

## RESEARCH ARTICLE

# Transcriptomics assisted proteomic analysis of *Nicotiana occidentalis* infected by *Candidatus Phytoplasma mali* strain AT

Toni Luge<sup>1</sup>, Michael Kube<sup>2</sup>, Anja Freiwald<sup>1</sup>, David Meierhofer<sup>1</sup>, Erich Seemüller<sup>3</sup> and Sascha Sauer<sup>1</sup>

<sup>1</sup> Otto Warburg Laboratory, Max Planck Institute for Molecular Genetics, Berlin, Germany

<sup>2</sup> Division Phytomedicine, Department of Crop and Animal Sciences, Humboldt-Universität zu Berlin, Berlin, Germany

<sup>3</sup> Institute for Plant Protection in Fruit Crops and Viticulture, Julius Kühn Institute Federal Research Centre for Cultivated Plants, Dossenheim, Germany

Phytoplasmas are pathogenic bacteria within the class of *Mollicutes*, which are associated with more than 1000 plant diseases. In this study, we applied quantitative mass spectrometry to analyse affected pathways of the model plant tobacco (*Nicotiana occidentalis*) upon *Candidatus Phytoplasma mali* strain AT infection. Using tissue obtained from leaf midribs, 1466 plant-assigned proteins were identified. For 1019 of these proteins, we could reproducibly quantify the expression changes of infected versus noninfected plants, of which 157 proteins were up- and 173 proteins were downregulated. Differential expression took place in a number of pathways, among others strong downregulation of porphyrin and chlorophyll metabolism and upregulation of alpha-linolenic acid metabolism, which was consistent with observed increased levels of jasmonic acid, a key signal molecule of plant defence. Our data shed light on the molecular networks that are involved in defence of plants against phytoplasma infection and provide a resource for further studies.

Received: December 13, 2013

Revised: May 20, 2014

Accepted: June 4, 2014

## Keywords:

Apple proliferation / Microbiology / Plant bacteria interaction / Proteogenomics / Proteomics informed by transcriptomics / RNA-Seq



Additional supporting information may be found in the online version of this article at the publisher's web-site

## 1 Introduction

*Candidatus Phytoplasma* represents a monophyletic taxon within the *Mollicutes*. Phytoplasmas are insect-transmitted phytopathogenic bacteria that mainly reside in the sieve-cells of the plant phloem and cause diseases in many agriculturally relevant plants [1–3]. For example, *Candidatus Phytoplasma mali* is the causative agent of apple proliferation, an important

plant disease in Europe. Research on these phytopathogens was hampered for a long time due to the lack of cell-free cultivation, which was reached recently [4].

A key challenge for improving plant defence to phytopathogens consists in comprehensively identifying the genes and proteins and molecular networks that are involved in the interaction of plants and bacteria [3]. Proteomics has been successfully applied to gain insight into the plant defence to pathogens. Many of the earlier studies in the field relied on 2DE to quantify protein expression and MALDI MS for protein identification [5–8]. But in these studies only few, mostly highly abundant proteins could be quantified. Recently, the introduction of labeling or label-free methods combined with nano-LC ESI-MS has paved the way for comprehensive proteomic analysis. For example, using a label-free

**Correspondence:** Dr. Sascha Sauer, Otto Warburg Laboratory, Max Planck Institute for Molecular Genetics, Ihnestr. 63-73, 14195 Berlin, Germany

**E-mail:** sauer@molgen.mpg.de

**Abbreviations:** *Ca. P. mali*, *Candidatus Phytoplasma mali*; EC, enzyme commission

quantitative shot-gun MS workflow, Monavarfeshani et al. could detect protein expression changes in Mexican lime tree infected with *Candidatus* Phytoplasma aurantifolia [9]. Label-free methods may allow higher throughput and inclusion of more replicates in a study. But on the other hand these procedures are rather difficult to control, less precise in terms of quantification and are in particular vulnerable for the detection of medium- to low-abundant proteins [10].

In this study, we focussed on the protein expression changes caused by *Ca. P. mali* strain AT infection on the established model plant tobacco (*Nicotiana occidentalis*). We applied dimethyl labeling [11, 12] for quantitative MS analysis to systematically detect protein expression changes in the leaf veins of tobacco in response to infection. Using obtained comprehensive quantitative datasets, we extracted relevant pathways underlying the response of the plant to the bacterial infection. In addition, we validated these results by comparing protein with gene expression data, and by testing functionally pathway-relevant changes of metabolite levels.

## 2 Materials and methods

### 2.1 Plants and sample preparation

*Ca. P. mali* strain AT was transmitted from *Catharanthus roseus* (periwinkle) via dodder bridges to *N. occidentalis* [13,14]. Infected plants were used for grafting on seed grown *N. occidentalis* plants. Successful transmission was confirmed by PCR using P1/P7 primer pair [15, 16]. All experiments were performed in an insect-protected greenhouse. For sequencing, RNA from the leaf veins of two phytoplasma positive plants was extracted after a Solution D protocol [17]. DNA digestion (NucleoSpin RNA II-Kit, Machery & Nagel) and depletion of ribosomal RNAs (RiboMinus Plant Kit for RNA-Seq, Invitrogen) were performed according to the relevant vendor protocols. Protein lysates sampled from biological triplicates from phytoplasma positive and negative plants were used as pools in discovery proteomics. Briefly, leaf vein tissue was placed in meshbags (BIOREBA) and disrupted in lysis buffer (9 M Urea, 4% SDS, 100 mM Tris-HCl pH 8, 2.5 mM EDTA, 2.5 mM EGTA, 1 mM PMSF from Sigma Aldrich, 10  $\mu$ M Pepstatin A from Serva, 10  $\mu$ M Leupeptin from Serva, 1 $\times$  cComplete Protease Inhibitor Cocktail from Roche) by making use of the hand homogenizer (BIOREBA). For quantitative RT-PCR, targeted proteomics, and metabolite quantification leaf veins of four phytoplasma positive and four phytoplasma negative plants were excised, immediately frozen in liquid nitrogen and grounded to powder using a mortar and pestle. Aliquots were immediately used to isolate total RNA using QIAshredder spin columns (QIAGEN), the RNeasy Mini Kit (QIAGEN), and the RNase-free DNase-Set (QIAGEN), proteins by acetone precipitation overnight followed by washing and sonication in lysisbuffer on ice and metabolites as described in Pan et al. [18].

### 2.2 RNA-Seq and transcriptome database construction

For sequencing the following Kits were applied: TruSeq RNA Sample Preparation Kit, Single-Read Cluster Generation Kit v4 and TruSeq SBS Kit v5 (all Illumina). Barcoded libraries were sequenced in a single read multiplex run over 100 cycles on one lane with the Genome Analyzer IIx (Illumina). From the obtained reads a transcriptome database was constructed as described in Evans et al. [19]. Briefly the transcriptome has been de novo assembled using Trinity [20] with default parameters. From the Trinity-generated transcripts open reading frames (>200 nucleotides) from all six frames were predicted using the EMBOSS tool getorf [21]. This file was subjected to construct a target-decoy database for protein identification and quantification together with a multiple FASTA file extracted from NRPROT database, www.ncbi.nlm.nih.gov, containing all 479 protein sequences of *Ca. P. mali* strain AT (tax id 37692, from 01.02.2013).

### 2.3 Discovery proteomics

Equal amounts of the three protein extracts were pooled prior to reduction and alkylation with 50 mM DTT and 100 mM iodoacetamide and separated in duplicates by standard SDS-PAGE [22]. Gel bands were excised and cutted from high to low molecular weight into 16 equal sized slices. The in gel tryptic digest was performed as described in Shevchenko et al. [23] with minor modifications. Peptides were extracted from the gel fractions and for relative quantification purposes dimethyl-labeled [11] in a label swap strategy. Briefly, concentrated peptides were redissolved in 100 mM triethylammonium bicarbonate buffer (Sigma Aldrich) and marked with the light and intermediate dimethyl label as described in Boersema et al. [12]. Finally, the peptides were dissolved in 5% ACN, 2% formic acid and the corresponding fractions of phytoplasma positive and negative plants mixed. LC-MS/MS was performed in technical duplicates via nanoflow RPLC (Agilent) in line with a linear ion trap (LTQ)-Orbitrap XL mass spectrometer (Thermo Scientific) and a Top10-method. Detailed information can be found in the Supplementary Material and methods.

### 2.4 Targeted proteomics

Protein lysates from biological quadruplicates were used and combined in equal amounts by pool alterations. Thus four protein pools from phytoplasma positive and four from phytoplasma-negative plants, each one consisting of a different combination of biological triplicates, could be obtained. In this way, it was possible to retain the benefits of analyzing pooled samples, while getting access to the biological variance. After reduction and alkylation of disulfide bonds with 50 mM DTT and 100 mM iodoacetamide,

proteins were subjected to in solution double digest with Lys-C and trypsin (Roche) as described in the Supplementary Material and methods. Desalted peptides were resuspended in Britton & Robinson buffer, pH11 and separated into three fractions using strong anion exchange in a StageTip format [24]. Prior to MS, desalted peptides were resuspended in 5% ACN and 2% formic acid and spiked with tryptic BSA peptides to monitor between-run reproducibility. Label-free MRM quantification was performed in technical triplicates on an Dionex Ultimate3000 (Thermo Scientific) coupled via nano-ESI source to a QTRAP® 6500 LC-MS/MS System (AB SCIEX) by applying the multiple reaction monitoring-initiated detection and sequencing workflow. The MRM method comprised nine proteins, including BSA, with 1–11 peptides per protein and 3–5 transitions per peptide resulting in a total cycle time of 2.8 s. Detailed information about the MS setup and MRM method development can be found in the Supplementary Material and methods.

## 2.5 Quantitative RT-PCR

Samples were pooled by pool alterations as described in Section 2.4. The four RNA pools from phytoplasma-positive plants and the four from phytoplasma-negative plants were reversely transcribed into cDNA applying the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) with random primers. Primer sequences, designed using NCBI's Primer Blast on the Trinity-generated transcripts that belong to the chosen genes, are listed in Supporting Information Table 2. Quantitative PCR with cDNA of the RNA pools and the appropriate minus RT controls was carried out on the ABI Prism 7900HT Sequence Detection System using the SYBR Green PCR Master Mix (Applied Biosystems). Briefly, after an initial denaturation at 95°C for 10 min, the cDNA was amplified by 40 cycles of PCR (95°C, 15 s; 60°C, 60 s). The relative gene expression levels were normalized using GAPDH and quantified by the  $2^{-\Delta\Delta C_t}$  method [25]. Primer specificity has been proven by melting curve analysis and by verification of the correct product size after gel electrophoresis of the PCR products.

## 2.6 Jasmonic acid quantification by LC-MS

Plant metabolite extracts were analyzed by MRM on an Agilent 1290 Infinity HPLC System (Agilent) coupled via ESI source to a QTRAP® 5500 LC-MS/MS System (AB SCIEX) as described in Pan et al. [18] and the Supplementary Material and methods. After pooling and altering the pool composition as described in Section 2.4, the samples were analyzed in technical quadruplicates by monitoring the jasmonic acid specific transition from 209 to 59 *m/z*. All MRM data were processed using MultiQuant 2.1.1 (AB SCIEX) with the MQ4 algorithm for peak integration and default parameters. All

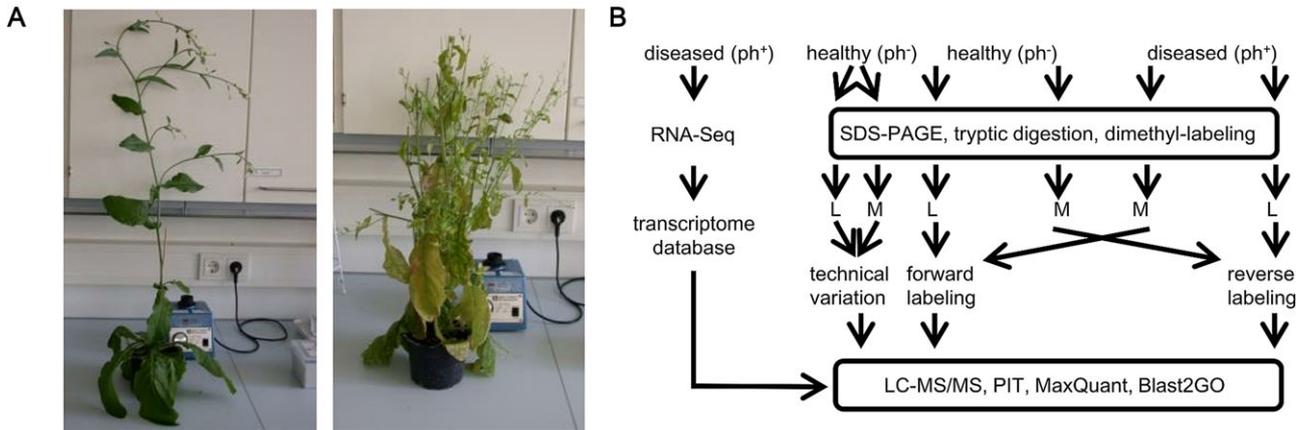
data were manually inspected to ensure correct peak detection and accurate integration.

## 2.7 Data analysis

Raw files from discovery proteomics were processed and analyzed using the MaxQuant computational proteomics platform [26] version 1.3.0.5 which makes use of the search engine Andromeda [27]. The peak list generated was searched against the self-build transcriptome database with a precursor mass tolerance of 6 ppm and a fragment mass tolerance of 0.5 Da. Tryptic peptides with minimum of seven amino-acid length and up to two missed cleavages were considered with both the peptide and protein FDR set to 1%. Only protein groups identified with at least two peptides, one of them unique, were included in the analysis. Quantification was based on only unmodified and unique peptides with at least two ratio counts. Common contaminants, reverse database hits, and protein groups only identified by site were excluded in further analysis steps performed in the R environment [28] and with Blast2GO (version 2.6.4) [29]. More detailed information can be found in the Supplementary Material and methods. Primary data analysis of targeted proteomics data was performed in Skyline version 2.1 [30]. Correct peak integration was checked manually based on sequence confirmation by enhanced product ion spectra with subsequent Mascot database search and retention time regression, that was calculated from spiked BSA peptides with the retention time predictor SRRCalc 3.0 (300 A) implemented in Skyline. Poor transitions have been removed and only those proteins were considered for relative quantification of which at least two proteotypic peptides with 3–5 transitions per peptide could be assayed. In this way, five of eight targeted proteins, beside BSA, were reproducibly quantified. Label-free, relative protein-level quantification, and testing for differential abundance were performed in the R environment [28] using a linear mixed-effects model as implemented in the MSstats package version 2.0.1 [31] with expanded scope on biological replication. Calculated *p*-values for each comparison were adjusted after Benjamini and Hochberg to control the FDR at a cutoff of 0.05.

## 3 Results and discussion

Seed-grown *N. occidentalis* plants were infected by graft-inoculation in the greenhouse. After three months, plants displayed the typical symptoms of infection such as decline, yellowing of plants, clustering of branches (witches' broom) and others (Fig. 1A). However, as could be observed by visual inspection symptom severity can vary considerably, and is much more difficult to control than for example in vitro cell culture experiments. Therefore, we have chosen a sample pooling approach as described in Section 2.



**Figure 1.** A large-scale quantitative proteomics experiment reveals altered expression levels of plant proteins upon phytoplasma infection. (A) *Ca. P. mali* strain AT infection causes phenotypical changes in *N. occidentalis* plants. Compared to healthy plants (phytoplasma negative,  $ph^-$ , shown left), infected plants (phytoplasma positive,  $ph^+$ , shown right) exhibited clear disease symptoms such as witches' broom and dwarfism. (B) The experimental design was arranged in three layers. RNA sequencing on two  $ph^+$  plants has been carried out to assemble a transcriptome database. Protein identification and quantification was based on this created database by applying the proteomics informed by transcriptomics (PIT) technique combined with de novo annotation of the sequences using Blast2GO. In the second step the robustness and technical variation of the applied stable isotope dimethyl-labeling technique with the dimethyl light (L) and intermediate (M) label was explored and threshold values were defined to indicate deregulated protein expression levels. The differential analysis of protein expression of infected versus uninfected plants was carried out using a label swap strategy.

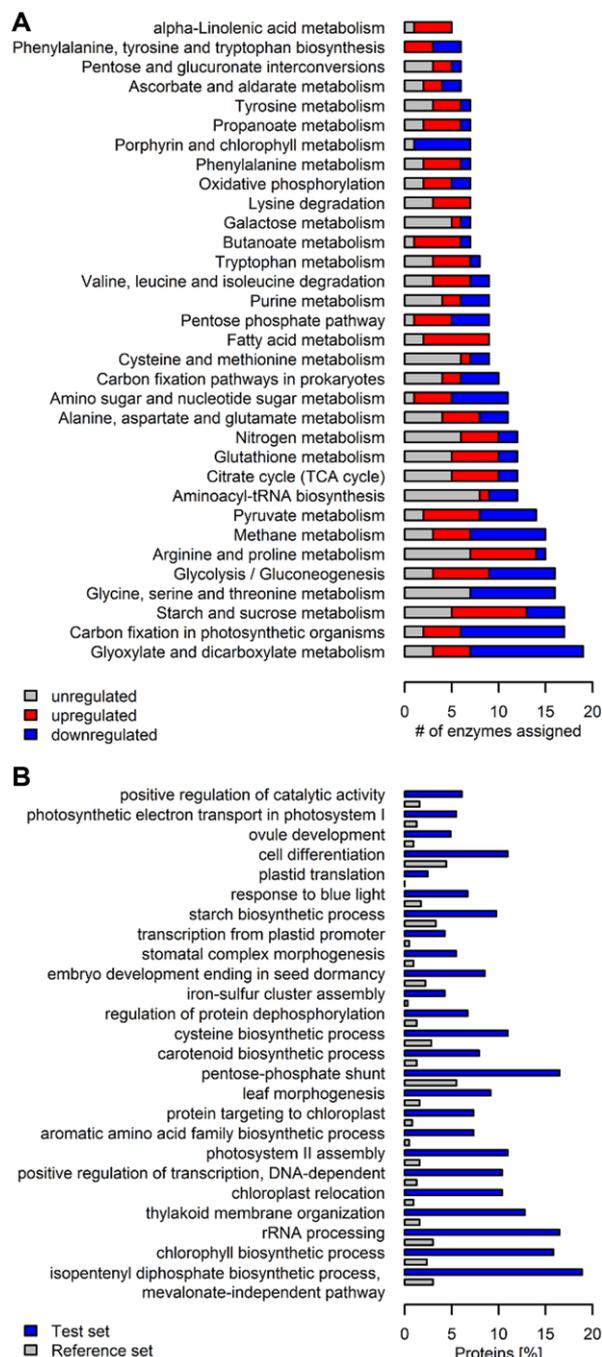
As for *N. occidentalis* almost no sequence and protein expression data were available, we performed RNA-sequencing (RNA-Seq) to establish a resource for constructing a target-decoy database used in protein identification and quantification as described by Evans et al. [19] (Fig. 1B).

To control the protein expression analysis based on dimethyl labeling, we first checked for technical variation of the method. Based on 442 identified proteins, we defined a threshold of  $\log_2(\text{ratio}) = 1.012$  for up- and for downregulated proteins (Fig. 1B and Supporting Information Fig. 1).

For the analysis of protein expression changes in phloem rich tissue we used forward and reverse labeling of proteins extracted from leaf veins of infected and uninfected plants. Using the MaxQuant software package [26, 27] we identified altogether 1466 proteins of which 1019 could be quantified in both labeling conditions in the infected versus the non-infected plants. To validate the forward-reverse labeling approach, we performed correlation analysis, which indicated good reproducibility of the experiments (Spearman's rank correlation coefficient  $\rho = 0.85$ ) (Supporting Information Fig. 2). For further analyses, we only considered those proteins that were consistent with the above-defined threshold criteria in both labeling conditions. As a result of the phytoplasma infection 157 proteins of *N. occidentalis* showed increased and 173 proteins decreased expression levels (Supporting Information Table 1).

Using our protein expression dataset and applying Blast2GO (www.blast2go.com) [29], we assigned proteins by their enzyme commission (EC) numbers to molecular KEGG pathways [32] potentially related to the plant response re-

sulting from the phytoplasma infection (Fig. 2A, Supporting Information Table 1). We observed a strong decrease in expression levels of proteins involved in anabolic pathways such as porphyrin and chlorophyll metabolism including the proteins protoporphyrin IX magnesium-chelatase, magnesium-protoporphyrin IX monomethyl ester (oxidative) cyclase, and NADPH-protochlorophyllide oxidoreductase. These data were consistent with the observed leaf yellowing. In particular, the expression of most of the proteins of the alpha-linolenic acid synthesis pathway such as allene oxide synthase, enoyl-CoA hydratase and 3-Ketoacyl-CoA-thiolase was increased (Fig. 2A and 3A). This finding is consistent with the trend of increased jasmonic acid levels in the leaf midribs, as detected by MRM (Fig. 3A). Furthermore gene expression analysis of seven exemplary genes of the porphyrin and chlorophyll metabolism and alpha-linolenic acid metabolism pathways on the RNA level widely supported the protein expression and metabolite data (Fig. 3B). Five out of seven selected genes and their corresponding proteins (protoporphyrin IX magnesium-chelatase, magnesium-protoporphyrin IX monomethyl ester (oxidative) cyclase, NADPH-protochlorophyllide oxidoreductase, allene oxide synthase, enoyl-CoA hydratase) showed similar trends in change of expression. Targeted, label-free quantification of proteins by MRM additionally confirmed discovery proteomics data. Expression level changes for five proteins could be detected, of which only the linoleate 13S-lipoxygenase (EC 1.13.11.12) was not concordant with the shotgun proteomic results (Fig. 3B). It was reported that apple trees infected with *Ca. P. mali* produce higher amounts of beta-caryophyllene

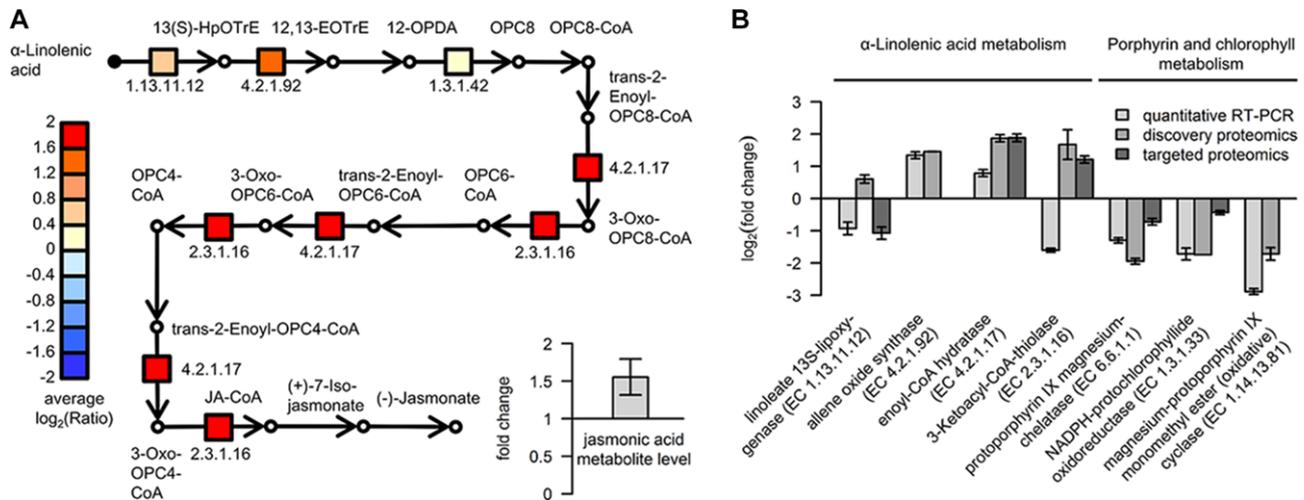


**Figure 2.** Analysis of altered pathways and biological processes. (A) Pathway mapping of quantified plant proteins illustrates deregulation in metabolism of *N. occidentalis* due to *Ca. P. mali* strain AT infection. Proteins were mapped on KEGG pathways by their enzyme commission (EC) numbers using Blast2GO. (B) Enriched GO-BP terms in the set of downregulated proteins. Test set (downregulated proteins, blue) was compared to all reproducibly quantified proteins (reference set, gray) using Fisher's exact test with multiple testing correction of FDR (Benjamini and Hochberg). Results are significant (FDR <0.05) and sorted from highest to lowest FDR. Bars show portions of the GO terms in the test compared to the reference set.

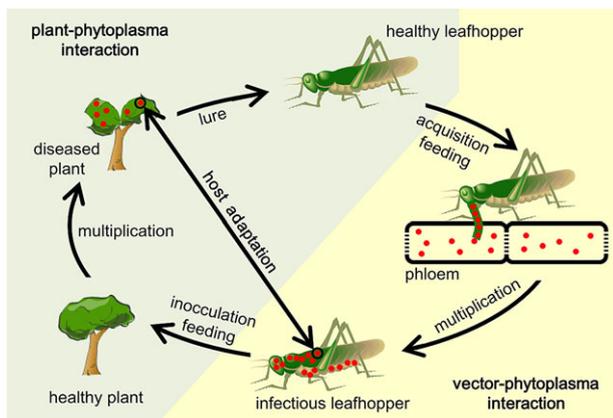
compared to uninfected plants, which may lure its insect vector *Cacopsylla picta* to the infected leaves and thereby support the spread of apple proliferation [33]. However, in tobacco we could not identify the underlying pathway due to an apparent lack of the beta-caryophyllene synthase (EC 4.2.3.57) in the generated transcriptomic data of *N. occidentalis*.

Alternatively to the pathway-centric view, gene ontology term enrichment analysis supported the observed results and further suggested striking downregulation of various transcriptional processes, protein targeting to chloroplast, photosystem II assembly, photosynthetic electron transport in photosystem I, as well as downregulation of starch biosynthetic process, cell differentiation and leaf morphogenesis (Fig. 2B). Among others, these data suggest that phytoplasma infection affected source-sink dynamics that may lead to phloem loading in the root tissue. This effect was consistent with the phenotype of rudimentary root systems, yellowing, dwarfism, and a general decline of phytoplasma infected plants. Moreover, phytoplasmas exhibit reduced metabolic capacities and rely on consuming nutrients from the phloem sap of the host plant [34]. Further research will elucidate if this nutrient depletion might be sufficient to alter the plant hosts metabolism or if phytoplasmas may rather directly intervene with plant processes by secreting effector proteins. For example, with the TENGU protein of *Candidatus Phytoplasma asteris* strain OY-M and the SAP11 protein of *Candidatus Phytoplasma asteris* strain AY-WB, two of such effectors were discovered, which induce disease-like symptoms when recombinantly expressed in model plants [35, 36]. A gene encoding a putative SAP11-like effector was recently identified by bioinformatical analysis in *Ca. P. mali* strain AT [37].

Recent proteomics studies discussed the benefits of analysing pooled versus individual samples, in particular for discovery approaches [38]. Interestingly, individual variations between pooled samples are usually averaged out. Consequently, even small differences in protein expression can be detected between groups, leading to higher numbers of identified differentially expressed proteins than using individual samples. However, only the measurement of multiple individual samples can reveal the underlying distributions and heterogeneity of protein expression. To counteract the high degree of biological heterogeneity between individuals and to narrow confidence intervals for a more precise estimation of population means that facilitates detection of statistical significance in particular for slightly regulated proteins, many more samples must be analysed than the usual “three biological replicates” known for conventional in vitro cell culture experiments. For example, even the data of the obviously regulated jasmonic acid levels would require an impractical high number of replicates. Especially in healthy plants the content of jasmonic acid varied strongly. Sample size determination indicated that more than 70 samples would be required to gain sufficient precision for differential analysis, which is unfeasible in practice due to logistic limitations (a simulated calculation is provided in Supporting Information Fig. 3).



**Figure 3.** Analysis of pathways related to alpha-Linolenic acid and chlorophyll metabolism. (A) Analysis of the alpha-Linolenic acid pathway indicates that expression levels of enzymes increase. Map displays selected steps from KEGG pathways map00592 alpha-linolenic acid metabolism and enzymes are given as EC numbers: 1.13.11.12 linoleate 13S-lipoxygenase, 4.2.1.92 allene oxide synthase, 1.3.1.42 12-oxophytodienoate reductase, 4.2.1.17 enoyl-CoA hydratase, 2.3.1.16 3-ketoacyl-CoA-thiolase. The amount of jasmonic acid, a phytohormone, and the endpoint metabolite of the pathway, was determined by LC-ESI-MS/MS in fresh leaf vein extracts of diseased and healthy *N. occidentalis* plants. Results were derived from biological and technical quadruplicates. Error bars indicate SEM. One-sided Welch's *t*-test for increase of jasmonic acid amount in leaves of diseased plants gives a *p*-value of 0.07, indicating upregulation of this plant hormone upon phytoplasma infection. (B) Discovery proteomics derived expression levels for selected proteins of the KEGG pathways map00592 alpha-linolenic acid metabolism and map00860 Porphyrin and chlorophyll metabolism were quantified by targeted proteomics and compared to the corresponding transcript levels as determined by quantitative RT-PCR. Biological quadruplicates were analyzed and the ratio is expressed as  $\log_2(\text{ph}^+/\text{ph}^-)$ . Error bars indicate SEM and all expression changes are significant (*p*-value of two-sided Student's *t*-test for PCR data and FDR for targeted proteomics data <0.05). For comparison the average  $\log_2(\text{ph}^+/\text{ph}^-)$ -protein ratios derived from forward and reverse labeling proteomics experiment with SEM are shown.



**Figure 4.** Simplified sketch of the host cycle of phytoplasmas. The focus of this study was on the interaction of plants with phytoplasma, whereas the roles of the vector insects need further investigation.

By integrating the list of identified regulated proteins into molecular pathways and biological processes we could counteract the problem of potential underlying biological variance in expression values of single proteins. This approach allowed us to build hypotheses of the response strategies of the plant to cope with phytoplasma infection. For example, the data sug-

gested pathways involved in upregulation of jasmonic acid metabolism or downregulation of chlorophyll production, which could be confirmed by quantitative RT-PCR, targeted proteomics, determination of the jasmonic acid metabolite levels and visual inspection of the tobacco leaves.

## 4 Concluding remarks

In this study, we focussed on the interaction of the model plant *N. occidentalis* with the phytopathogenic bacterium *Ca. P. mali* strain AT (Fig. 4). Our data show the adverse effects of phytoplasma infection on the plant metabolism and upregulation of the stress-response metabolite jasmonic acid.

Moreover, by embedding sample specific RNA-Seq data into the proteomic analysis the workflow applied in this study does not rely on a genomic sequence [19]. Therefore, this emerging approach facilitates to gain first insights into specific species to go beyond analyzing the few sequenced model systems.

Further studies may focus on the interaction of the vector insects with phytoplasmas to potentially add complementary data to provide a more detailed view on the infection cycle of these pathogens (Fig. 4). Such knowledge will enable strategies to restore or prevent phytoplasma-induced disease symptoms that are responsible for major economic losses in

agriculture [39]. These strategies may be of increasing importance when facing a warming world, where crop pests and pathogens spread into new territories [40].

Our results provide a detailed picture on the plant host response to infection by *Ca. P. mali*. To build and test new hypotheses the analyzed datasets, including raw data, transcriptomics derived search database, MaxQuant and Blast2GO result tables, as well as the MRM method for targeted quantification of the proteins, can be accessed via the PeptideAtlas database (<http://www.peptideatlas.org/>) using the dataset identifier PASS00266.

*Our work is supported by the German Ministry for Education and Research (BMBF, grant number 0315082, 01EA1303 to S.S.), the European Union (FP7/2007-2013, under grant agreement n° 262055 (ESGI) to S.S.), the German Research Foundation (through projects KU 2679/2-1 and BU 890/21-1) and the Max Planck Society. This study is part of the Ph.D. thesis of T.L.*

*The authors have declared no conflict of interest.*

## 5 References

- Christensen, N. M., Axelsen, K. B., Nicolaisen, M., Schulz, A., Phytoplasmas and their interactions with hosts. *Trends Plant Sci.* 2005, *10*, 526–535.
- Hogenhout, S. a, Oshima, K., Ammar, E.-D., Kakizawa, S., Namba, S., Phytoplasmas: bacteria that manipulate plants and insects. *Mol. Plant Pathol.* 2008, *9*, 403–423.
- Lee, I., Davis, R. E., Gundersen-Rindal, D. E., Phytoplasma: phytopathogenic mollicutes. *Annu. Rev. Microbiol.* 2000, *54*, 221–255.
- Contaldo, N., Bertaccini, A., Paltrinieri, S., Windsor, H. M., Windsor, G. D., Axenic culture of a plant pathogenic phytoplasmas. *Phytopathol. Mediterr.* 2012, *51*, 607–617.
- Ji, X., Gai, Y., Zheng, C., Mu, Z., Comparative proteomic analysis provides new insights into mulberry dwarf responses in mulberry (*Morus alba* L.). *Proteomics* 2009, *9*, 5328–5339.
- Margarita, P., Palmano, S., Response of the *Vitis vinifera* L. cv. “Nebbiolo” proteome to *Flavescence dorée* phytoplasma infection. *Proteomics* 2011, *11*, 212–224.
- Margarita, P., Abbà, S., Palmano, S., Novel aspects of grapevine response to phytoplasma infection investigated by a proteomic and phospho-proteomic approach with data integration into functional networks. *BMC Genomics* 2013, *14*, 38–52.
- Taheri, F., Nematzadeh, G., Zamharir, M. G., Nekouei, M. K. et al., Proteomic analysis of the Mexican lime tree response to *Candidatus* Phytoplasma aurantifolia infection. *Mol. Biosys.* 2011, *7*, 3028–3035.
- Monavarfeshani, A., Mirzaei, M., Sarhadi, E., Amirkhani, A. et al., Shotgun proteomic analysis of the Mexican lime tree infected with *Candidatus* Phytoplasma aurantifolia. *J. Proteome Res.* 2013, *12*, 785–795.
- Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., Kuster, B., Quantitative mass spectrometry in proteomics: a critical review. *Anal. Bioanal. Chem.* 2007, *389*, 1017–1031.
- Hsu, J., Huang, S., Chow, N., Chen, S., Stable-isotope dimethyl labeling for quantitative proteomics. *Anal. Chem.* 2003, *75*, 6843–6852.
- Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., Heck, A. J. R., Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat. Protoc.* 2009, *4*, 484–494.
- Marwitz, R., Petzold, H., Özel, M., Untersuchungen zur Übertragbarkeit des möglichen Erregers der Triebstucht des Apfels auf einen krautigen Wirt. *J. Phytopathol.* 1974, *81*, 85–91.
- Lauer, U., Seemüller, E., Physical map of the chromosome of the apple proliferation phytoplasma. *J. Bacteriol.* 2000, *182*, 1415–1418.
- Deng, S., Hiruki, C., Amplification of 16S rRNA genes from culturable and nonculturable Mollicutes. *J. Microbiol. Methods* 1991, *14*, 53–61.
- Schneider, B., Seemüller, E., Smart, C. D., Kirkpatrick, B. C., in: Razin S., Tully J. G. (Eds.), *Molecular and Diagnostic Procedures in Mycoplasmaology*, Academic Press, San Diego 1995, pp. 369–380.
- Chomczynski, P., Single-step method of RNA isolation by acid Guanidinium extraction. *Anal. Biochem.* 1987, *159*, 156–159.
- Pan, X., Welti, R., Wang, X., Quantitative analysis of major plant hormones in crude plant extracts by high-performance liquid chromatography-mass spectrometry. *Nat. Protoc.* 2010, *5*, 986–992.
- Evans, V. C., Barker, G., Heesom, K. J., Fan, J. et al., De novo derivation of proteomes from transcriptomes for transcript and protein identification. *Nat. Methods* 2012, *9*, 1207–1211.
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z. et al., Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 2011, *29*, 644–652.
- Rice, P., The European molecular biology open software suite EMBOSS: the European molecular biology open software suite. *Trends Genet.* 2000, *16*, 2–3.
- Laemmli, U. K., Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970, *227*, 680–685.
- Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V., Mann, M., In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* 2006, *1*, 2856–2860.
- Wiśniewski, J. R., Zougman, A., Mann, M., Combination of FASP and StageTip-based fractionation allows in-depth analysis of the hippocampal membrane proteome. *J. Proteome Res.* 2009, *8*, 5674–5680.
- Livak, K. J., Schmittgen, T. D., Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001, *25*, 402–408.
- Cox, J., Mann, M., MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 2008, *26*, 1367–1372.

- [27] Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A. et al., Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* 2011, *10*, 1794–1805.
- [28] R Development Core Team, R: a language and environment for statistical computing. *R Foundation Stat. Comput.* 2011, *1*, 409.
- [29] Conesa, A., Götz, S., García-Gómez, J. M., Terol, J. et al., Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 2005, *21*, 3674–3676.
- [30] MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M. et al., Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 2010, *26*, 966–968.
- [31] Clough, T., Thaminy, S., Ragg, S., Aebersold, R., Vitek, O., Statistical protein quantification and significance analysis in label-free LC-MS experiments with complex designs. *BMC Bioinformatics* 2012, *13*, 1–17.
- [32] Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M., Hirakawa, M., KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res.* 2010, *38*, D355–D360.
- [33] Mayer, C. J., Vilcinskas, A., Gross, J., Phytopathogen lures its insect vector by altering host plant odor. *J. Chem. Ecol.* 2008, *34*, 1045–1049.
- [34] Kube, M., Mitrovic, J., Duduk, B., Rabus, R., Seemüller, E., Current view on phytoplasma genomes and encoded metabolism. *Sci. World J.* 2012, *2012*, 1–25.
- [35] Hoshi, A., Oshima, K., Kakizawa, S., Ishii, Y. et al., A unique virulence factor for proliferation and dwarfism in plants identified from a phytopathogenic bacterium. *Proc. Natl. Acad. Sci. USA* 2009, *106*, 6416–6421.
- [36] Sugio, A., MacLean, A. M., Grieve, V. M., Hogenhout, S. A., Phytoplasma protein effector SAP11 enhances insect vector reproduction by manipulating plant development and defense hormone biosynthesis. *Proc. Natl. Acad. Sci. USA* 2011, *108*, 1–10.
- [37] Siewert, C., Luge, T., Duduk, B., Seemüller, E. et al., Analysis of expressed genes of the bacterium *Candidatus Phytoplasma mali* highlights key features of virulence and metabolism. *PLoS One* 2014, *9*, e94391.
- [38] Schröder, C., Srinivasan, H., Sill, M., Linseisen, J. et al., Plasma protein analysis of patients with different B-cell lymphomas using high-content antibody microarrays. *Proteomics Clin. Appl.* 2013, *7*, 802–12.
- [39] Strauss, E., Phytoplasma research begins to bloom. *Science* 2009, *325*, 388–390.
- [40] Bebbber, D. P., Ramotowski, M. A. T., Gurr, S. J., Crop pests and pathogens move polewards in a warming world. *Nat. Clim. Change* 2013, *3*, 985–988.