi, C., Adams, S. R., Getzoff, E. D., et al. (2014). FRET-based reporters for the direct visualization of abscisic acid concentration changes and distribution in *Arabidopsis*. *eLife* 3:e01739. doi: 10.7554/eLife.01739
- Wang, P., Du, Y., Hou, Y. J., Zhao, Y., Hsu, C. C., Yuan, F., et al. (2015). Nitric oxide negatively regulates abscisic acid signaling in guard cells by S-nitrosylation of OST1. *Proc. Natl. Acad. Sci. U.S.A.* 112, 613–618. doi: 10.1073/pnas.1423481112
- Wang, R. S., Pandey, S., Li, S., Gookin, T. E., Zhao, Z., Albert, R., et al. (2011). Common and unique elements of the ABA-regulated transcriptome of *Arabidopsis* guard cells. *BMC Genomics* 12:216. doi: 10.1186/1471-2164-12-216
- Wang, X. Q., Ullah, H., Jones, A. M., and Assmann, S. M. (2001). G protein regulation of ion channels and abscisic acid signaling in *Arabidopsis* guard cells. *Science* 292, 2070–2072. doi: 10.1126/science.1059046
- Wasilewska, A., Vlad, F., Sirichandra, C., Redko, Y., Jammes, F., Valon, C., et al. (2008). An update on abscisic acid signaling in plants and more. *Mol. Plant* 1, 198–217. doi: 10.1093/mp/ssm022
- Watkins, J. M., Hechler, P. J., and Muday, G. K. (2014). Ethylene-induced flavonol accumulation in guard cells suppresses reactive oxygen species and moderates stomatal aperture. *Plant Physiol.* 164, 1707–1717. doi: 10.1104/pp.113.233528
- Weise, A., Lalonde, S., Kuhn, C., Frommer, W. B., and Ward, J. M. (2008). Introns control expression of sucrose transporter LeSUT1 in trichomes, companion cells and in guard cells. *Plant Mol. Biol.* 68, 251–262. doi: 10.1007/s11103-008-9366-9
- Willmer, C. M., Don, R., and Parker, W. (1978). Levels of short-chain fatty acids and of abscisic acid in water-stressed and non-stressed leaves and their effects on stomata in epidermal strips and excised leaves. *Planta* 139, 281–287. doi: 10.1007/BF00388642
- Winkel-Shirley, B. (2002). Biosynthesis of flavonoids and effects of stress. *Curr. Opin. Plant Biol.* 5, 218–223. doi: 10.1016/S1369-5266(02)00256-X
- Worrall, D., Liang, Y. K., Alvarez, S., Holroyd, G. H., Spiegel, S., Panagopoulos, et al. (2008). Involvement of sphingosine kinase in plant cell signalling. *Plant J.* 56, 64–72. doi: 10.1111/j.1365-313X.2008.03579.x
- Wu, Y., Kuzma, J., Maréchal, E., Graeff, R., Lee, H. C., Foster, R., et al. (1997). Abscisic acid signaling through cyclic ADP-ribose in plants. *Science* 278, 2126–2130. doi: 10.1126/science.278.5346.2126
- Xia, X. J., Gao, C. J., Song, L. X., Zhou, Y. H., Shi, K., and Yu, J. Q. (2014). Role of H₂O₂ dynamics in brassinosteroid-induced stomatal closure and opening in *Solanum lycopersicum*. *Plant Cell Environ.* 37, 2036–2050. doi: 10.1111/pce.12275
- Yamashita, T. (1952). Influences of potassium supply upon various properties and movement of guard cell. *Sielboldia Acta Biologica* 1, 51–70.
- Yamazaki, D., Yoshida, S., Asami, T., and Kuchitsu, K. (2003). Visualization of abscisic acid-perception sites on the plasma membrane of stomatal guard cells. *Plant J.* 35, 129–139. doi: 10.1046/j.1365-313X.2003.01782.x
- Yan, X. F., and Chen, S. (2007). Regulation of plant glucosinolate metabolism. *Planta* 226, 1343–1352. doi: 10.1007/s00425-007-0627-7
- Yu, Y., and Assmann, S. M. (2014). Metabolite transporter regulation of ABA function and guard cell response. *Mol. Plant* 7, 1505–1507. doi: 10.1093/mp/ssu093
- Yuan, X., Li, Y., Liu, S., Xia, F., Li, X., and Qi, B. (2014). Accumulation of eicosapolyenoic acids enhances sensitivity to abscisic acid and mitigates the effects of drought in transgenic *Arabidopsis thaliana*. *J. Exp. Bot.* 65, 1637–1649. doi: 10.1093/jxb/eru031
- Zeiger, E., and Zhu, J. (1998). Role of zeaxanthin in blue light photoreception and the modulation of light-CO₂ interactions in guard cells. *J. Exp. Bot.* 49, 433–442. doi: 10.1093/jxb/49.Special_Issue.433
- Zeng, W., and He, S. Y. (2010). A prominent role of the flagellin receptor FLAGELLIN-SENSING2 in mediating stomatal response to *Pseudomonas syringae* pv tomato DC3000 in *Arabidopsis*. *Plant Physiol.* 153, 1188–1198. doi: 10.1104/pp.110.157016
- Zhang, Y., Zhu, H., Zhang, Q., Li, M., Yan, M., Wang, R., et al. (2009). Phospholipase Dα1 and phosphatidic acid regulate NADPH oxidase activity and production of reactive oxygen species in ABA-mediated stomatal closure in *Arabidopsis*. *Plant Cell* 21, 2357–2377. doi: 10.1105/tpc.108.062992

- Zhang, W., He, S. Y., and Assmann, S. M. (2008). The plant innate immunity response in stomatal guard cells invokes G-protein-dependent ion channel regulation. *Plant J.* 56, 984–996. doi: 10.1111/j.1365-313X.2008.03657.x
- Zhang, W., Jeon, B. W., and Assmann, S. M. (2011). Heterotrimeric G-protein regulation of ROS signalling and calcium currents in *Arabidopsis* guard cells. *J. Exp. Bot.* 62, 2371–2379. doi: 10.1093/jxb/erq424
- Zhang, T., Chen, S., and Harmon, A. (2014). Protein phosphorylation in stomatal movement. *Plant Signal. Behav.* 9:e972845. doi: 10.4161/15592316.2014.972845
- Zhang, X. L., Jiang, L., Xin, Q., Liu, Y., Tan, J. X., and Chen, Z. Z. (2015). Structural basis and functions of abscisic acid receptors PYLs. *Front. Plant Sci.* 6:88. doi: 10.3389/fpls.2015.00088
- Zhang, X., Dong, F. C., Gao, J. F., and Song, C. P. (2001). Hydrogen peroxide-induced changes in intracellular pH of guard cells precede stomatal closure. *Cell Res.* 11, 37–43. doi: 10.1038/sj.cr.7290064
- Zhao, Z., and Assmann, S. M. (2011). The glycolytic enzyme, phosphoglycerate mutase, has critical roles in stomatal movement, vegetative growth, and pollen production in *Arabidopsis thaliana*. *J. Exp. Bot.* 62, 5179–5189. doi: 10.1093/jxb/err223
- Zhao, Z., Zhang, W., Stanley, B. A., and Assmann, S. M. (2008). Functional proteomics of *Arabidopsis thaliana* guard cells uncovers new stomatal signaling pathways. *Plant Cell* 20, 3210–3226. doi: 10.1105/tpc.108.063263
- Zhu, G. H., Liu, Y. G., Ye, N. H., Liu, R., and Zhang, J. H. (2011). Involvement of the abscisic acid catabolic gene CYP707A2 in the glucose-induced delay in seed germination and post-germination growth of *Arabidopsis*. *Physiol. Plant.* 143, 375–384. doi: 10.1111/j.1399-3054.2011.01510.x
- Zhu, M., Dai, S., Zhu, N., Booy, A., Simons, B., Yi, S., et al. (2012). Methyl jasmonate responsive proteins in *Brassica napus* guard cells revealed by iTRAQ-based quantitative proteomics. *J. Proteome Res.* 11, 3728–3742. doi: 10.1021/pr300213k
- Zhu, M., Simons, B., Zhu, N., Oppenheimer, D. G., and Chen, S. (2010). Analysis of abscisic acid responsive proteins in *Brassica napus* guard cells by multiplexed isobaric tagging. *J. Proteomics* 73, 790–805. doi: 10.1016/j.jprot.2009.11.002
- Zhu, M., Zhu, N., Song, W. Y., Harmon, A. C., Assmann, S. M., and Chen, S. (2014). Thiol-based redox proteins in abscisic acid and methyl jasmonate signaling in *Brassica napus* guard cells. *Plant J.* 78, 491–515. doi: 10.1111/tpj.12490

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Misra, Acharya, Granot, Assmann and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Single cell-type comparative metabolomics of epidermal bladder cells from the halophyte *Mesembryanthemum crystallinum*

Bronwyn J. Barkla^{1*} and Rosario Vera-Estrella²

¹ Southern Cross Plant Science, Southern Cross University, Lismore, NSW, Australia, ² Departamento de Biología Molecular de Plantas, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Mexico

OPEN ACCESS

Edited by:

Sixue Chen,
University of Florida, USA

Reviewed by:

Steffen Neumann,
Leibniz Institute of Plant Biochemistry,
Germany
Jedrzej Jakub Szymanski,
Weizmann Institute of Science, Israel

*Correspondence:

Bronwyn J. Barkla,
Southern Cross Plant Science,
Southern Cross University, Military
Road, PO Box 157, Lismore, NSW
2480, Australia
bronwyn.barkla@scu.edu.au

Specialty section:

This article was submitted to
Plant Systems and Synthetic Biology,
a section of the journal
Frontiers in Plant Science

Received: 02 May 2015

Accepted: 27 May 2015

Published: 10 June 2015

Citation:

Barkla BJ and Vera-Estrella R (2015)
Single cell-type comparative
metabolomics of epidermal bladder
cells from the halophyte
Mesembryanthemum crystallinum.
Front. Plant Sci. 6:435.
doi: 10.3389/fpls.2015.00435

One of the remarkable adaptive features of the halophyte *Mesembryanthemum crystallinum* are the specialized modified trichomes called epidermal bladder cells (EBC) which cover the leaves, stems, and peduncle of the plant. They are present from an early developmental stage but upon salt stress rapidly expand due to the accumulation of water and sodium. This particular plant feature makes it an attractive system for single cell type studies, with recent proteomics and transcriptomics studies of the EBC establishing that these cells are metabolically active and have roles other than sodium sequestration. To continue our investigation into the function of these unusual cells we carried out a comprehensive global analysis of the metabolites present in the EBC extract by gas chromatography Time-of-Flight mass spectrometry (GC-TOF) and identified 194 known and 722 total molecular features. Statistical analysis of the metabolic changes between control and salt-treated samples identified 352 significantly differing metabolites (268 after correction for FDR). Principal components analysis provided an unbiased evaluation of the data variance structure. Biochemical pathway enrichment analysis suggested significant perturbations in 13 biochemical pathways as defined in KEGG. More than 50% of the metabolites that show significant changes in the EBC, can be classified as compatible solutes and include sugars, sugar alcohols, protein and non-protein amino acids, and organic acids, highlighting the need to maintain osmotic homeostasis to balance the accumulation of Na⁺ and Cl⁻ ions. Overall, the comparison of metabolic changes in salt treated relative to control samples suggests large alterations in *M. crystallinum* epidermal bladder cells.

Keywords: salinity, salt-tolerance, metabolomics, halophyte, trichomes, epidermal bladder cells, Crassulacean acid metabolism, *Mesembryanthemum crystallinum*

Introduction

Single cell-type studies allow insight into the highly specific processes and functions of differentiated cells. In plants, studies into the role of specialized cells have been undertaken with a particular focus on those cells in which the isolation procedure is relatively uncomplicated. This includes free moving cell-types such as pollen grains, but also cells of the epidermis, such as root

hairs, guard cells, and trichomes (Misra et al., 2014). Trichomes are widely conserved across the plant kingdom but are highly divergent in relation to morphology and function. They can be simple unicellular structures like the hairs of *Arabidopsis*, or complex multicellular appendages made up of differentiated basal, stalk, and apical cells; with the possibility of multiple trichome types co-existing on the same organ, as is observed for tomato (Kang et al., 2009). They produce, store and, in the case of glandular trichomes, secrete a diverse range of chemical compounds including polar and non-polar metabolites, which have been implicated in adaptive responses of the plant to their biotic and abiotic environments (Schilmiller et al., 2008). Trichomes can function to deter insects and herbivores, attract pollinators, aid in seed dispersal, regulate leaf temperature, but also to store unwanted xenobiotics including heavy metals and salts (Wagner et al., 1994).

In the Aizoaceae, specialized single cell trichomes called epidermal bladder cells (EBC) are a distinctive feature of this family (Klak et al., 2003). EBC of the succulent desert halophyte and facultative Crassulacean acid metabolism plant *Mesembryanthemum crystallinum* are abundant on leaves and stems from an early developmental age; however, cell morphology changes with development and metabolic/stress state of the plant (Adams et al., 1992). In young unstressed plants the EBC are small and tightly appressed to the leaf surface, whereas in adult salt-treated plants the EBC become enlarged and can be balloon- or sausage-like (Oh et al., 2015). They have been shown to accumulate high concentrations of sodium and thought to be important salt-adaptive features of the plant (Adams et al., 1998; Barkla et al., 2002). “Omics” approaches have helped to define a more encompassing role for the EBC, with proteomics and transcriptomics analysis suggesting these cells are metabolically active (Barkla et al., 2012), and show pronounced alternations in response to salt in a number of precisely defined pathways including significant changes in transcripts from networks representing Gene Ontology (GO) terms for ion transport, osmolyte accumulation, and stress signaling (Oh et al., 2015). Here we continue our systems wide integrative investigation of EBC in the facultative CAM plant *M. crystallinum* by carrying out non-targeted metabolite profiling of EBC extracts from plants under control and salinity treatment regimens to obtain a snapshot of EBC metabolism. Overall, the comparison of metabolic changes in salt treated relative to control samples suggested very large perturbations in metabolites between the treatment conditions and highlighted 13 significantly enriched biochemical pathways.

Material and Methods

Plant Materials and Growth Conditions

M. crystallinum L. plants were grown from seed in soil (MetroMix 510; Sun Gro Horticulture, Bellevue, WA) in a propagation tray as previously described (Barkla et al., 2009). Three weeks following germination, individual seedlings were transplanted to pots containing the soil mixture, with two plants per 15-cm-diameter pot. Plants were watered daily and one-half strength Hoagland’s medium (Hoagland and Arnon, 1938) was supplied

weekly. NaCl (200 mM) treatment was initiated 6 weeks after germination for a period of 14 d. Plants were grown in a glasshouse under natural irradiation and photoperiod, with the greenhouse photosynthetic photon flux density reaching a peak value of $1300 \text{ mmol m}^{-2} \text{ s}^{-1}$ during the middle of the day. Temperature was maintained at $25^\circ\text{C} \pm 3^\circ\text{C}$.

Extraction of Bladder Cell Extract

Bladder cell extract was obtained from individual cells on the leaf abaxial epidermal surface and stems by vacuum aspiration of the cell contents using a fine gage insulin needle (27 G, 13 mm) attached to a collection reservoir maintained on ice. The needle was oriented horizontally to the leaf or stem axis to avoid removing sap from underlying tissue and the procedure was visualized using a Nikon SMZ645 stereo microscope equipped with a dual arm Nikon MKII fiber optic light source (Nikon, México). The extracted liquid from approximately 2000 EBC estimated by counts of cell yield per leaf/stem section from a single control or salt-treated plant was pooled to obtain a single biological replicate. Seven biological replicates were collected per treatment. Plants were maintained in the dark prior to extraction and collection was undertaken early in the morning.

Sample Preparation for Metabolomics Analysis

Bladder cell extract was aliquoted into 1.5 ml Eppendorf tubes and the extract was evaporated in a Labconco Centrivap concentrator (Kansas City, MO, USA) to complete dryness. The dried extract was then resuspended in the indicated pre-chilled (to -20°C) nitrogen degassed extraction solution. Samples were vortexed for 10 s, and placed on an orbital shaker at 4°C for 6 min and finally dried in the Labconco Centrivap cold trap concentrator to complete dryness prior to derivitization. Two experimental replicates were obtained for each biological replicate.

Gas Chromatography Time-of-flight Mass Spectrometry (GC-TOF)

GC-TOF was performed by West Coast Metabolomics at the UC Davis Genome Centre (Davis CA, USA). Mass spectrometry was performed using a Pegasus III Time Of Flight (TOF) mass spectrometer (LECO, St. Joseph, MI, USA) coupled to an Agilent 6890N gas chromatograph (Agilent Technologies) equipped with a Gerstel autosampler, including a MPS PrepStation and Automated Liner EXchange (ALEX) (Gerstel, Muehlheim, Germany). Liner was changed after every 10 samples, (using the Maestro1 Gerstel software vs. 1.1.4.18). Before and after each injection, the $10 \mu\text{l}$ injection syringe was washed three times with $10 \mu\text{l}$ ethyl acetate. The Gas Chromatograph was equipped with a 30 m long, 0.25 mm i.d. Rtx-5Sil MS column (0.25 μm 95% dimethyl 5% diphenyl polysiloxane film) with additional 10 m integrated guard column (Restek, Bellefonte PA, USA). Pure helium (99.9999%) with built-in purifier (Airgas, Radnor PA, USA) was set at constant flow of 1 ml/min. The oven temperature was held constant at 50°C for 1 min and then ramped at $20^\circ\text{C}/\text{min}$ to 330°C at which it was held constant for 5 min. The injector was set to a temperature of 250°C and was used in split mode with 10:1 ratio. The injection volume

was 1 μ l. The transfer line temperature was set to 280°C (GC re-entrant temperature). The mass spectrometer, controlled by Leco ChromaTOF software vs. 2.32 (St. Joseph, MI, USA), was operated in positive EI mode. All EI spectra were collected using an electron energy of 70 eV, trap current of 250 mA, emission current of 600 mA, filament current of 4.5 A, and source temperature of 250°C. Acquisition rate was 17 spectra/s, with a scan mass range of 85–500 Da. Spectra were deconvoluted from co-eluting peaks with ChromaTOF, this software detects peaks in an unbiased way and exports one deconvoluted spectrum per peak.

Metabolite Identifications

Following acquisition of data, a BinBase relational database system was used to allow for automated metabolite annotation. Metabolites were grouped into either a list of, “identified compounds” or a list of “unknown compounds” and assigned as previously described Lee and Fiehn (2008). Accordingly both groups were unambiguously assigned using the identification criteria of retention index and mass spectrum to BinBase identifier numbers. Additional confidence criteria were given by mass spectral metadata, using the combination of unique ions, apex ions, peak purity, and signal/noise ratios. Each BinBase identifier was routinely matched against the Fiehn lab mass spectral library (Fiehn et al., 2005). For named BinBase compounds, PubChem numbers and KEGG identifiers were added. In addition, for all reported compounds (identified and unknown metabolites) the quantification ion and the full mass spectrum encoded as string is reported.

Data Analysis and Accessibility

Statistical analyses were conducted on natural logarithm transformed metabolic parameters, and data summaries are presented for raw data values. HCA of samples, and PCA were implemented on autoscaled data (mean centered and scaled by the standard deviation, z-scaled). The data obtained in this study will be accessible at the NIH Common Fund’s Data Repository and Coordinating Center (supported by NIH grant, U01-DK097430) website, <http://www.metabolomicsworkbench.org>.

Results

Extraction of Metabolites

Untargeted, global metabolite profiling studies aim to analyse as many small-molecular-weight species with a single sample extraction protocol as possible. However, due to the large variety of metabolites with different chemical stability, solubility, and polarity, the choice of sample-preparation method can greatly influence the observed metabolite content of a cellular extract. Ideally the method should be as non-selective as possible to ensure maximal depth of coverage. We first established the optimal sample extraction protocol for GC-TOF of EBC extracts that would allow us to obtain the greatest coverage of the EBC metabolome. Three solvent extraction methods were tested: 9:1 methanol/chloroform, 5:2:2 methanol/chloroform/water and 10:3:1 methanol/chloroform/water. Varying the ratio of aqueous

and organic solvents can influence the recovery of polar and non-polar metabolites. Supplementary Table 1 shows the metabolite profile of EBC’s extracted using the different solvent regimes and two different volumes (200 and 500 μ l). In all extracts we were able to identify 668 different molecular features (175 known and 493 unknown). However, the amounts extracted varied between the different procedures (Supplementary Table 1) suggesting a variation in recovery efficiency similar to what was shown for yeast metabolite extraction (Tambellini et al., 2013). Based on this information we selected 5:2:2 methanol/chloroform/water solvent to water ratio which has been shown to extract comparatively high levels of both polar metabolites and non-polar fatty acids (Lee and Fiehn, 2008).

Alteration in Metabolites in Salt-Treated EBC Samples

To explore the effect of salinity on EBC metabolite levels, samples from 7 independent biological replicates for untreated and salinity-treated plants were collected, extracted with 5:2:2 methanol/chloroform/water, and metabolic parameters (194 known and 722 total molecular features) were compared Supplementary Table 2. A two-sample Student’s *t*-Test was used to assess the significance of the difference between the control and salt-treated samples. The probability level for the *p*-values were adjusted to allow for a maximum 5% probability ($q = 0.05$) of false positives (Benjamini and Hochberg, 1995). The FDR was also directly estimated as the *q*-value, for all comparisons (Klaus and Strimmer, 2012). Statistical test results are reported along with metabolite averages, standard deviations and fold changes of means in Supplementary Table 2. A total of 352 significantly differing metabolites (268 after correction for FDR) were identified (Supplementary Table 2). Of the known metabolites, 28 showed a statistically significant down-regulation of more than 1.5 fold (Table 1) and 28 metabolites showed a statistically significant up-regulation of more than 1.5 fold (Table 1). The top 10 known, significant down- and up-regulated metabolites are depicted in box plots (Figures 1A,B). Of the significantly changing unknown metabolites, nine showed more than a 5-fold increase in the salt-treated samples compared to the control samples, while an additional 13 were shown to be significantly down-regulated by more than 5-fold (Supplementary Table 2).

Hierarchical Cluster Analysis

Hierarchical cluster analysis (HCA) was used to group samples and metabolites based on similarities in auto scaled values and correlations, respectively (Sugimoto et al., 2012). Distance was calculated based on the Euclidean method, linkage was done using the Ward method (Ward, 1963), and variable similarities were based on Spearman correlations. Multivariate sample similarities, displayed as a heatmap, are shown in Figure 2. Separation of experimental treatments or sample classes into non-overlapping clusters suggests large metabolic differences between the comparisons. While separation of similar classes of samples into clusters may suggest high biological or analytical variability.

TABLE 1 | Metabolites which were significantly down- or up-regulated by more than 1.5-fold in EBC extracts from salt-treated plants compared to control plants.

Metabolite name	Metabolite classification	Fold change ^a	p-values ^b	Adjusted p-values ^c	q-values ^d
DOWN-REGULATED					
Tyramine	amine	0.36	0.00008	0.00151	0.00035
Putrescine	amine	0.40	0.00128	0.00947	0.00221
Phenylethylamine	amine	0.44	0.01377	0.04125	0.00961
Ornithine	amine	0.44	0.01629	0.04700	0.01095
Tryptophan	amino acid	0.25	0.00739	0.02681	0.00624
Histidine	amino acid	0.49	0.00559	0.02307	0.00537
Noradrenaline	hormone	0.11	0.00165	0.01048	0.00244
Cysteine-glycine	non-protein amino acid	0.32	0.00150	0.00982	0.00229
Kynurenine	non-protein amino acid	0.33	0.00000	0.00024	0.00006
Alpha ketoglutaric acid	organic acid	0.12	0.00002	0.00047	0.00011
Ascorbic acid	organic acid	0.20	0.01459	0.04317	0.01005
Isocitric lactone	organic acid	0.25	0.00229	0.01342	0.00312
Oxalic acid	organic acid	0.27	0.01407	0.04197	0.00977
2-hydroxyglutaric acid	organic acid	0.39	0.00065	0.00582	0.00136
Dihydroxymalonic acid	organic acid	0.44	0.00073	0.00618	0.00144
2-hydroxyadipic acid	organic acid	0.46	0.00299	0.01574	0.00367
Ethanol phosphate	other	0.09	0.00568	0.02319	0.00540
Glycerol-alpha-phosphate	phosphate	0.38	0.00007	0.00129	0.00030
Adenine	purine	0.20	0.00115	0.00880	0.00205
Uridine	pyrimidine	0.31	0.00664	0.02539	0.00591
Inulobiose	sugar	0.30	0.00011	0.00164	0.00038
Maltose	sugar	0.37	0.00064	0.00582	0.00136
Cellobiose	sugar	0.42	0.00072	0.00618	0.00144
Beta-gentiobiose	sugar	0.50	0.00026	0.00289	0.00067
Lactobionic acid	sugar acid	0.21	0.00000	0.00006	0.00001
Galactonic acid	sugar acid	0.49	0.00028	0.00303	0.00071
Erythritol	sugar alcohol	0.30	0.00030	0.00311	0.00072
Ribitol	sugar alcohol	0.47	0.00165	0.01048	0.00244
UP-REGULATED					
Triethanolamine	Amine	2.15	0.00239	0.01370	0.00319
Threonine	Amino acid	2.20	0.00011	0.00163	0.00038
Methionine	Amino acid	2.49	0.00129	0.00947	0.00221
Adenosine	Amino acid	2.59	0.01693	0.04775	0.01112
Asparagine	Amino acid	2.95	0.00219	0.01320	0.00307
Valine	Amino acid	3.09	0.00189	0.01166	0.00272
Sinapyl alcohol	Cell wall lignification	1.79	0.00674	0.02539	0.00591
1-monopalmitin	Fatty acid	1.78	0.02282	0.05604	0.01305
Dodecanol	Fatty alcohol	2.22	0.00335	0.01670	0.00389
4-hydroxybutyric acid	Non-protein amino acid	1.62	0.01725	0.04807	0.01120
Beta-alanine	Non-protein amino acid	2.27	0.00778	0.02766	0.00644
N-methylalanine	Non-protein amino acid	2.50	0.00258	0.01423	0.00331
Proline	Non-protein amino acid	5.94	0.00002	0.00047	0.00011
Pipecolic acid	Non-protein amino acid	12.71	0.00429	0.01923	0.00448
Dehydroascorbic acid	Organic acid	2.36	0.00267	0.01462	0.00340
Itaconic acid	Organic acid	2.81	0.00777	0.02766	0.00644
Maleic acid	Organic acid	9.31	0.02157	0.05427	0.01264
Butyrolactam	Other	2.36	0.01847	0.04975	0.01159
1-methyl-1-(4-methyl-3-cyclohexen-1-yl)ethoxy	Other	2.69	0.00410	0.01892	0.00441
Xanthine	Purine	5.28	0.01299	0.03990	0.00929

(Continued)

TABLE 1 | Continued

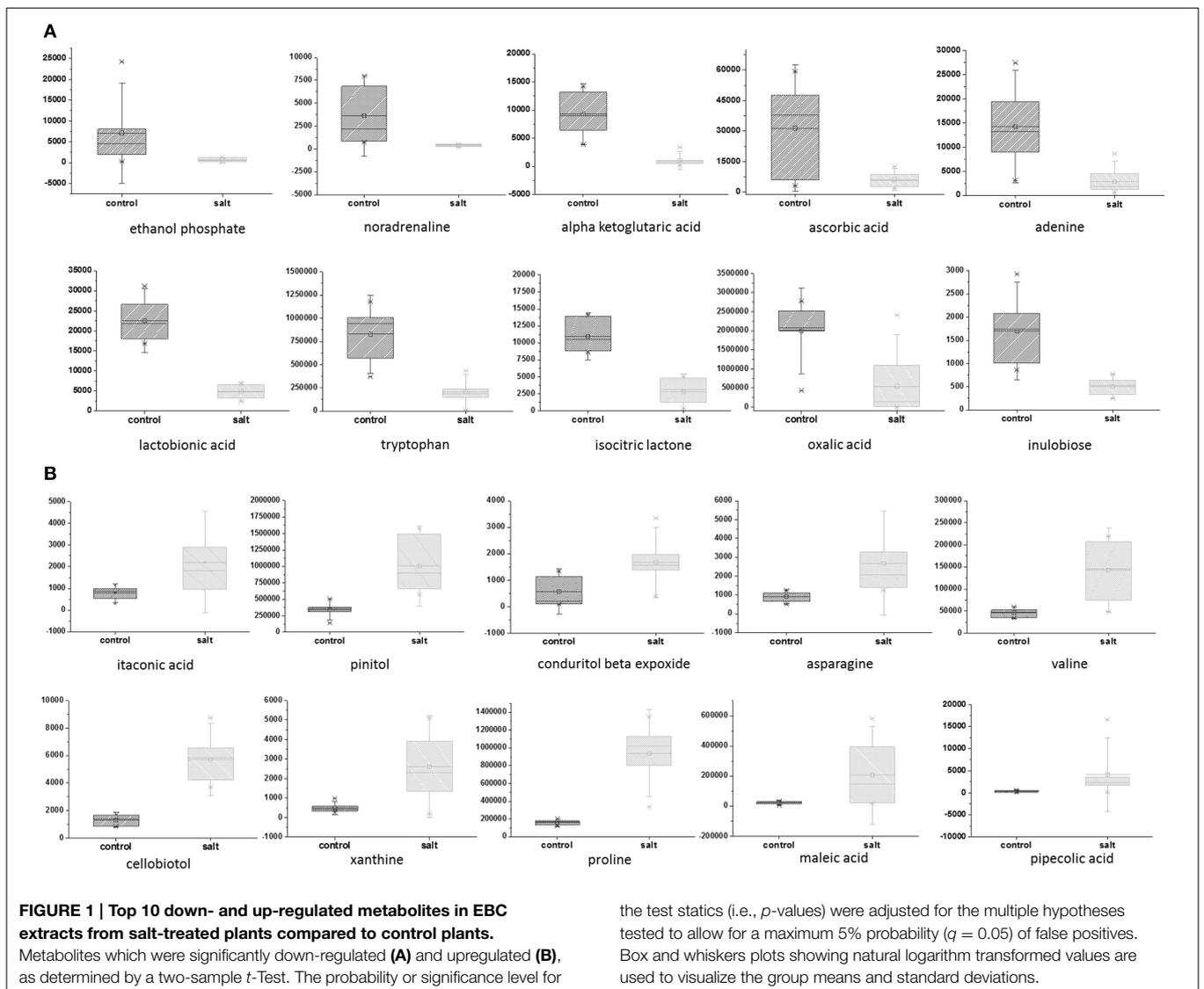
Metabolite name	Metabolite classification	Fold change ^a	<i>p</i> -values ^b	Adjusted <i>p</i> -values ^c	<i>q</i> -values ^d
Uracil	Pyrimidine	2.53	0.01100	0.03497	0.00815
Xylose	Sugar	1.73	0.00893	0.03042	0.00709
Fructose	Sugar	2.01	0.01239	0.03873	0.00902
Glycerol	Sugar alcohol	1.82	0.00094	0.00743	0.00173
1,2-anhydro-myo-inositol	Sugar alcohol	2.68	0.00005	0.00095	0.00022
Pinitol NIST	Sugar alcohol	2.91	0.00033	0.00338	0.00079
Conduritol beta expoxide	Sugar alcohol	2.92	0.01335	0.04081	0.00950
Cellobiotol	Sugar alcohol	4.32	0.00000	0.00007	0.00002

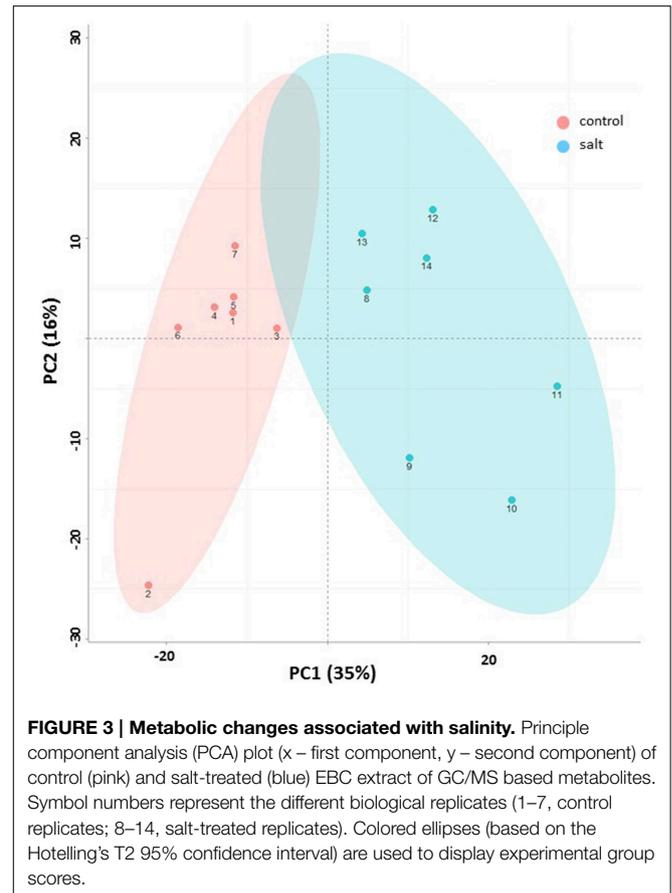
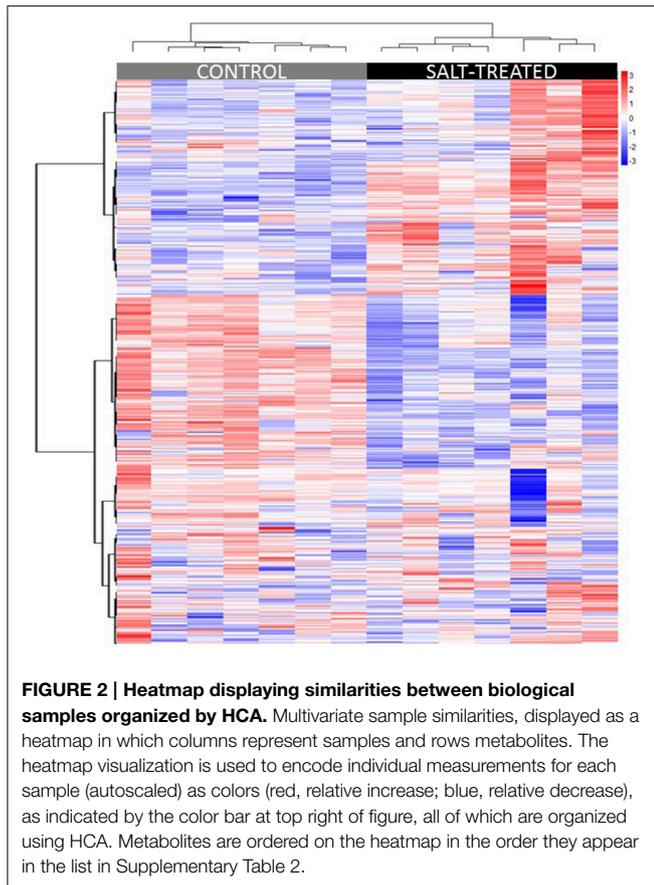
^a Fold change—mean salt-treated value/mean control value.

^b *p*-values—Significance levels for the two-sample Student's *t*-Test were computed for assessing the significance of difference between salt-treated samples and control samples.

^c Adjusted *p*-values—the significance level for the test statistics were adjusted for the multiple hypotheses tested to allow for a maximum 5% probability of false positives (*p*FDR-adjusted *p*-values).

^d *q*-values—estimate of the FDR for all comparisons.





Principal Component Analysis

Principal component analysis (PCA) provided an unbiased evaluation of the data variance structure. **Figure 3** shows the PCA for the dataset of EBC extracts from control and salt treated plants for the first two principal components. As can be observed the experiments are clearly separated into two distinct groups which reflect the treatment conditions. The first principal component (PC1)—explained the greatest variance (35%) across the data and separates the samples based on treatment. The second principal component (PC2) separated the components based on sample replicates and accounted for 16% of the variance.

Biochemical Pathway Enrichment

Pathway enrichment analysis was used to identify if the significantly changed metabolites, between control and salt-treated EBC samples, were significantly enriched in members of biochemical pathways as defined by the KEGG Database. Significant enrichment in KEGG pathways was identified using MBRole (Chagoyen and Pazos, 2011). The hypergeometric test coupled with adjustment for FDR (Benjamini and Hochberg, 1995) was used to identify 13 significantly enriched pathways (**Table 2**); which included aminoacyl-tRNA biosynthesis, glutathione metabolism, biosynthesis of alkaloids derived from histidine and purine, C₅-branched dibasic acid metabolism, biosynthesis of alkaloids

derived from ornithine, lysine and nicotinic acid, beta-alanine metabolism, biosynthesis of plant hormones, biosynthesis of phenylpropanoids, butanoate metabolism, galactose metabolism, cyanoamino acid metabolism and glyoxylate and dicarboxylate metabolism.

Discussion

Within the plant kingdom it is believed there are an estimated 200,000 metabolites (Fiehn, 2002), however species specificity, specialized organ, tissue and cellular distribution, spatiotemporal factors, and rapid metabolite turnover, combined with the difficulty to capture and detect the metabolites, severely limits the number of metabolites that can be detected in an organism at any one time. Employing multiple analytical approaches can increase the number of metabolites detected, but also simply optimizing the extraction protocol can influence the success of detection. In this study, following extraction optimization steps, we were able to identify a total of 722 metabolites in single cell-type EBC extracts consisting of 194 known and 528 unidentified molecular features. The amount of known metabolites detected in EBC extracts was significantly higher than previous reports of single cell-type global metabolite profiling of leaf trichomes, including glandular trichomes from *Solanum* (119 known compounds/LC-MS/3:3:2 propanol:acetonitrile:water) and

TABLE 2 | Pathways which are significantly enriched for the metabolites.

Metabolic pathways
Aminoacyl-tRNA biosynthesis
Glutathione metabolism
Biosynthesis of alkaloids derived from histidine and purine
C5-Branched dibasic acid metabolism
Biosynthesis of alkaloids derived from ornithine, lysine and nicotinic acid
beta-Alanine metabolism
Biosynthesis of plant hormones
Biosynthesis of phenylpropanoids
Butanoate metabolism
Galactose metabolism
Cyanoamino acid metabolism
Glyoxylate and dicarboxylate metabolism

Significant using adjusted *p*-value ≤ 0.05 . *Arabidopsis thaliana* was used as a background set.

simple trichomes from *Arabidopsis* (119 compounds/GC-TOF/ 1:2.5:1 chloroform:methanol:water) (Ebert et al., 2010; McDowell et al., 2011). This may be a result of selecting a more optimized extraction protocol, or related to cell-type specific metabolic diversity, but could also be due to improvements in bioinformatics pipelines for the chemical assignment of unknown ions in metabolome data.

In order to gain greater insight into the biochemical function and underlying biological roles of the EBC we studied the effect of salt treatment on the EBC metabolites. Statistical analysis detected significant changes in 37% of the total metabolites (Supplementary Table 2). Of the 194 known metabolites identified we were able to detect significant changes in 57 (30%) of these upon salt-treatment (Table 1), suggesting an extensive alteration in the metabolome in response to salinity. The most highly increased metabolite (more than 12 fold) was the non-protein amino acid pipercolic acid, shown to be a critical regulator of inducible plant immunity and resistance to bacterial pathogens (Návarová et al., 2012), but also known to accumulate upon salt stress in *Limonium vulgare* (Stewart and Larher, 1980). Pipercolic acid is an analog of proline, sharing a similar chemical structure (Ganapathy et al., 1983). Proline is another non-protein amino acid that showed high accumulation in salt stressed EBC (Table 1). In fact, more than 50% of the metabolites that show significant changes in the EBC, can be classified as compatible solutes and include sugars, sugar alcohols, protein and non-protein amino acids, and organic acids (Figures 4A,B, Table 1) (Slama et al., 2014). This data correlates well with gene ontology enrichment analysis of differentially expressed transcripts in the EBC, with GO term networks “small molecule metabolic process” and “carboxylic acid metabolic process”, significantly enriched in the salt-treated EBC transcripts as compared to the control (Oh et al., 2015). The accumulation of osmotically active non-toxic organic compounds in the cytosol increases the cellular osmotic potential to provide a balance between the cytoplasm and the vacuolar lumen, which, in *M. crystallinum* can accumulate up to 1 M Na⁺ (and Cl⁻) (Adams et al., 1992). Both

proline and pinnitol have been shown to increase previously in EBC from salt-treated plants (Paul and Cockburn, 1989; Adams et al., 1998), and RNAseq results suggested that salt treatment induces metabolic pathways that lead to synthesis, accumulation, transport and conversion of compatible solutes (Oh et al., 2015). Specifically, increases in pinitol can be linked to increases in transcripts encoding myo-inositol-1-phosphate synthase and myo-inositol O-methyltransferase 1, two key enzymes in the pathway leading to pinitol synthesis via ononitol, which were highly abundant and significantly upregulated in the EBC transcriptome (up to 170-fold) following salinity treatment (Oh et al., 2015). Expression of transcripts encoding two key proline biosynthesis enzymes, delta-1-pyrroline-5-carboxylate synthase and pyrroline-5-carboxylate reductase were also salt-induced in the EBC (Oh et al., 2015) and link changes in other metabolites, including ornithine and alpha ketoglutaric acid (Table 1) to proline accumulation (Figure 4C).

A high number of organic acids were identified as significantly altered in the EBC following salt-treatment (Table 1, Figure 4B). As well as their possible function as metabolically active solutes for osmotic adjustment, organic acid metabolism is of fundamental importance at the cellular level for several biochemical pathways, including as photosynthetic intermediates in CAM (Cushman and Bohnert, 1997), and for the formation of precursors for amino-acid biosynthesis (López-Bucio et al., 2000). It is likely that changes in abundance of organic acids reflects all these roles. One of the most highly altered organic acids was the C₄ dicarboxylic acid maleate (maleic acid, 9.41-fold increase), a trans-isomer of fumaric acid, it is considered an uncommon plant metabolite (Fiehn et al., 2000), and its biological role is not clear, other than as a possible citric acid cycle intermediate. It was shown to be negatively correlated with Na⁺ concentration in roots from salt-treated barley seedlings (Wu et al., 2013), and was a cold shock inducible metabolite identified in *Arabidopsis* shoots (Kaplan et al., 2004), as well as being identified in *Arabidopsis* and cotton seed trichomes (Ebert et al., 2010; Naoumkina et al., 2013); but these reports give little insight into its function in the plant.

Amino acids are the predominant form of solute accumulated by a phylogenetically diverse range of salt-tolerant organisms such as salt-tolerant bacteria, halophytes, marine invertebrates, and hagfishes (Yancey et al., 1982). However, only selected protein amino acids appear to be utilized. In EBC there was an overall net increase in amino acids (Table 1), with the highest increase (approx. 3-fold) observed for asparagine and valine. Both these amino acids have been reported to increase upon salt-treatment (Rabe, 1990; Martinelli et al., 2007; Nedjimi, 2011; Zhang et al., 2011) and are implicated as nitrogen stores during stress to maintain metabolic homeostasis (Mansour, 2000; Rabe, 1990).

Both ascorbic acid and its oxidized form dehydroascorbic (DHA) acid are significantly altered in salt-treated EBC, with ascorbate decreasing and dehydroascorbic increasing. In plants ascorbate is the most abundant water-soluble antioxidant (Smirnoff, 2000) and participates in cellular redox state homeostasis by maintaining ROS below toxic levels but also through temporal-spatial coordination of ROS for a multitude

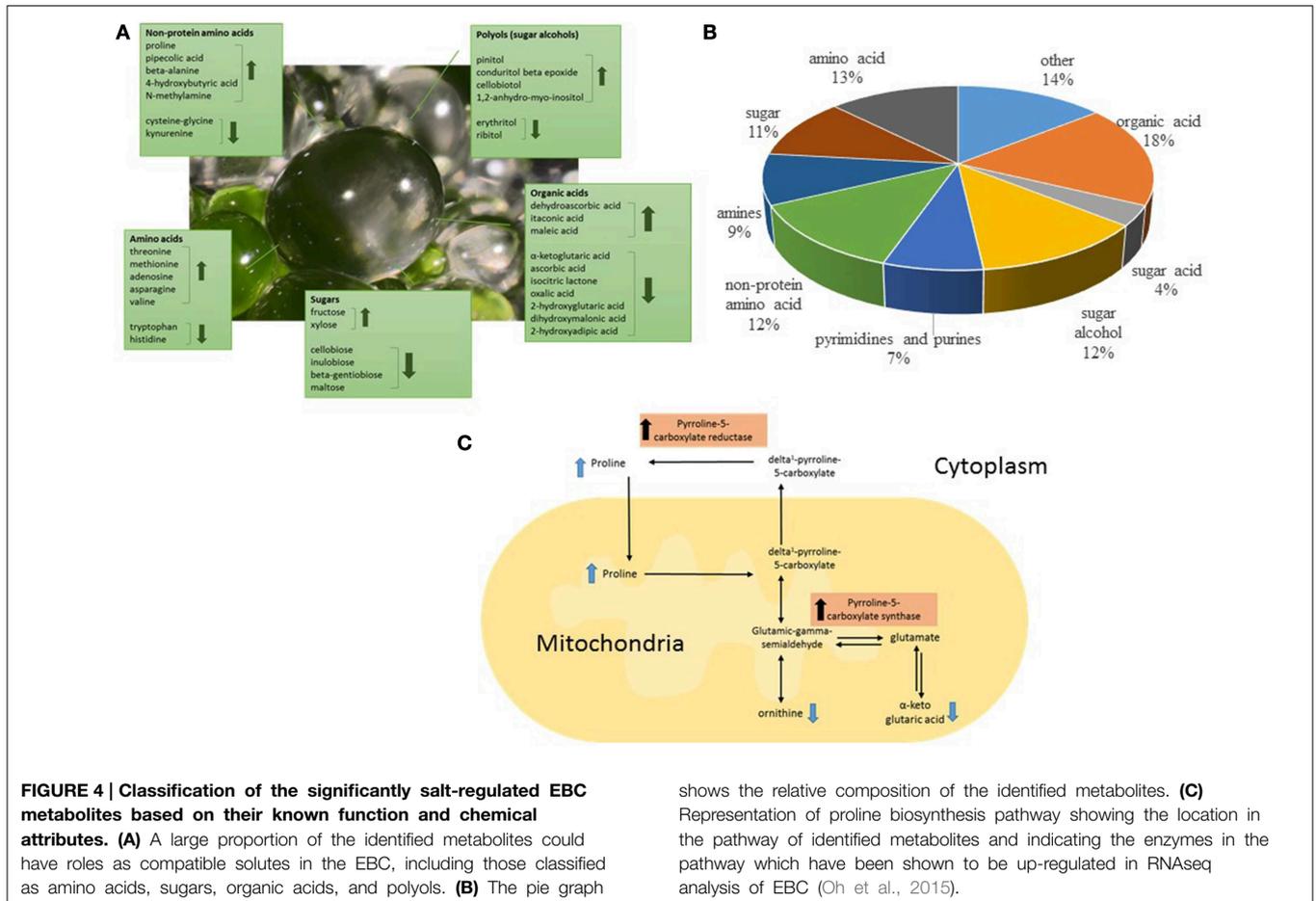


FIGURE 4 | Classification of the significantly salt-regulated EBC metabolites based on their known function and chemical attributes. (A) A large proportion of the identified metabolites could have roles as compatible solutes in the EBC, including those classified as amino acids, sugars, organic acids, and polyols. **(B)** The pie graph

shows the relative composition of the identified metabolites. **(C)** Representation of proline biosynthesis pathway showing the location in the pathway of identified metabolites and indicating the enzymes in the pathway which have been shown to be up-regulated in RNAseq analysis of EBC (Oh et al., 2015).

of signaling pathways (Baxter et al., 2013). Salinity stress results in the increased production of ROS (Abogadallah, 2010) and the recycling of ascorbic acid may be critical for maintaining ROS at a level within the EBC that minimizes oxidative damage but permits signaling function (Gallie, 2013). Ascorbic acid reacts with ROS and is oxidized to the radical MDHA which can rapidly disproportionate non-enzymatically to produce dehydroascorbic acid and recycled ascorbic acid (Smirnoff, 2000). The accumulation of dehydroascorbic acid in the EBC may reflect this recycling process. Oxidation of ascorbate has also been implicated in cell growth through the generation of MDHA radicals, and downstream stimulation of the plasma membrane proton ATPase which acidifies the extra cellular space resulting in cell wall loosening (Kato and Esaka, 2008).

Sinapyl alcohol, one of the main building blocks of lignin (Vanholme et al., 2010) is increased in the EBC from salt-treated plants. Lignin is deposited in the secondary cell walls of all vascular plants forming cross-links with other cell wall components to create rigidity. Increased lignification may be essential for the structural integrity of the swelling bladder cells to provide increased support as their volume increases with the accumulation of ions and water. Additionally, increases in the fatty alcohol dodeconol, a component of the waxy cuticular

film that is present on the aerial surface of plants (Kunst and Samuels, 2009) would help to protect the swelling bladder cells, limiting non-stomatal water loss, while increases in fatty acids, such as is observed for monopalmitin under salt-stress, would serve as precursors to fatty alcohol synthesis in the EBC (Kunst et al., 2006). Salt induced increases in transcripts for enzymes involved in biosynthesis of plant cuticular wax, including fatty acid hydrolase, were also observed in EBC transcriptome (Oh et al., 2015).

Comparison of global metabolic changes upon salt stress with other halophytes, including the legume *Lotus creticus* (Sanchez et al., 2011), and the salt-tolerant *Arabidopsis relative Thellungiella salsuginea* (Lugan et al., 2010), highlights a similarity in the classes of metabolites which show highest alteration upon salinity treatment, in particular, sugars, sugar alcohols, organic acids, and amino acids. This is not surprising as compatible solute synthesis is a well-documented tolerance response in most plants to salinity (Hasegawa et al., 2000), but there was little agreement in the specific metabolites that change, implying the existence of differential metabolic arrangements between species and possibly cell types, to compensate for ion imbalance (Flowers and Colmer, 2008).

Of the chemically unidentified metabolites, 21 showed changes of more than 5-fold in the salt-treated EBC samples

compared to the control samples. These could potentially make important contributions to the cell-type specific adaptations to salt in *M. crystallinum* EBC. Recent advances in linking functionally identified genes to unknown metabolites by genome wide association studies, integrating the data from genetic associations and metabolic networks with biochemical pathway information, has been successful in both humans (Krumstiek et al., 2012) and plants (Luo, 2015) for annotation of unidentified metabolites and will provide a means in the future to accelerate research in this area.

References

- Abogadallah, G. M. (2010). Antioxidative defense under salt stress. *Plant Signal. Behav.* 5, 369–374. doi: 10.4161/psb.5.4.10873
- Adams, P., Nelson, D. E., Yamada, S., Chmara, W., Jensen, R. G., Bohnert, H. J., et al. (1998). Growth and development of *Mesembryanthemum crystallinum* (Aizoaceae). *New Phytol.* 138, 171–190. doi: 10.1046/j.1469-8137.1998.00111.x
- Adams, P., Thomas, J. C., Vernon, D. M., Bohnert, H. J., and Jensen, R. G. (1992). Distinct cellular and organismic responses to salt stress. *Plant Cell Physiol.* 33, 1215–1223.
- Barkla, B. J., Vera-Estrella, R., Camacho-Emiterio, J., and Pantoja, O. (2002). Na⁺/H⁺ exchange in the halophyte *Mesembryanthemum crystallinum* is associated with cellular sites of Na⁺ storage. *Funct. Plant Biol.* 29, 1017–1024. doi: 10.1071/FP02045
- Barkla, B. J., Vera-Estrella, R., Hernández-Coronado, M., and Pantoja, O. (2009). Quantitative proteomics of the tonoplast reveals a role for glycolytic enzymes in salt tolerance. *Plant Cell* 21, 4044–4058. doi: 10.1105/tpc.109.069211
- Barkla, B. J., Vera-Estrella, R., and Pantoja, O. (2012). Protein profiling of epidermal bladder cells from the halophyte *Mesembryanthemum crystallinum*. *Proteomics* 12, 2862–2865. doi: 10.1002/pmic.201200152
- Baxter, A., Mittler, R., and Suzuki, N. (2013). ROS as key players in plant stress signalling. *J. Exp. Bot.* 65, 1229–1240. doi: 10.1093/jxb/ert375
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J. Royal Stat. Soc. Ser. B Methodol.* 57, 289–300.
- Chagoyen, M., and Pazos, F. (2011). MBRole, enrichment analysis of metabolomic data. *Bioinformatics* 27, 730–731. doi: 10.1093/bioinformatics/btr001
- Cushman, J. C., and Bohnert, H. J. (1997). Molecular genetics of crassulacean acid metabolism. *Plant Physiol.* 113, 667–676.
- Ebert, B., Zöller, D., Erban, A., Fehrlé, I., Hartmann, J., Niehl, A., et al. (2010). Metabolic profiling of *Arabidopsis thaliana* epidermal cells. *J. Exp. Bot.* 61, 1321–1335. doi: 10.1093/jxb/erq002
- Fiehn, O. (2002). Metabolomics - the link between genotypes and phenotypes. *Plant Mol. Biol.* 48, 155–171. doi: 10.1023/A:1013713905833
- Fiehn, O., Kopka, J., Trethewey, R. N., and Willmitzer, L. (2000). Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry. *Anal. Chem.* 72, 3573–3580. doi: 10.1021/ac991142i
- Fiehn, O., Wohlgemuth, G., and Scholz, M. (2005). Setup and annotation of metabolomic experiments by integrating biological and mass spectrometric metadata. *Data Integration Life Sci. Proc.* 3615, 224–239. doi: 10.1007/11530084_18
- Flowers, T. J., Colmer, T. D. (2008). Salinity tolerance in halophytes. *New Phytol.* 179, 945–963. doi: 10.1111/j.1469-8137.2008.02531.x
- Gallie, D. R. (2013). L-Ascorbic acid: a multifunctional molecule supporting plant growth and development. *Scientifica* 2013:795964. doi: 10.1155/2013/795964
- Ganapathy, V., Roesel, R. A., Howard, J. C., and Leibach, F. H. (1983). Interaction of proline, 5-oxoproline, and piperolic acid for renal transport in the rabbit. *J. Biol. Chem.* 258, 2266–2272.
- Hasegawa, P. M., Bressan, R. A., Zhu, J.-K., and Bohnert, H. J. (2000). Plant cellular and molecular responses to high salinity. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 51, 463–499. doi: 10.1146/annurev.arplant.51.1.463
- Hoagland, D. R., and Arnon, D. I. (1938). The water culture method for growing plants without soil. *Univ. Calif. Exp. Stn. Circ.* 347, 1–39.
- Kang, J.-H., Shi, F., Jones, A. D., Marks, M. D., and Howe, G. A. (2009). Distortion of trichome morphology by the hairless mutation of tomato affects leaf surface chemistry. *J. Exp. Bot.* 61, 1053–1064. doi: 10.1093/jxb/erp370
- Kaplan, F., Kopka, J., Haskell, D. W., Zhao, W., Schiller, K. C., Gatzke, N., et al. (2004). Exploring the temperature-stress metabolome of Arabidopsis. *Plant Physiol.* 136, 4159–4168. doi: 10.1104/pp.104.052142
- Kato, N., and Esaka, M. (2008). Changes in ascorbate oxidase gene expression and ascorbate levels in cell division and cell elongation in tobacco cells. *Physiol. Plant.* 105, 321–329. doi: 10.1034/j.1399-3054.1999.105218.x
- Klak, C., Khunou, A., Reeves, G., and Hedderson, T. (2003). A phylogenetic hypothesis for the Aizoaceae (Caryophyllales) based on four plastid DNA regions. *Am. J. Bot.* 90, 1433–1445. doi: 10.3732/ajb.90.10.1433
- Klaus, B., and Strimmer, K. (2012). *fdrtool, Estimation of (local) False Discovery Rates and Higher Criticism, R Package Version 1.2.10*. Available online at: <http://cran.r-project.org/web/packages/fdrtool/index.html>
- Krumstiek, J., Suhre, K., Evans, A. M., Mitchell, M. W., Mohney, R. P., Milburn, M. V., et al. (2012). Mining the unknown, a systems approach to metabolite identification combining genetic and metabolic information. *PLoS Genet.* 8:e1003005. doi: 10.1371/journal.pgen.1003005
- Kunst, L., Jetter, R., and Samuels, A. L. (2006). Biosynthesis and transport of plant cuticular waxes. *Annu. Plant Rev.* 23, 182–207. doi: 10.1002/9780470988718.ch5
- Kunst, L., and Samuels, L. (2009). Plant cuticles shine, advances in wax biosynthesis and export. *Curr. Opin. Plant Biol.* 12, 721–727. doi: 10.1016/j.pbi.2009.09.009
- Lee, D. Y., and Fiehn, O. (2008). High quality metabolomic data for *Chlamydomonas reinhardtii*. *Plant Methods* 4:7. doi: 10.1186/1746-4811-4-7
- López-Bucio, J., Nieto-Jacobo, M. F., Ramírez-Rodríguez, V., and Herrera-Estrella, L. (2000). Organic acid metabolism in plants, from adaptive physiology to transgenic varieties for cultivation in extreme soils. *Plant Sci.* 160, 1–13. doi: 10.1016/S0168-9452(00)00347-2
- Lugan, R., Niogret, M.-F., Lepout, L., Guégan, J.-P., Larher, F. R., Savouré, A., et al. (2010). Metabolome and water homeostasis analysis of *Thellungiella salsuginea* suggests that dehydration tolerance is a key response to osmotic stress in this halophyte. *Plant J.* 64, 215–229. doi: 10.1111/j.1365-313X.2010.04323.x
- Luo, J. (2015). Metabolite-based genome-wide association studies in plants. *Curr. Opin. Plant Biol.* 24, 31–38. doi: 10.1016/j.pbi.2015.01.006
- Mansour, M. M. F. (2000). Nitrogen compounds and adaptation of plants to salinity stress. *Biol. Plantarum* 43, 491–500. doi: 10.1023/A:1002873531707
- Martinelli, T., Whittaker, A., Bochicchio, A., Vazzana, C., Suzuki, A., and Masclaux-Daubresse, C. (2007). Amino acid pattern and glutamate metabolism during dehydration stress in the 'Resurrection' plant *Sporobolus stapfianus*: a comparison between desiccation-sensitive and desiccation-tolerant leaves. *J. Exp. Bot.* 58, 3037–3046. doi: 10.1093/jxb/erm161
- McDowell, E. T., Kapteyn, J., Schmidt, A., Li, C., Kang, J.-H., Descour, A., et al. (2011). Comparative functional genomic analysis of *Solanum* glandular trichome types. *Plant Physiol.* 155, 524–539. doi: 10.1104/pp.110.167114

Acknowledgments

Funding for this work was provided in part by DGAPA IN202514 and CONACYT P-178232 to R.V.-E.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2015.00435/abstract>

- Misra, B. B., Assmann, S. M., and Chen, S. (2014). Plant single-cell and single-cell-type metabolomics. *Trends Plant Sci.* 19, 637–646. doi: 10.1016/j.tplants.2014.05.005
- Naoumkina, M., Hinchliffe, D. J., Turley, R. B., Bland, J. M., and Fang, D. D. (2013). Integrated metabolomics and genomics analysis provides new insights into the fiber elongation process in Ligon lintless-2 mutant cotton (*Gossypium hirsutum* L.). *BMC Genomics* 14:155. doi: 10.1186/1471-2164-14-155
- Návarová, H., Bernsdorff, F., Döring, A.-C., and Zeier, J. (2012). Pipecolic acid, an endogenous mediator of defense amplification and priming, is a critical regulator of inducible plant immunity. *Plant Cell* 24, 5123–5141. doi: 10.1105/tpc.112.103564
- Nedjimi, B. (2011). Is salinity tolerance related to osmolytes accumulation in *Lygeum spartum* L. seedlings? *J. Saudi Soc. Agric. Sci.* 10, 81–87. doi: 10.1016/j.jssas.2011.03.002
- Oh, D.-H., Barkla, B. J., Vera-Estrella, R., Pantoja, O., Lee, S.-Y., Bohnert, H. J., et al. (2015). Cell type-specific responses to salinity – the epidermal bladder cell transcriptome of *Mesembryanthemum crystallinum*. *New Phytol.* doi: 10.1111/nph.13414. [Epub ahead of print].
- Paul, M. J., and Cockburn, W. (1989). Pinitol, a compatible solute in *Mesembryanthemum crystallinum* L.? *J. Exp. Bot.* 40, 1093–1098. doi: 10.1093/jxb/40.10.1093
- Rabe, B. (1990). Stress physiology: the functional significance of the accumulation of nitrogen containing compounds. *J. Hortic. Sci.* 65, 231–243.
- Sanchez, D. H., Pieckenstein, F. L., Escaray, F., Erban, A., Kraemer, U., Udvardi, M. K., et al. (2011). Comparative ionomics and metabolomics in extremophile and glycophytic *Lotus* species under salt stress challenge the metabolic pre-adaptation hypothesis. *Plant Cell Environ.* 34, 605–617. doi: 10.1111/j.1365-3040.2010.02266.x
- Schillmiller, A. L., Last, R. L., and Pichersky, E. (2008). Harnessing plant trichome biochemistry for the production of useful compounds. *Plant J.* 54, 702–711. doi: 10.1111/j.1365-313X.2008.03432.x
- Smirnov, N. (2000). Ascorbic acid, metabolism and functions of a multifaceted molecule. *Curr. Opin. Plant Biol.* 3, 229–235. doi: 10.1016/S1369-5266(00)00069-8
- Slama, I., Abdelly, C., Bouchereau, A., Flowers, T., and Savouré, A. (2014). Diversity, distribution and roles of osmoprotective compounds accumulated in halophytes under abiotic stress. *Ann. Bot.* 115, 433–447. doi: 10.1093/aob/mcu239
- Stewart, G. R., and Larher, F. (1980). “Accumulation of amino acids and related compounds in relation to environmental stress,” in *The Biochemistry of Plants*, ed H. J. Mifflin (New York, NY: Academic Press), 609–635. doi: 10.1016/b978-0-12-675405-6.50023-1
- Sugimoto, M., Kawakami, M., Robert, M., Soga, T., and Tomita, T. (2012). Bioinformatics tools for mass spectroscopy-based metabolomic data processing and analysis. *Curr. Bioinform.* 7, 96–108. doi: 10.2174/157489312799304431
- Tambellini, N. P., Zarembeg, V., Turner, R. J., and Weljie, A. M. (2013). Evaluation of extraction protocols for simultaneous polar and non-polar yeast metabolite analysis using multivariate projection methods. *Metabolites* 3, 592–605. doi: 10.3390/metabo3030592
- Vanholme, R., Demedts, B., Morreel, K., Ralph, J., and Boerjan, W. (2010). Lignin biosynthesis and structure. *Plant Physiol.* 153, 895–905. doi: 10.1104/pp.110.155119
- Wagner, G. J., Wang, E., and Shepherd, R. W. (1994). New approaches for studying and exploiting an old protuberance, the plant trichome. *Ann. Bot.* 93, 3–11. doi: 10.1093/aob/mch011
- Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* 58, 236–244. doi: 10.1080/01621459.1963.10500845
- Wu, D., Shen, Q., Cai, S., Chen, Z.-H., Dai, F., and Zhang, G. (2013). Ionic responses and correlations between elements and metabolites under salt stress in wild and cultivated barley. *Plant Cell Physiol.* 54, 1976–1988. doi: 10.1093/pcp/pct134
- Yancey, P. H., Clark, M. E., Hand, S. C., Bowlus, R. D., and Somero, G. N. (1982). Living with water stress, evolution of osmolyte systems. *Science* 217, 1214–1222. doi: 10.1126/science.7112124
- Zhang, J., Zhang, Y., Du, Y., Chen, S., and Tang, H. (2011). Dynamic metabolomic responses of tobacco (*Nicotiana tabacum*) plants to salt stress. *J. Proteome. Res.* 10, 1904–1914. doi: 10.1021/pr101140n

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Barkla and Vera-Estrella. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read,
for greatest visibility



COLLABORATIVE PEER-REVIEW

Designed to be rigorous
– yet also collaborative,
fair and constructive



FAST PUBLICATION

Average 85 days from
submission to publication
(across all journals)



COPYRIGHT TO AUTHORS

No limit to article
distribution and re-use



TRANSPARENT

Editors and reviewers
acknowledged by name
on published articles



SUPPORT

By our Swiss-based
editorial team



IMPACT METRICS

Advanced metrics
track your article's impact



GLOBAL SPREAD

5'100'000+ monthly
article views
and downloads



LOOP RESEARCH NETWORK

Our network
increases readership
for your article

Frontiers

EPFL Innovation Park, Building I • 1015 Lausanne • Switzerland
Tel +41 21 510 17 00 • Fax +41 21 510 17 01 • info@frontiersin.org
www.frontiersin.org

Find us on

