Functional Proteomics of Arabidopsis thaliana Guard Cells Uncovers New Stomatal Signaling Pathways

Zhixin Zhao, Wei Zhang, Bruce A. Stanley, and Sarah M. Assmann

Biology Department, Penn State University, University Park, Pennsylvania 16802
Section of Research Resources, Pennsylvania State University College of Medicine, Hershey, Pennsylvania 17033

We isolated a total of $3 \times 10^8$ guard cell protoplasts from 22,000 Arabidopsis thaliana plants and identified 1734 unique proteins using three complementary proteomic methods: protein spot identification from broad and narrow pH range two-dimensional (2D) gels, and 2D liquid chromatography–matrix assisted laser desorption/ionization multidimensional protein identification technology. This extensive single-cell-type proteome includes 336 proteins not previously represented in transcriptome analyses of guard cells and 52 proteins classified as signaling proteins by Gene Ontology analysis, of which only two have been previously assessed in the context of guard cell function. THIOGLUCOSIDE GLUCOHYDROLASE1 (TGG1), a myrosinase that catalyzes the production of toxic isothiocyanates from glucosinolates, showed striking abundance in the guard cell proteome. tgg1 mutants were hyposensitive to abscisic acid (ABA) inhibition of guard cell inward K+ channels and stomatal opening, revealing that the glucosinolate-myrosinase system, previously identified as a defense against biotic invaders, is required for key ABA responses of guard cells. Our results also suggest a mechanism whereby exposure to abiotic stresses may enhance plant defense against subsequent biotic stressors and exemplify how enhanced knowledge of the signaling networks of a specific cell type can be gained by proteomics approaches.

INTRODUCTION

Multicellular organisms develop specialized cell types, each with unique functions with regard to its specific role in the organism. The importance of single-cell-type transcriptomics studies in elucidating the functions of specialized cell types is uncontested (Dinneny et al., 2008). Single-cell-type proteomics studies are also essential to unravel the functions of specialized cells, particularly for cell types, like the guard cell (GC), where essential responses to stimuli can occur within seconds (Assmann and Grantz, 1990) and thus are unlikely to be mediated by transcriptomic changes. However, there have been very few single-cell-type proteomics studies in either plant or metazoan systems to date, in part owing to the greater complexity of the proteome and the greater technical challenges of proteomic methodologies (Tyers and Mann, 2003).

The most common subjects for single-cell-type proteomic studies have been cultured mammalian cell lines, where material is not limiting for proteomic analyses (Schirle et al., 2003; Diks and Peppelenbosch, 2004). In addition, studies have been done on human red blood cells (Pasini et al., 2006) and mouse red blood cells (Pasini et al., 2008), where 593 and 668 proteins were identified. By contrast, for plant cells, only the proteomes of blood cells (Pasini et al., 2008), where 593 and 668 proteins were identified. By contrast, for plant cells, only the proteomes of

1 Address correspondence to sma3@psu.edu.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Sarah M. Assmann (sma3@psu.edu).

Open Access articles can be viewed online without a subscription. www.plantcell.org/cgi/doi/10.1105/tpc.108.063263
and pump activity can be attained by electrophysiological assays. Ions (Ca$^{2+}$, K$^+$, Cl$^-$, and NO$_3^-$), signaling elements (e.g., G-proteins, phospholipase C [PLC], phospholipase D [PLD], inositol 1,4,5-trisphosphate, nitric oxide [NO], and reactive oxygen species [ROS]), and plant hormones (e.g., abscisic acid [ABA], auxin, and ethylene) regulate stomatal movements. With the advent of systems biology techniques, a dynamic model for induction of stomatal closure by the drought and stress hormone, ABA, has been developed (Li et al., 2006).

However, despite these achievements, basic questions remain unanswered: Are there proteins or sets of proteins that are preferentially expressed in GCs? How does the GC proteome compare with that of other plant cell types, and, importantly, how can proteomics inform studies of GC signaling and reveal new functions and relationships? These questions are addressed by this study.

**RESULTS**

**Proteomic Methods and GC Proteins**

A major challenge in single-cell-type proteomics is obtaining a sufficient quantity of highly pure cells. For GCs, as for some other plant cell types (Birnbaum et al., 2003), such purity can only be achieved by isolation of protoplasts (Figure 1A). Previous studies have validated that GC protoplasts (GCPs) retain key physiological responses present in GC in situ, including responsiveness to environmental signals such as light, ABA, and CO$_2$. To obtain enough GCPs for our proteomic study, we adapted our preparation method (Pandey et al., 2002) by increasing the plant material from 100 to 300 leaves per isolation and performed a total of ~100 GCP isolations.

GC proteins were isolated from GCPs and subjected to three complementary proteomics methods (Figure 1B): broad pH range two-dimensional (2D) gels (BR; Figure 1C), narrow pH range 2D gels (NR; Figure 1D), and 2D liquid chromatography-matrix assisted laser desorption/ionization multidimensional protein identification technology (2D LC-MALDI MudPIT). Gel-based and gel-free methods both have advantages and drawbacks. Gel-based methods can simultaneously separate and visualize hundreds of proteins on a single gel but have poor resolving power for membrane and low abundance proteins (Roe and Griffin, 2006). Gel-free methods have greater dynamic range in peptide identification compared with gel-based methods; however, these methods are biased against identification of small proteins (Baerenfaller et al., 2008).

A total of 1712 unique proteins (see Supplemental Table 1 online) was identified from two independent biological samples of GCP subjected to LC-MALDI MudPIT using Mascot with global false discovery rate (FDR) < 0.015 and Protein Pilot with local FDR < 0.05 (see Methods). Protein identification acceptance criteria were confidence interval (CI) ≥ 98% for proteins identified from multiple peptides and CI ≥ 99.9% for proteins identified from a single peptide. In total, 1489 proteins were identified by multiple peptides (nonbold proteins in Supplemental Table 2 online), 19 proteins were identified from a single peptide but from two independent replicates (bold proteins in Supplemental Table 2 online), and 204 unique proteins were identified from single peptides and from only one MudPIT replicate (peptide sequences are provided in Supplemental Table 3 online). For any of the 204 singletons that appear in Tables 2 or 3 (discussed below), protein identity was verified by manual inspection of spectra (see Supplemental Figure 3 online for these spectra). There were 13 peptides that could be assigned to a small group of homologous proteins but could not be assigned unambiguously to one particular protein within that group (peptide sequences are provided in Supplemental Table 3 online).

Since no single proteomic method is able to provide a comprehensive picture of the proteome (Kleffmann et al., 2007), we also pursued protein identifications using two gel-based methods that can provide information complementary to gel-free methods such as LC-MALDI MudPIT. For these methods, stained spots were in-gel digested by trypsin, identified by MALDI-tandem time-of-flight (TOF/TOF), and analyzed with Mascot software (see Methods). In two independent biological replicates of BR 2D gels, 138 spots were consistently detected, and from these, 85 spots, representing 58 unique proteins, were accepted as identified, meeting a very stringent acceptance criterion of requiring CI > 99.5% for identification.

To identify a greater number of proteins, GCP proteins were first prefractionated into five fractions with pH ranges 3 to 4.6, 4.6 to 5.4, 5.4 to 6.2, 6.2 to 7.0, and 7.0 to 10.0 using an isoelectric focusing (IEF) fractionator. Each fraction was applied to a corresponding single pH 2D gel (see Methods). A total of 250 spots was consistently detected from two biological replicates, from which 120 spots were identified with CI > 99.5%; these spots represent 59 unique proteins. By far the largest protein spot on either BR or NR 2D gels was the myrosinase THIOGLUCOHYDROLASE1 (TGG1) (Figure 1C, spot 12, and Figure 1D, spot 3E12). TGG1 was also detected in 11 and 36 more spots on the BR and NR 2D gels, respectively (see Supplemental Tables 4 and 5 online). Fourteen and 19 proteins were identified in multiple spots on the BR and NR 2D gels, respectively, representing a total of 28 unique proteins (see Supplemental Table 5 online).

In total, 1734 unique Arabidopsis GC proteins were identified from the combined application of all methods (see Figure 1B for summary; see Supplemental Tables 1 and 4 online for all identified GC proteins and Supplemental Table 6 online for the proteins identified from any two of the three methods). Of these proteins, 336 proteins are not represented in previously reported GC transcriptomes (Leonhardt et al., 2004; Yang et al., 2008), illustrating the complementarity of proteomic and transcriptomic approaches.

**Global Bioinformatic Analyses of the GC Proteome**

Enrichment of Gene Ontology (GO) Biological Process (http://www.Arabidopsis.org/tools/bulk/go/index.jsp) categories in our GC proteome compared with the conceptually predicted whole Arabidopsis proteome was determined using the topGO package (Alexa et al., 2006; Baerenfaller et al., 2008). The numbers of proteins present in the published pavement epidermal cell and trichome proteomes are too low for valid topGO analysis, but it was possible to perform topGO analysis of the leaf proteome
Figure 1. Three Proteomic Methods Identified 1734 Unique Proteins of the Arabidopsis GC Proteome.
PLD proteins (Table 3). Of these, only two proteins, the phospholipase were present in our GC proteome (Table 3). In addition, functional proteome (see Supplemental Table 7 online). from roots, and one from seeds) were present in our GC nine of these proteins (three from flowers, three from siliques, two Baerenfaller et al. (2008) as biomarkers of specific organ types, tissues by mutant analysis (Table 2). Of the proteins identified by have unique roles in GC function, particularly since, to date, only in cultured plant cells (Rajjou et al., 2004; Dixon et al., 2005; Job et al., 2007; Carroll et al., 2008), and the other four (At1g50010, At3g23840, and At5g54190) were identified in studies on specific in the literature: four of the eight (At3g16080, At3g18040, At3g23840, and At5g54190) were identified in studies on specific organelles (Friso et al., 2004; Heazlewood et al., 2004; Philippar et al., 2007; Carroll et al., 2008), and the other four (At1g50010, At5g08690, AtMg01190, and At5g08670) were identified by proteomic analyses of certain mutants or certain treatments, or in cultured plant cells (Rajjou et al., 2004; Dixon et al., 2005; Job et al., 2005; Sorin et al., 2006). The remaining 53 proteins may have unique roles in GC function, particularly since, to date, only 10 of these proteins have been shown to function in any other tissues by mutant analysis (Table 2). Of the proteins identified by Baerenfaller et al. (2008) as biomarkers of specific organ types, nine of these proteins (three from flowers, three from siliques, two from roots, and one from seeds) were present in our GC proteome (see Supplemental Table 7 online).

Twelve proteins previously shown to play a role in GC function were present in our GC proteome (Table 3). In addition, functional classification by GO analysis showed that 52 proteins out of our identified GC proteome are predicted as signal transduction proteins (Table 3). Of these, only two proteins, the phospholipase PLDα1 (Mishra et al., 2006) and the calcium-dependent protein kinase CPK3/CDPK6 (Mori et al., 2006), have been previously studied in the context of GC function, where they have been shown to participate in ABA signaling.

Functional Analysis of One of the Most Abundant Proteins in GCs, the Myrosinase TGG1

The plant glucosinolate-myrosinase system is known as a defense system against bacteria, pathogens, and herbivores. When tissue is damaged (e.g., by insect chewing), glucosinolates are thought to be released from the vacuole and hydrolyzed by myrosinases into a variety of toxic small molecules, including thiocyanate, isothiocyanate, and nitrile, which are active against biotic intruders (Wittstock and Halkier, 2002; Barth and Jander, 2006).

Although our gel-based analyses contributed only incrementally to the total number of GC proteins identified, these analyses revealed a remarkable abundance in GC of the myrosinase, TGG1. TGG1 was by far the largest spot on either the BR (Figure 1C, spot 12) or the NR 2D gels (Figure 1D, spot 3E12) and was identified in multiple spots (see Supplemental Table 5 online; 12 spots on the BR and 37 spots on the NR gels), suggesting high abundance and multiple posttranslational modifications.

TGG1 protein was not identified in proteomic analyses of trichome and epidermal pavement cells (Wienkoop et al., 2004), consistent with previous reporter gene analysis demonstrating strong expression of the TGG1 gene in GC and no expression in other epidermal cell types (Husebye et al., 2002; Barth and Jander, 2006). However, the presence of the TGG1 protein in GC has not been previously shown, and no function of TGG1 in GC has been described previously. The other Arabidopsis myrosinase, TGG2, is not expressed in GC by reporter gene analysis (Barth and Jander, 2006). TGG1 and TGG2 are indistinguishable by the probes used on Affymetrix microarrays, and the only other demonstrated locus of expression of these two TGG genes is in plhemo idioblasts (Husebye et al., 2002; Barth and Jander, 2006). Different masses of trypsin-digested peptides and amino acid sequences are generated from TGG1 versus the related myrosinase, TGG2, and can be detected by mass spectrometry and tandem mass spectrometry, respectively. Thus, TGG1 and TGG2 can be clearly distinguished by proteomic methods. Unlike TGG1, TGG2 was not found in any of the gel-based studies and was identified only in one replicate of the LC-MALDI MudPIT method using the Protein Pilot search algorithm. While 37 unique peptides were identified from TGG1, only two unique peptides were identified from TGG2, plus one peptide was identified that is identical in TGG1 and TGG2. These results suggest a substantially lower abundance of TGG2 in GC compared with TGG1.

Figure 1. (continued).

(A) Images showing high-purity GCP preparations (∼100; Inset magnification ×400).
(B) A total of 1712, 58, and 59 unique proteins were identified from 2D LC-MALDI MudPIT, BR, and NR methods, respectively; 19 proteins were identified by all three methods. For each method, two independent biological samples were analyzed.
(C) A 2D gel image from the broad pH range method. The first dimension was run using a 24-cm, pH 3 to 10 IPG strip. In total, 138 protein spots were detected via Coomassie blue staining. Twelve spots were identified as TGG1. Identifications of numbered spots can be found in Supplemental Table 5 online.
(D) 2D gel images from the narrow pH range method. Proteins were first fractionated into five fractions, and each protein fraction was separated on a narrow pH range IPG strip. From B1 to B5, the pH ranges are 3 to 6, 4.5 to 5.5, 5.3 to 6.3, 6.1 to 7.1, and 6 to 10 respectively. Thirty-seven spots were identified as TGG1. Identifications of numbered spots can be found in Supplemental Table 5 online.
TGG1 is predicted to be a cytoplasmic protein by SubLoc (http://www.bioinfo.tsinghua.edu.cn/SubLoc/), a secreted protein by Target P (http://wolfpsort.org/), and a chloroplast protein by WoLF PSORT (http://wolfpsort.org/). TGG1 has been identified in proteomic studies of the chloroplast proteome (Kleffmann et al., 2004), vacuole proteome (Carter et al., 2004), and ribosome proteome (Giavalisco et al., 2005). This lack of consensus regarding the subcellular localization of TGG1 notwithstanding, previous studies by Jander and colleagues have clearly shown that degradation of glucosinolates is slower in tgg1 mutant leaves compared with the wild type (Barth and Jander, 2006).

If the primary role of myrosinases is the deterrence of herbivory, why would TGG1 expression be limited to GC and not extend throughout all cell types of the epidermal layer? The abundance of TGG1 in GCs suggested to us that TGG1 might have as yet undiscovered roles in GCs. Accordingly, we evaluated TGG1 functions in GCs using two independent tgg1 mutants, tgg1-1 and tgg1-3 (Barth and Jander, 2006; Ueda et al., 2006). tgg1-1 (SALK_130474) and tgg1-3 (SAIL_786_B08) are T-DNA insertional mutants. tgg1-3 has been shown to lack full-length TGG1 transcript by RT-PCR analysis (Barth and Jander, 2006). In in vitro assays, aboveground tissue homogenates from tgg1-1 and tgg1-3 mutants exhibit only ~5 to 8% of wild-type levels of myrosinase activity (Barth and Jander, 2006; Ueda et al., 2006). As previously reported (Barth and Jander, 2006), these tgg1 mutants showed no obvious whole-plant phenotypes or developmental defects.

Given the roles of myrosinases in plant–herbivore interactions, we first assessed the GC response of wild-type and tgg1 mutants to a uniform wounding stimulus (Bailey et al., 2005). We found that wounding induces stomatal closure (Figure 2). tgg1 GCs showed a moderate disruption of wound-induced stomatal closure (Figure 2). Wounding induces methyl jasmonate (MJ) accumulation in leaf tissue (Maffei et al., 2007), and MJ has been reported to promote stomatal closure (Suhita et al., 2004; Munemasa et al., 2007). Therefore, we next evaluated MJ regulation of stomatal movements in Columbia (Col) and tgg1 mutants; however, we could not find any consistent effects of MJ on these responses, even in wild-type plants (see Supplemental Figure 1 online). This difference from published reports may be due to different plant growth conditions. Suhita et al. (2004) grew plants hydroponically in sand with short days (8-h photoperiod) and 300 μmol m⁻²s⁻¹ light intensity, Munemasa et al. (2007) grew plants with long days (16-h photoperiod) and 8000 Lux (~160 μmol m⁻²s⁻¹) light intensity, and we grew plants in potting mix with an 8-h photoperiod and 110 μmol m⁻²s⁻¹ of light (see Supplemental Methods online for details). Indeed, the effect of MJ on stomatal closure differs among published studies, consistent with the supposition that this effect may be altered by plant growth conditions; for example, the extent of MJ promotion of stomatal closure in the Suhita et al. (2004) study (~3 μm) is greater than that reported by the Munemasa et al. (2007) study (<1 μm). It has also been reported that the sitiens mutant of tomato (Solanum lycopersicum), an ABA biosynthetic mutant, shows little MJ-induced decrease in transpiration unless ABA is exogenously supplied (Herde et al., 1997), suggesting an interplay between these two hormones in the regulation of stomatal aperture.

Because ABA regulation of stomatal movements is central to GC function, stomatal responses to ABA were evaluated in the tgg1 mutants. tgg1 mutant stomata showed wild-type responses to ABA promotion of stomatal closure (Figure 3A) but were hyposensitive to ABA inhibition of stomatal opening (Figure 3B). These results imply that TGG1 is a positive regulator in ABA inhibition of stomatal opening and thus has a heretofore unrecognized role in plant abiotic stress responses.

In wild-type Col plants, ABA is known to inhibit the GC K⁺ in channels that mediate K⁺ uptake during stomatal opening (Schroeder et al., 1987; Schwartz et al., 1994; Wang et al.,

<table>
<thead>
<tr>
<th>Rank</th>
<th>GO ID</th>
<th>Term</th>
<th>Annotated</th>
<th>Significant</th>
<th>Expected</th>
<th>Log10 (1/P Value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GO:0009409</td>
<td>Response to cold</td>
<td>202</td>
<td>64</td>
<td>15.7</td>
<td>23</td>
</tr>
<tr>
<td>2</td>
<td>GO:0006412</td>
<td>Translation</td>
<td>561</td>
<td>123</td>
<td>43.7</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>GO:0006096</td>
<td>Glycolysis</td>
<td>56</td>
<td>28</td>
<td>4.4</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>GO:0015979</td>
<td>Photosynthesis</td>
<td>133</td>
<td>56</td>
<td>10.4</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>GO:0019684</td>
<td>Photosynthesis, light reaction</td>
<td>69</td>
<td>28</td>
<td>5.4</td>
<td>14</td>
</tr>
<tr>
<td>6</td>
<td>GO:0006457</td>
<td>Protein folding</td>
<td>228</td>
<td>52</td>
<td>17.8</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>GO:0006633</td>
<td>Fatty acid biosynthetic process</td>
<td>105</td>
<td>31</td>
<td>8.2</td>
<td>11</td>
</tr>
<tr>
<td>8</td>
<td>GO:0006520</td>
<td>Amino acid metabolic process</td>
<td>321</td>
<td>86</td>
<td>25.2</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 1. Comparison of topGO Analyses of Our GC Proteome and the Previously Identified Leaf Proteome
### Table 2. Fifty-Three Proteins Were Identified in Our GC Proteome and Not in Other Proteomes

#### Ten Proteins Have Been Shown to Function in Other Tissues by Mutant Analysis

<table>
<thead>
<tr>
<th>Accession Number</th>
<th>Protein Name</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>At1g07780</td>
<td>PAI1</td>
<td>Melquist and Bender (2004)</td>
</tr>
<tr>
<td>At1g48270</td>
<td>GCR1</td>
<td>Pandey and Assmann (2004)</td>
</tr>
<tr>
<td>At1g63680</td>
<td>A FMO family protein</td>
<td>Hansen et al. (2007)</td>
</tr>
<tr>
<td>At2g29980</td>
<td>FAD3</td>
<td>Browse et al. (1993)</td>
</tr>
<tr>
<td>At2g46800</td>
<td>At-MTP1</td>
<td>Desbrosses-Forrouge et al. (2005)</td>
</tr>
<tr>
<td>At3g52590</td>
<td>UBQ1</td>
<td>Johnson et al. (2004)</td>
</tr>
<tr>
<td>At3g54560</td>
<td>HTA11</td>
<td>March-Diaz et al. (2008)</td>
</tr>
<tr>
<td>At3g55830</td>
<td>EPC1</td>
<td>Bown et al. (2007)</td>
</tr>
<tr>
<td>At5g04110</td>
<td>DNA topoisomerase II family protein</td>
<td>Wall et al. (2004)</td>
</tr>
<tr>
<td>At5g59820</td>
<td>RHL41</td>
<td>Davletova et al. (2005)</td>
</tr>
</tbody>
</table>

#### Forty-Three Proteins Have Not Been Studied

<table>
<thead>
<tr>
<th>Accession Number</th>
<th>Protein Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>At1g06030</td>
<td>pfkB-type carbohydrate kinase family protein</td>
</tr>
<tr>
<td>At1g06630</td>
<td>F-box family protein</td>
</tr>
<tr>
<td>At1g07350</td>
<td>Transformer Ser/Arg-rich ribonucleoprotein, putative</td>
</tr>
<tr>
<td>At1g08135</td>
<td>CHX6B</td>
</tr>
<tr>
<td>At1g10770</td>
<td>Invertase/pectin methylesterase inhibitor family protein</td>
</tr>
<tr>
<td>At1g18140</td>
<td>WAKL3, similar to wall-associated kinase protein family</td>
</tr>
<tr>
<td>At1g21540</td>
<td>AMP binding protein, putative</td>
</tr>
<tr>
<td>At1g32120</td>
<td>Unknown protein</td>
</tr>
<tr>
<td>At1g50700</td>
<td>CPK33 (calcium-dependent protein kinase 33)</td>
</tr>
<tr>
<td>At1g51490</td>
<td>Glycosyl hydrolase family 1 protein</td>
</tr>
<tr>
<td>At1g53290</td>
<td>Galactosyltransferase family protein</td>
</tr>
<tr>
<td>At1g57620</td>
<td>Emp24/gp25L/p24 family protein</td>
</tr>
<tr>
<td>At1g63310</td>
<td>Unknown protein</td>
</tr>
<tr>
<td>At1g67580</td>
<td>Protein kinase family protein</td>
</tr>
<tr>
<td>At1g73670</td>
<td>At-MPK15</td>
</tr>
<tr>
<td>At2g02540</td>
<td>At-HB21/ZFHD4</td>
</tr>
<tr>
<td>At2g46320</td>
<td>Mitochondrial substrate carrier family protein</td>
</tr>
<tr>
<td>At3g05030</td>
<td>NHX2</td>
</tr>
<tr>
<td>At3g06410</td>
<td>Nucleic acid binding</td>
</tr>
<tr>
<td>At3g17070</td>
<td>Peroxidase, putative</td>
</tr>
<tr>
<td>At3g26430</td>
<td>GDSL-motif lipase/hydrolase family protein</td>
</tr>
<tr>
<td>At3g26740</td>
<td>CCL (CCR-LIKE)</td>
</tr>
<tr>
<td>At3g28530</td>
<td>UDP-glucose 4-epimerase</td>
</tr>
<tr>
<td>At3g42170</td>
<td>Transposase-like gene</td>
</tr>
<tr>
<td>At3g42950</td>
<td>Glycosid hydrolase family 28 protein</td>
</tr>
<tr>
<td>At4g02420</td>
<td>Lectin protein kinase, putative</td>
</tr>
<tr>
<td>At4g05590</td>
<td>Unknown protein</td>
</tr>
<tr>
<td>At4g11660</td>
<td>At-HSFB2B</td>
</tr>
<tr>
<td>At4g16590</td>
<td>At-CSLA01</td>
</tr>
<tr>
<td>At4g17100</td>
<td>Ser protease protein-related</td>
</tr>
<tr>
<td>At4g23900</td>
<td>NDK4</td>
</tr>
<tr>
<td>At4g28650</td>
<td>Leu-rich repeat transmembrane protein kinase, putative</td>
</tr>
<tr>
<td>At4g31130</td>
<td>Unknown protein</td>
</tr>
<tr>
<td>At5g15210</td>
<td>ZFHD3</td>
</tr>
<tr>
<td>At5g50230</td>
<td>Nucleotide binding</td>
</tr>
<tr>
<td>At5g53320</td>
<td>Leu-rich repeat transmembrane protein kinase, putative</td>
</tr>
<tr>
<td>At5g55040</td>
<td>DNA binding bromodomain-containing protein</td>
</tr>
<tr>
<td>At5g58340</td>
<td>DNA binding</td>
</tr>
<tr>
<td>At5g60720</td>
<td>Unknown protein</td>
</tr>
<tr>
<td>At5g62020</td>
<td>At-HSFB2A</td>
</tr>
<tr>
<td>AtCg00860</td>
<td>Unknown protein</td>
</tr>
<tr>
<td>AtCg01000</td>
<td>Unknown protein</td>
</tr>
<tr>
<td>AtMg00480</td>
<td>Subunit 8 of the mitochondrial F0 ATP synthase complex</td>
</tr>
</tbody>
</table>

Proteins in italics were identified by a single peptide from one MudPIT replicate. Corresponding peptide sequences and spectra for these proteins are provided in Supplemental Table 3 and Supplemental Figure 3 online, respectively.
<table>
<thead>
<tr>
<th>AGI</th>
<th>Name</th>
<th>AGI</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>At1g11260</td>
<td>STP1</td>
<td>At3g45780</td>
<td>PHOT1</td>
</tr>
<tr>
<td>At1g37130</td>
<td>NIA2</td>
<td>At3g53720</td>
<td>CHX20</td>
</tr>
<tr>
<td>At1g48270</td>
<td>GCR1</td>
<td>At4g18290</td>
<td>KAT2</td>
</tr>
<tr>
<td>At2g18960</td>
<td>OST2</td>
<td>At4g23650</td>
<td>CPK3 or CDPK6</td>
</tr>
<tr>
<td>At2g21660</td>
<td>GRP7</td>
<td>At5g23060</td>
<td>CAS</td>
</tr>
<tr>
<td>At3g15730</td>
<td>PLDα1</td>
<td>At5g58140</td>
<td>PHOT2</td>
</tr>
</tbody>
</table>

**Twelve Proteins Previously Shown to Have a Role in GC Function Were Identified in Our GC Proteome**

**Fifty-Two Proteins in Our GC Proteome Are Predicted to Be Signaling Proteins by GO Software**

(Continued)
In the absence of ABA treatment, \textit{tgg1} mutant GC had similar K$_{\text{in}}$ current amplitudes and kinetics as Col; however, ABA inhibition of K$_{\text{in}}$ current was abolished in both independent \textit{tgg1} mutants (Figures 3C and 3D). Consistent with the fact that \textit{tgg1} mutants show a wild-type response in ABA promotion of stomatal closure, outward K$^+$ currents, which mediate K$^+$ efflux during stomatal closure, were statistically identical in Col versus \textit{tgg1} mutants either with or without ABA treatment (Figure 3C; see Supplemental Figure 2 online).

Alterations in the glucosinolate profiles of \textit{tgg} mutants have already been characterized at the whole-leaf level, with significant increases in aliphatic and indole glucosinolates primarily observed in \textit{tgg1 tgg2} double mutants (Barth and Jander, 2006). Since myrosinases hydrolyze glucosinolates, one key question that arises is, what is functioning in the ABA inhibition of K$_{\text{in}}$ channels: myrosinase itself, glucosinolates, or the hydrolyzed products of glucosinolates? As a first step toward addressing this question, glucosinolates, myrosinase, or a combination of glucosinolates and myrosinase was directly applied to the cytosol of Col and \textit{tgg1} mutant GCs via the patch pipette solution (Figures 3C and 3D). Whole-cell patch clamp data showed that glucosinolate administration resulted in inhibition of K$_{\text{in}}$ channels in Col GCs but not in GCs of \textit{tgg1} mutants, indicating that the glucosinolates themselves do not suffice to inhibit channel activity. By contrast, coadministration of glucosinolates and myrosinase resulted in a similar extent of inhibition of K$_{\text{in}}$ channels in both Col and \textit{tgg1} mutants, suggesting that it is the hydrolyzed products of glucosinolates that evoke ion channel inhibition. However, myrosinase addition alone had no effect on K$_{\text{in}}$ channels in either Col or \textit{tgg1} mutants, suggesting that, in the absence of an appropriate triggering event, glucosinolate substrates are not available for myrosinase action. Most importantly, even though myrosinase application alone had no effect on K$_{\text{in}}$ currents, application of myrosinase restored K$_{\text{in}}$ sensitivity to ABA in \textit{tgg1} GCs (Figures 3C and 3D). These results indicate that the hydrolysis of glucosinolates by myrosinases is induced in some manner by ABA in \textit{Arabidopsis} GCs (see Figure 4 and Discussion) and is a necessary component of ABA-mediated K$_{\text{in}}$ channel inhibition.

**DISCUSSION**

Proteomics, an important postgenomic approach, has been applied to many fields [e.g., identification of protein expression profile changes under stress conditions (Hashimoto and Komatsu, 2007), analysis of posttranslational modifications (Kwon et al., 2006), and determination of protein–protein interactions (Parrish et al., 2007)]. Quantitative proteomic methods are also emerging.
but such quantifications will have greatest correspondence to biologically significant cellular protein amounts in the context of single cell type proteomes, as opposed to mixed tissues or organs where the abundance of a given protein may vary greatly among the different cell types present and thus mask abundance, or changes in abundance, within any single cell type.

Although fava bean (Vicia faba) GCs were used as material for an in-gel kinase assay approximately a decade ago, leading to identification of a Ca$^{2+}$-independent ABA-activated protein kinase by mass spectrometry (Li and Assmann, 1996, 2000), a major bottleneck for the characterization of the GC proteome has been obtaining enough highly pure GC from a species with a
Comparison of the Three Proteomic Methods

In our study, the gel-free method (2D LC-MALDI MudPIT) yielded by far the largest number of protein identifications: 1712 proteins were identified by this method, and, of these, 1638 were not identified by either of the gel-based approaches. The two gel-based methods importantly allowed recognition of the high abundance of TGG1 in GC. In addition, gel-based methods provide more reliable inference of posttranslational modifications (see Supplemental Table 5 online). Moreover, 22 proteins of the GC proteome were exclusively identified by gel-based methods, indicating that gel-free methods cannot completely replace traditional gel-based methods (Lambert et al., 2005). Although we evaluated the 22 proteins for common characteristics (e.g., molecular weight, pI, predicted subcellular localization, and the predicted trypsin-digestion patterns of these protein sequences) that might have enhanced their detection by gel-based over gel-free methods, no such characteristics could be identified.

Figure 4. Speculative Model of Interactions between Glucosinolates, Myrosinase (TGG1), and K+ Channels in GC ABA Signaling.

Left and right GCs indicate events before and after, respectively, activation of signaling pathways. Abiotic stresses increase ABA delivery to GCs. ABA either repositions glucosinolates (G) from the vacuole to the cytosol (arrow 1) or enhances myrosinase (M) activity or substrate affinity (arrow 2). Inhibition of inward K+ channel (K+in) activity might result from channel modification by the resultant hydrolyzed products of glucosinolates (e.g., isothiocyanates [ITC]), as in mammalian cells (arrow 3) or might result from decreased ROS scavenging (arrow 4) by ascorbate (Asc). Hydrolyzed products of glucosinolates may diffuse through the stomatal pore (shaded arrows) and thus deter biotic invaders. Reported biotic (wound) induced increases in ABA (Schmelz et al., 2003) may also provide positive feedback to the glucosinate-myrosinase system.

sequence

Comparison of the GC Proteome with Other Proteomes

Although we used GCPs as starting material, 29 of the identified GC proteins were identified by a previous cell wall proteomic study (Bayer et al., 2006): these proteins may localize to multiple subcellular compartments, including both apoplastic and symplastic destinations, or be present in secretory vesicles that have not yet fused with the plasma membrane (Lee et al., 2004). Indeed, further GO analysis predicted that 23 of the 29 (79%) proteins also localize to non–cell wall subcellular compartments.

Our topGO analysis revealed Biological Processes that were enriched in the GC proteome (Table 1) relative to the entire predicted Arabidopsis proteome and to the documented leaf proteome (Lee et al., 2007). Four of the GC-enriched Biological Processes were also enriched in the published leaf proteome (Lee et al., 2007) and thus may typify leaf cell types in general but not GC in particular. The GO category “amino acid metabolic process” was also fairly highly ranked in leaves (rank = 10). More interesting are the remaining GC-enriched Biological Processes: glycolysis (rank 3 in GC but 14 in leaf), light reactions of photosynthesis (rank 5 in GC but 29 in leaf), and fatty acid biosynthesis (rank 7 in GC but 142 in leaf).

Stomatal movement is estimated to have a high energetic requirement (Assmann and Zeiger, 1987), consistent with the topGO prediction of the importance of glycolysis and photophosphorylation to this cell type. Indeed, biochemical assays have shown disproportionately high rates of photophosphorylation in GC relative to their very low rates of carbon fixation (Shimazaki et al., 2007), and red light–stimulated stomatal opening is known to be inhibited by photosynthetic inhibitors, such as DCMU (Schwartz and Zeiger, 1984). The enrichment in the GC proteome of proteins involved in fatty acid biosynthesis may reflect not only the importance of lipids to cuticle formation (Samuels et al., 2008), but also the importance of lipids and lipid metabolites as signaling entities in GC. For example, the guard cell–specific HIGH CARBON DIOXIDE (HIC) gene encodes a likely 3-keto acyl CoA synthase, involved in the synthesis of very-long-chain fatty acids, yet hic mutants also show a dramatic
CO2-dependent alteration in GC production (Gray et al., 2000). Other lipid-related molecules with demonstrated roles in GC related to ABA signaling include phosphatic acid (Jacob et al., 1999; Zhang et al., 2004), sphingosine-1-phosphate (Ng et al., 2001; Couso et al., 2003), and inositol phosphates (Lee et al., 1996; Jung et al., 2002; Hunt et al., 2003). Our topGO analysis of the GC proteome suggests that additional study of mutants in enzymes related to fatty acid synthesis may uncover GC-related phenotypes.

Of the 53 proteins that were identified in our GC proteome (Table 2) but not in other known proteomes, some may be specific to GCs and thus can be considered as candidate GC biomarkers. In this regard, it would be of particular interest to characterize the seven proteins of unknown function in this set (Table 2). Others may be more abundant in GCs than elsewhere; thus, we succeeded in identifying these proteins as part of the GC proteome while they were missed in other proteomic analyses (e.g., the likely G-protein coupled receptor, GCR1, which confers ABA hypersensitivity to both stomatal movements and root growth) (Pandey and Assmann, 2004). Finally, some proteins may be specifically present in GC plus a few other specialized cell types and thus missed in proteome analyses of whole organs. Conversely, our identification in the GC proteome of proteins previously thought to be biomarkers for specific organs (see Supplemental Table 7 online), including roots and seeds, which lack GC, indicates the importance of single-cell-type proteome analysis in determinations of protein distribution.

The GC Proteome and GC Signaling

We identified 67 proteins from the literature as previously shown to function in mature Arabidopsis GCs/GCPs (see Supplemental Table 8 online). Of these, 51 participate in GC responses to ABA, light, and/or CO2, while six of the remaining 16 proteins function in solute transport. Twelve of the 67 proteins were present in our GC proteome (Table 3); the other 55 may be low abundance proteins or induced under specific conditions. The 12 identified proteins are involved in light (PHOT1, PHOT2, and CHX20) and ABA (CPK3, GCR1, GRP7, OST2, NIA2, and PLDx1) signaling and in solute transport (CHX20, STP1, KAT2, and OST2) (Table 3).

GO analysis categorizes 52 proteins in our GC proteome as signal transduction proteins (Table 3). Of these, 50 have yet to be studied in the context of GC function, highlighting the usefulness of proteome analysis in identifying targets for further functional analyses. Thirteen are protein kinases, including one CDPK (CPK3/CDPK6), seven LRR protein kinases, and five MAP kinases; four of these kinases (At1g73670, At3g18040, At4g28650, and At5g53320) are also in the list of proteins only found to date in GCs (Table 2), while five of the 53 proteins are involved in auxin signaling, suggesting that auxin may be more important in GC physiology than previously recognized (Acharya and Assmann, 2008).

TGG1 Function in GCs

While many interesting candidates for further downstream analysis were revealed in our GC proteome analysis, we chose to perform an in-depth functional study of one protein. TGG1 was chosen due to its superabundance in GC: TGG1 comprises 40 to 50% of the total protein identified on 2D gels and was identified by >30 unique peptides (not shared with TGG2) in our MudPIT analysis. However, no roles for TGG1 in GCs had been previously demonstrated.

We examined the effect of wounding on stomatal apertures and found that wounding of the epidermis did stimulate a stomatal response and that TGG1 seemed to participate in this effect (Figure 2). However, TGG1 appeared to play a more essential role in GC ABA signaling: tgg1 mutant plants lacking this enzyme were unresponsive to ABA inhibition of stomatal opening and K+in channel regulation. Intracellular application of myrosinase alone did not restore K+in channel inhibition in tgg1 GCs or cause channel inhibition in wild-type GCs, yet application of myrosinase in the presence of ABA restored channel inhibition to the tgg1 mutants.

Our electrophysiological results are consistent with the following scenario (Figure 4): (1) myrosinase and its substrates, the glucosinolates, are localized in distinct subcellular compartments in GCs, as has been proposed for other cell types (Grubb and Abel, 2006). (2) ABA induces the relocation of glucosinolates to the cytosol (arrow 1 in Figure 4). (3) The hydrolysis of available glucosinolates by TGG1 leads to inhibition of K+in channels in GCs, and this is one component of ABA inhibition of stomatal opening. However, we note that there are also alternative interpretations that are consistent with our data (e.g., ABA signaling might somehow increase the activity of myrosinase or its affinity for its substrate) (arrow 2 in Figure 4).

ABA regulation of glucosinolate compartmentalization and thus availability for hydrolysis by myrosinases, as hypothesized here, could help to explain nondefensive developmental changes in glucosinolate concentrations (Petersen et al., 2002; Brown et al., 2003) that occur in the absence of the tissue disruption that brings substrate and enzyme together during herbivory. Such a phenomenon could also provide a mechanism whereby exposure to abiotic stress could strengthen defenses against subsequent biotic stressors. In addition, given the mechanisms described here, biotic (wound)-induced increases in ABA, as reported to occur in leaf tissue (Schmelz et al., 2003), might also provide positive feedback to the glucosinolate-myrosinase defense pathway (Figure 4), priming plant defense mechanisms against abiotic and biotic (Beckers and Conrath, 2007) stressors.

Alteration in membrane potential is a rapid response to wounding (Maffei et al., 2007), and our observations suggest that K+ channel regulation may contribute to these early electrical events. Since application of glucosinolates to wild-type GCs, or application of glucosinolates plus myrosinase to tgg1 mutant GCs, also inhibits the K+in channels, we infer that it is the reaction catalyzed by myrosinase (e.g., the hydrolyzed products of glucosinolates) that evoke channel inhibition. Because the inward K+ channels of GCs are, on a sequence homology basis, similar to metazoan Shaker channels (Pilot et al., 2003;
Pandey et al., 2007), animal Shaker K⁺ channels may also be targets for these plant secondary compounds, which are known to have both toxic and anticarcinogenic effects in mammals (Halkier and Gershenzon, 2006). Preharvest growth conditions, harvesting processes, and storage conditions all can affect plant glucosinolate concentrations (Johnson, 2002). The implication from our data that ABA may regulate glucosinolate sequestration and stability may suggest modifications to extant agronomic protocols, with implications for food quality.

Our results (Figures 3C and 3D) demonstrate an interconnection between the products of myrosinase activity and K⁺ channel regulation, but further research will required to determine the exact mechanistic basis of this response. It may be relevant that recent studies have shown that isothiocyanates alter activity of the mammalian pain-sensing TRP1A channels by covalent modification at Cys residues (Hinman et al., 2006). Alternatively, or in addition, the fact that myrosinase binds the antioxidant ascorbic acid and can catalyze formation of a condensation product of ascorbic acid with methylindole (Burmeister et al., 2000) may suggest that myrosinase activity promotes ABA signaling via decreasing the ability of the GC to scavenge ROS. ROS elevation is a key signal transduction element in ABA signaling (Pei et al., 2000; Kwak et al., 2003), and plants engineered for increased ascorbate levels exhibit decreased levels of ROS and decreased GC responsiveness to ABA (Chen and Gallie, 2004).

While it has recently been demonstrated that small molecules, such as NO and ROS, are shared between ABA and defense signaling pathways, including in GCs (Melotto et al., 2006; Ali et al., 2007; Zhang et al., 2008), enzymes of secondary metabolism, such as TGG1, have not previously been implicated in ABA signaling. The interconnection discovered here between ABA and the glucosinolate-based biotic defense mechanism suggests a mechanism whereby exposure to abiotic stresses may enhance plant defense against subsequent biotic invaders. One general property of the hydrolyzed products of glucosinolates is volatility (Yan and Chen, 2007). Speculatively, the localized and extremely high abundance of TGG1 in GC might facilitate evaporation of the hydrolyzed products from stomatal pores and thus maximize both deterrence of would-be herbivores and attraction of their parasites and predators (Bradburne and Mithen, 2000) as well as possibly initiate between- or within-plant defense signaling mechanisms (Baldwin et al., 2006; Frost et al., 2007).

In conclusion, the GC is a model system in plant cell biology, and the discovered GC proteome can be used to inform the reconstruction of GC signaling networks in silico (Li et al., 2006) and in planta. Assessment of candidate proteins identified by our GC proteomic analysis also has the potential to enhance our understanding of how plants interact with the local climate and biotic environment. In particular, demonstration of TGG1 involvement in ABA signaling demonstrates novel roles for this known enzyme and highlights an interplay between biotic and abiotic stress responses in plants. The strategy applied here, beginning with global protein identification within a single cell type and ending with discovery of novel signaling pathways by functional analysis of protein candidates identified from proteomic analyses, will be a powerful approach for future single cell type studies in both plants and metazoans.
Spot Cutting, Trypsin Digestion, and Spotting on MALDI Plates

All visible spots in both BR 2D gels were cut manually, and all spots in both NR 2D gels were cut using a spot-cutter (Bio-Rad EXQuest spot cutter). All spots were digested with trypsin (Promega Sequencing Grade) according to http://www.hmc.psu.edu/core/proteins_MassSpec/MassSpec/sampleprep.htm, desalted with SCX Ziptips (Millipore), and then spotted on MALDI plates. After the samples were dried, each spot was overlaid sampleprep.htm, desalted with SCX Ziptips (Millipore), and then spotted All spots were digested with trypsin (Promega Sequencing Grade) ac-

Mass Spectrometry and Data Analysis

All peptides were analyzed using a 4700 or 4800 proteomic analyzer MALDI-TOF/TOF tandem system (Applied Biosystems). Two different software packages were used: GPS Explorer (Applied Biosystems/MDS Sciex), using as an underlying search algorithm a locally installed copy of the Mascot software programs, version 2.1 (Matrix Science; http://www.matrixscience.com), or Protein Pilot software version 2.0 (Applied Bio-
systems/MDS Sciex), using the Paragon algorithm (Shilov et al., 2007) for searching and the ProFound algorithm for protein inference and grouping from tandem mass spectrometry (MS/MS) spectral/peptide data.

All MS and MS/MS data obtained from gel-based methods were analyzed using GPS Explorer (Applied Biosystems). Candidate protein IDs from individual gel spots were accepted if they had a GPS Explorer protein Cl > 99.5% (equivalent to a Mascot Score of P < 0.005). MS/MS data from 2D LC-MALDI MudPIT experiments were analyzed using both Mascot and Protein Pilot software version 2.0. For both algorithms, protein identification acceptance criteria were CI ≥ 98% (equal to a Protein Pilot unused score of 1.7) for proteins identified with multiple peptides and CI ≥ 99.9% for proteins detected from a single peptide, plus acceptable estimated FDRs (see Supplemental Methods online for de-
tails). In addition to the much more stringent CI requirements for accep-
tance of protein identifications that were identified solely from single

Stomatal Aperture Measurement
tgg1 mutants were generously provided by Georg Jander, Cornell Uni-

Electrophysiology

Arabidopsis GCP isolation and standard whole-cell K+ recording were as previously described (Wang et al., 2001; Coursoli et al., 2003). For ABA treatment, 50 μM ABA was added in basic solution for ≥1.5 h pretreat-

Accession Number

Sequence data from this article can be found in the Arabidopsis Genome Initiative or GenBank/EMBL databases under the following accession number: TGG1 (At1g26000, P37702).

Supplemental Data

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

Supplemental Table 1. Proteins Identified by Single Peptides from One MudPIT Experi-

Supplemental Table 2. Proteins Identified by Multiple Peptides and Proteins Identified by Single Peptides but from Both Replicates of the MudPIT Method.

Supplemental Table 3. Peptide Sequences for Proteins Identified by a Single Peptide and in One Replicate of the MudPIT Method.
Supplemental Table 4. Proteins Identified in the GC Proteome by Gel-Based Methods.

Supplemental Table 5. Twenty-Eight Proteins Were Identified in Multiple Spots from the Gel-Based Methods.

Supplemental Table 6. GC Proteins Identified by Any Two Proteome Methods.

Supplemental Table 7. Nine Previously Identified Organ Biomarker Proteins Were Identified in Our GC Proteome.

Supplemental Table 8. Sixty-Seven Proteins Demonstrated to Function in Mature Arabidopsis Guard Cells Based on Published Literature.

Supplemental Methods.

ACKNOWLEDGMENTS

We acknowledge the College of Medicine Mass Spectrometry and Proteomics Facility and the Proteomics and Mass Spectrometry Core Facility at Penn State University. We thank Dr. Georg Jander for tgg mutant seeds and Drs. Hong Ma, Daniel Jones, and Jian Li for comments on the manuscript. This work was supported by National Science Foundation Grants MCB-0209694 and 6-2066-01 and USDA Grant 2006-35100-17254 to S.M.A.

Received September 13, 2008; revised November 26, 2008; accepted December 15, 2008; published December 29, 2008.

REFERENCES


Shimazaki, K.I. (1988). Ribulosebisphosphate carboxylase activity and


