Review

Plant proteome analysis

Francisco M. Cánovas[1]
Eliane Dumas-Gaudot[2]
Ghislaine Recorbet[2]
Jesus Jorrin[3]
Hans-Peter Mock[4]
Michel Rossignol[5]

[1]Departamento de Biología Molecular y Bioquímica, Instituto Andaluz de Biotecnología, Unidad Asociada UMA-CSIC, Universidad de Málaga, Málaga, Spain
[2]UMR 1088 INRA/U, Bourgogne & FRE CNRS 2625, INRA/CMSE, PME, Dijon, France
[3]University of Cordoba, ETSIAM, Dept. Bioquim & Biol Mol, Cordoba, Spain
[4]Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany
[5]UR 1199, Protéomique, INRA, Montpellier, France

Proteome analysis is becoming a powerful tool in the functional characterization of plants. Due to the availability of vast nucleotide sequence information and based on the progress achieved in sensitive and rapid protein identification by mass spectrometry, proteome approaches open up new perspectives to analyze the complex functions of model plants and crop species at different levels. In this review, an overview is given on proteome studies performed to analyze whole plants or specific tissues with particular emphasis on important physiological processes such as germination. The chapter on subcellular proteome analysis of plants focuses on the progress achieved for plastids and mitochondria but also mentions the difficulties associated with membrane-bound proteins of these organelles. Separate chapters are dedicated to the challenging analysis of woody plants and to the use of proteome approaches to investigate the interaction of plants with pathogens or with symbiotic organisms. Limitations of current techniques and recent conceptual and technological perspectives for plant proteomics are briefly discussed in the final chapter.

Keywords: Plant proteome / Review / Symbiosis / Tree physiology / Two-dimensional gel electrophoresis
PRO 0602

Contents

Correspondence: PD Dr. Hans-Peter Mock, Institute of Plant Genetics and Crop Plant Research, Corrensstrasse 3, D-06466 Gatersleben, Germany
E-mail: mock@ipk-gatersleben.de
Fax: +49-39482-5139

Abbreviation: AM, arbuscular mycorrhizal

1 Introduction

The progress made in completing the sequencing of whole genomes and the rapid increase in the availability of expressed sequence tag (EST) sequences has opened new and exciting prospects for analyzing biological systems and their complex functions at different levels. In parallel to the accumulation of vast nucleic acid data, technological developments have permitted the establishment of systems for multiparallel analysis of transcript as well as of protein levels. The implementation of sensitive and rapid methods for protein identification and the continuous technical improvement of the so far largely descriptive analysis of protein patterns by two-dimensional gel electrophoresis (2-DE) have transformed the combination of both techniques into a powerful tool for

functional analysis now also more and more used in plant studies. Along with general limitations of the current available technologies, plant proteome approaches face specific challenges. Sample preparation is often more difficult due to the rigidity of plant cell walls or can be compromised by the accumulation of large quantities of secondary compounds in the central vacuole, which upon tissue disruption can lead to protein precipitation. Due to the availability of complete genomic sequence information and of large mutant collections, a number of recent proteome studies have focused on *Arabidopsis* as a model plant. With the completion of the rice genome and the progress of EST sequencing projects for other many plant species, the increased use of crop and other model plants can be predicted.

In this review, we summarize proteome approaches performed on the whole plant and tissue level (Section 2) and subcellular proteomic studies (Section 3). Special emphasis will be paid to woody plants (Section 4) which comprise many important crop species. The proteomic analysis of the biotic interactions of plants constitutes another challenging but rewarding field, with regard to symbiotic systems as well as to the interaction of plants with pathogens (Section 5). Finally, conceptual and technological perspectives for plant proteomics will be presented (Section 6). Different methodological approaches and techniques will be briefly mentioned making reference, when necessary, to key original papers or reviews. The reader is also referred to other recent reviews covering various aspects of plant proteomics [1–4].

2 Proteome analysis of plant organs and tissues

Proteomic analyses of plant organs or tissues were applied to monitor developmental changes or the influence of environmental stimuli on protein patterns but were also used to compare lines with different genetic backgrounds. A study of the maize leaf proteome based on 2-DE for protein separation resolved almost 900 spots when a pH 4–7 gradient was used for isoelectric focusing and 200 spots were found with gels covering the pH range from 6–11. Protein spot identification by mass spectrometry (MS) was based on protein and EST databases. Remarkably for some of the most abundant leaf proteins, for over 50% of the spots identified none, or only a hypothetical function, has been assigned [5].

Studies of the developmental processes occurring during seed germination were performed in *Arabidopsis* [6, 7]. Two-dimensional gel electrophoresis (2-DE) revealed about 1300 proteins in seeds, from which 74 showed modified abundance during germination. Specific changes in protein patterns were also observed at the stage of radicle protrusion and when applying a priming treatment intended to allow synchronous germination of a population of seeds. Several of the relevant proteins were identified by MS [6]. In a subsequent paper, the role of gibberellins (GAs) during seed germination was investigated by comparing protein patterns of a GA-deficient line, wild-type controls, and wild-type seeds treated with an inhibitor of GA biosynthesis during germination [7]. The analysis indicated that GA plays a very specific role during the initial stages of germination, as out of 46 changes observed for protein abundance only one was GA-dependent. Both studies underline the potential of the proteome approach for a better understanding of the complex cellular events during germination, a process of enormous agronomic impact. Grain filling and seed maturation are other important processes intensively studied in crop plants. Proteome analysis of barley seed development revealed one set of proteins present throughout this process, and other sets associated with early grain filling, with the later phase of desiccation, or showing transient accumulation during this developmental process. Identification of relevant spots demonstrated accumulation of low-M_r α-amylase/trypsin inhibitors, serine protease inhibitors, and of enzymes of the antioxidative defence system. The presence of several proteins during this process was experimentally demonstrated for the first time and cultivar-specific spot variations were monitored. The results demonstrate the capacity of proteome analysis to reveal new insights into the cellular mechanisms underlying seed development [8]. Complementary to the proteome studies on the tissue level, a sub-proteome analysis of the endoplasmic reticulum was performed for developing and germinating seeds of castor bean [9].

The second higher plant for which the complete genomic sequence is available is rice [10], which, as one of the most important crop plants holds great promise for functional genomics including proteome approaches as tools to detect novel traits for breeding. A survey of the proteome complement of rice root, leaf, and seed tissues was obtained by application of both 2-D gel electrophoresis and LC-based separation methods for complex peptide mixtures after digestion of protein extracts. Based on both methods, the identification of more than 2500 individual proteins was achieved [11]. Only a small fraction of the proteins (7.5%) were expressed in all three tissues analyzed, indicating tissue-specific regulation of metabolic pathways. A comparison of the contribution of the methods used to establish the protein inventory of all tissues demonstrated 2363 proteins were identified by LC-MS, whereas 556 proteins were identified using 2-DE. This result indicates the potential of non-gel based approaches in proteomic studies. Other rice proteome

studies were performed to identify embryo proteins [12] and to monitor the consequences of metal stress treatments [13].

In a number of studies, a proteome approach was undertaken to compare the protein patterns of mutants with the complement in wild-type lines. Investigation of developmental *Arabidopsis* mutants revealed that the quantity of an actin isoform correlated with the hypocotyl length [14]. An analysis of the consequences of iron deficiency in root tissue by 2-DE was performed using wild-type tomato and the *chloronerva* mutant affected in iron acquisition which revealed the induction of proteins involved in stress defence such as ascorbate peroxidase. Analysis of leaf tissue of the *chloronerva* mutant also showed a major reduction in the copper containing superoxide dismutase isoforms and in plastocyanin; the plastocyanin content could be restored by providing plants with additional Cu [15]. A proteome approach was also used to analyze remodelling of the photosynthetic apparatus in *Chlamydomonas* in response to Fe-deficiency [16]. The *cri1* mutant, accumulating cytokinins, was identified based on 2-DE patterns obtained from developmental mutants of *Arabidopsis* and statistical treatment of the data [17]. Analysis of the *Arabidopsis pasticcino* mutants by 2-DE revealed a considerable percentage of variable spots relative to wild-type controls; evaluation of responses to different hormone treatments indicated that the mutants were affected in cytokinin responses [18]. A mutant was also used to study cytokinin effects on chloroplast division in the moss *Physcomitrella* at the protein level [19]. Comparison of protein patterns in leaves of the late flowering *Arabidopsis* mutant *fy* and wild-type demonstrated qualitative differences, however further analysis of the F2 plants from crosses between wild-type and mutant showed no cosegregation between protein pattern differences and the late-flowering phenotype [20].

A 2-DE approach was used to monitor the target proteins of the *opaque 2* transcription factor in maize [21, 22]. A range of pairs of isogenic or near-isogenic lines was analyzed for modified protein expression during grain development. Thirty-six proteins differed in abundance between wild-type and o2 mutant throughout the different genetic backgrounds. A number of novel genes coding for enzymes in various metabolic pathways were thus identified not previously known to be controlled by Opaque-2 [22]. The analysis of modified protein expression included a thorough statistical treatment of the data [21, 22].

The resolving power of 2-DE has been widely used to assess genetic variability at the level of expressed proteins (see [23] for detailed discussion). Closely related lines were successfully differentiated in wheat cultivars, barley and rice lines, maize genotypes, and a number of other crop species. Analysis of maize lines demonstrated that the quantitative variability of proteins gave a pattern of relationships between genotypes [24]. Changes in protein patterns as a consequence of water deficit were investigated in two unrelated maize lines and their F1 hybrid [25]. Image analysis and statistical treatment of the data revealed that 40 proteins already visible in control plants showed increased amounts after water stress and 10 others were detectable only upon stress treatment. Another set of proteins was decreased in intensity. Comparison of the genotypes showed that several induced proteins were specific to one of the parental lines. Identification of selected spots revealed proteins already known to be associated with drought stress, but also others were identified such as caffeate *O*-methyltransferase detected only in stressed plants indicating increased lignification in response to the stress treatment [25]. These and many other studies have demonstrated the capacity of 2-DE to document genetic variability and to distinguish between lines and varieties, *e.g.*, when analyzing barley seed and malt [26, 27] or wheat grains [28]. Positional shifts of proteins were observed in the 2-D gel analysis of segregating families of maize, barley, pea, and maritime pine [29]. In maize, 42 pairs of proteins showed a 1:2:1 segregation in the F2 population indicative for a monogenic inheritance. Two linkage maps were constructed from RFLP and position shift loci which revealed that protein markers were interdispersed between the RFLP markers on all chromosomes [29]. In many cases position shift variants correspond to the same protein as shown by microsequencing [30, 31]. It is expected that the maps of expressed genes obtained by 2-DE will be crucial for the candidate gene strategy of quantitative trait loci (QTL) characterization [23]. QTL analysis has been applied to map genes controlling protein quantity for spots on 2-D gels [30, 32] and the loci have been termed PQL for protein quantity loci [1]. Colocalization of a protein quantity locus (PQL) and its protein-coding locus would indicate that the expression level of the protein is a consequence of allelic differences, whereas colocalization between a PQL and a QTL for a different trait would point to an association of a candidate gene and the variation observed for a trait. In extension of these pioneering works, strategies combining the power of genetics with the tools of functional genomics were recently suggested and the term genetical genomics was coined [2, 23, 33, 34].

In parallel to proteome analysis, transcript profiling is a large-scale technique becoming more and more accessible to study plant systems and used, for example, to analyze *Arabidopsis* seed development [35] or to monitor drought and cold stress responses [36]. It can be antici-

pated that these techniques will be complementary and the combined use of both approaches on the same material will indicate which of the regulatory control processes under study are exerted at the level of transcription or translation.

3 Subcellular proteome analysis in plants

The identification of those proteins recruited to fulfill the specific function of subcellular compartments gives an additional dimension to the proteome analysis. During the last years of the preceding decade, a few pioneering efforts resulted in the first catalogues of proteins present in the cell wall [37], the plasma membrane [38], the mitochondrion, the endoplasmic reticulum, and the Golgi apparatus [39]. In most cases, proteins were characterized using Edman chemistry, a technology not really appropriate to the identification of protein mixtures such as those obtained after SDS-PAGE. In addition, concerning the membrane proteins that were solubilized with the help of currently used detergents and resolved by 2-DE, it became rapidly clear that most integral proteins did not enter the gel, although some of them were detected using specifically selected detergents [40]. Both the sequencing of the *Arabidopsis* genome, the generalized use of MS and the introduction of alternative procedures to resolve hydrophobic proteins have since paved the way towards efficient proteome analysis of organelles and associated membranes. In addition, several bioinformatics resources became recently available, providing catalogs of predicted membrane proteins, such as *Aramemnon* (http://aramemnon.botanik.uni-koeln.de), the *Arabidopsis* Membrane Protein Library (http://www.cbs.umn.edu/arabidopsis/) and the Rice Membrane Protein Library (http://www.cbs.umn.edu/rice/), thus providing further support for proteomics efforts. Most efforts in recent years concerned the mitochondrion and the chloroplast, with additional attention towards the endoplasmic reticulum [9] and the vacuolar membrane (Szponarski *et al.*, this issue).

3.1 Chloroplasts

In the case of the chloroplast, three lines of work addressed the protein composition of thylakoid and chloroplast envelope respectively, and that of some chloroplast protein complexes:

(i) The classical combination of 2-DE with Edman sequencing and MS was first used to characterize purified fractions of thylakoid membrane and luminal compartment [41]. This allowed the identification of more than 60 proteins in pea, and detailed sequence analysis demonstrated that approximately half of the lumen proteins displayed the so-called twin-arginine motif for translocation pathway. In addition, this work provided novel information to refine the prediction of transit peptides. More recently, a similar approach was used to compare thylakoid lumen proteins from *Arabidopsis* and spinach [42]. A good correlation between the two proteomes was observed, allowing the theoretical estimation of *ca.* 80 proteins for the *Arabidopsis* thylakoid luminal proteome. In addition, this work confirmed that half of the luminal precursors in *Arabidopsis* displayed the translocation twin-arginine motif. The use of 2-DE after protein extraction with solvents revealed abundant membrane proteins having more than ten transmembrane domains, like some components of the photosystem I, in the thylakoid membrane of the eukaryotic green alga *Chlamydomonas reinhardti* [43]. Furthermore, this work provided evidence that some photosystem I proteins undergo differential processing and maturation, thus accounting for the high complexity of the corresponding proteome. In pea and spinach, an alternative procedure, based on RP-HPLC separation of proteins, was used to resolve the proteins constituting the photosystem II [44]. Using this method, the authors succeeded in identifying *ca.* 40 gene products from their intact mass tag. In addition, the procedure allowed the detection of some post-translational modifications. However, although well-suited for the detection of minor populations showing specific modifications, this approach seems limited to proteomes of moderate complexity.

(ii) Another recent major progress concerns the characterization of hydrophobic proteins from the envelope. Using chloroform/methanol extraction, followed by SDS-PAGE and MS/MS analysis, more than 40 proteins, including one-half of novel envelope proteins, were first characterized from spinach [45, 46]. In *Arabidopsis*, a similar approach led to the identification of nearly 60 proteins, including again one-half of new envelope proteins [47]. In both cases, most proteins were highly hydrophobic. Using this experimental basis, the authors were able to derive a set of features proven to be shared by most envelope integral proteins, and further used these features to identify in the *Arabidopsis* genome more than 50 new candidates for envelope transporters.

(iii) Two kinds of chloroplast protein complexes have been characterized to date. The protein composition of the plastid ribosomal 30 S and 50 S subunits was investigated using a combination of various protein separation and characterization procedures [48, 49]. Nearly 60 proteins were identified in spinach, demonstrating that most of them are orthologues of *Escherichia coli* ribosomal proteins, whereas 10% are plastid-specific proteins. Another chloroplast complex that was characterized recently is the ClpP protease complex [50]. Using nondenaturing

Blue-Native gel electrophoresis to resolve the complex and MS to identify components, ten different Clp isoforms were identified in *Arabidopsis*, demonstrating an unexpected complexity.

3.2 Mitochondria

As for other organelles, the mitochondrial proteome was first investigated using 2-DE of total protein extracts [51]. Over 50 proteins were characterized, including 20% of unidentified proteins not previously described in plant mitochondria. Similar results were obtained overall when fractionating the protein population into soluble, membrane peripheral, and membrane integral proteins [52], although a higher proportion of proteins was found to correspond to orphan proteins. Not surprisingly, in both cases, almost no very hydrophobic protein was detected. More recently, another more sophisticated electrophoretical approach, based on separation of complexes by Blue-Native electrophoresis prior to 2-DE, was introduced [53]. This method was used to resolve the components of several complexes encompassing various isoforms and hydrophobic proteins. However, none of the numerous carrier proteins, known to control communication with the cytosol, could be identified by any of these approaches. Very recently, the SDS-PAGE patterns obtained after carbonate stripping of membrane to remove peripheral proteins or chloroform/methanol extraction of membranes were compared [54]. After characterization of bands by MS/MS, the two-thirds of the carriers highly represented in EST databases (such as adenine nucleotide translocator, dicarboxylate/tricarboxylate carrier or phosphate transporter) could be identified from carbonate-stripped membranes but not from solvent extracts. Simultaneously, numerous other hydrophobic proteins were identified, thus exemplifying the efficiency of the procedure.

Several efforts based on current proteome technologies have led in a few years to various methodological alternatives now allowing access to membrane proteins. In this respect, both the mitochondrial and chloroplast examples demonstrate that organelle proteomics is presently able to identify most proteins constituting a specialized subcellular compartment in plants. In addition, the data thus generated offer complementary opportunities to exploit the genome information. For instance, several of the surveys have helped improve genome annotation, correcting errors in automatic intron and exon prediction [47, 50]. In the same way, these data allow the evaluation of current predictions of transit peptides and protein targeting [41, 52], and even to provide evidence for alternative transit peptides [55]. In addition to improvements in interpreting genomic sequence data, organelle proteomics can help predict sets of orphan proteins likely to be sharing the same subcellular location [47], thus gaining a new functional value.

4 Proteomic analysis of trees

4.1 Economic and ecological importance of trees

Woody plants comprise a large group of angiosperm and gymnosperm species of economic importance [56] that play a crucial role in the lives of humans and in the functioning of ecosystems. Despite their economic and ecological relevance, progress from molecular, biochemical, and physiological studies in trees has been constrained by various problems, including their large physical size, the usually large genome in many species of economic interest, the long life cycle of perennial species, the recalcitrance for genetic transformation and regeneration *in vitro*, and difficulties in sample preparation for molecular and biochemical analyses.

4.2 Development of genomic and proteomic studies in trees

The difficulties of woody plants as experimental models have negatively influenced the development of genomic and proteomic studies. For example, the isolation of nucleic acids or the preparation of protein samples is particularly troublesome because the usual interfering substances in plants such as polysaccharides, pigments and phenolics, are especially abundant in lignified tissues of trees. Nevertheless, the separation of polypeptides by 2-DE and further protein identification by using standard techniques have been widely used by researchers during the last decade in a variety of applications in trees. These approaches have been applied, for example, to study changes in the protein profiles during early stages of loblolly pine development [57], to characterize the response of pines to biotic [58] and abiotic stresses [59], to differentiate isoenzymes involved in pine nitrogen metabolism [60] or for determination of genetic polymorphisms and genomic mapping in maritime pine [61–63].

In the last few years, the completion of the *Arabidopsis* genome sequence, the availability of genome drafts for rice and the continuous progress in the genomics of other important herbaceous species open the possibility for genomics initiatives in woody models and the subsequent development of major breakthrough technologies for global analysis of gene expression. In this way, structural/functional genomic and proteomic studies are now

feasible in trees. Two major models have been proposed: poplar for angiosperms [64] and pine for gymnosperms [65]. Large-scale EST sequencing projects have been initiated for poplar [66] and loblolly [67] and maritime pines [68]. ESTs databases from different woody tissues have been established and the information is accessible on the internet (http://www.biochem.kth.se/PopulusDB/; http://pinetree.ccgb.umn.edu; http://cbi.labri.fr/outils/SPAM/index.php).

Transcriptome analysis of wood formation has received special attention. Large-scale cDNA sequencing and expression profiling allow the identification of key molecular players in the process [56, 66–69]. ESTs derived from these studies have shown similarity to proteins of known function revealing the biochemical processes involved in xylem differentiation. Structural proteins (actin, α- and β-tubulin), cell wall-related proteins (arabinogalactan proteins, cellulases, extensins and glycine-rich like proteins, enzymes for lignin biosynthesis: phenylalanine ammonia-lyase, cinnamate-4-hydroxylase, caffeoyl-CoA-*O*-methyltransferase, cinnamyl alcohol dehydrogenase, housekeeping proteins (ubiquitin, peptidylprolyl isomerase, elongation factor 1-α), metallothioneins and energy production proteins (flavoprotein, ADP-ATP carrier protein) were among the most abundant expression products. Glycine hydroxymethyltransferase and enzymes of the C1 metabolism were also abundant and may reflect the demand for methyl groups in wood forming tissue. Transcription factors such as zinc finger proteins, myb, homeobox, and LIM-like proteins could play important regulatory roles in differentiating xylem. Several other highly abundant transcripts corresponded to a translationally controlled tumor protein (whose function in wood formation remains unknown), and a blue copper protein (possibly involved in oxidative polymerization reactions of lignin monomers). Interestingly, 12–25% of the isolated ESTs from woody xylem showed no similarity to sequences in the databases and may have unique functions in wood formation.

The establishment of EST databases from woody plant tissues has enabled the development of global analysis methods to study changes in gene expression associated with important processes in trees such as xylem differentiation, embryogenesis or tree growth [56, 68–72].

However, proteins are the final products of the gene expression and, therefore, large-scale analyses of proteins are needed to complement the data derived from transcriptome analysis. In fact, only differences affecting the proteome can be directly attributed to changes in function. 2-DE and image analysis have been used for proteome analysis in maritime pine [73]. Proteins from needles and xylem, two important tissues for growth and wood formation, were separated and characterized by internal peptide microsequencing. Identified proteins are accessible in a proteomic database on the internet (http://www.pierroton.inra.fr/genetics/2D/). The differential accumulation of proteins specifically associated to compression wood has also been studied in the same pine species by proteomic approaches [74]. Upregulated proteins included 1-aminocyclopropane-1-carboxylate oxidase (an ethylene forming enzyme), a putative transcription factor, two lignification enzymes (caffeate *O*-methyltransferase and caffeoyl CoA-*O*-methyltransferase), members of the *S*-adenosyl-L-methionine synthase family, and enzymes involved in nitrogen and carbon assimilation (glutamine synthetase and fructokinase). Since compression wood is biochemically characterized by the relative high content of lignin, the results suggest the functional regulation of proteins corresponding to specialized pathways participating in lignin biosynthesis during wood formation. Similar studies have also been performed during wood formation in poplar. Proteins differentially produced in developing poplar xylem were identified by 2-DE and microsequencing [75]. Proteins were identified by similarity to other previously characterized polypeptides and correspondence to ESTs in the poplar database [66]. The results were quite similar to those found in pine, and again, enzymes involved in cell wall and lignin biosynthesis, *S*-adenosyl-L-methionine-synthases, glycine hydroxymethyltransferases, and glutamine synthetase were among the most abundant xylem proteins in poplar.

In summary, a combination of high-throughput proteomic and transcriptomic techniques is allowing the identification of key molecular players in the process and providing new insights on the metabolic pathways involved in wood formation. This new knowledge can be complemented by the use of *Arabidopsis* as a genetic system to investigate the molecular controls regulating the identity of the vascular cambium and the development of secondary xylem and phloem. Under appropriate growth conditions the secondary xylem of *Arabidopsis* closely resembles the anatomy of the wood of an angiosperm tree. Chaffey *et al.* [76] have recently shown that basic questions about wood formation can be addressed by using *Arabidopsis* as a model for wood formation.

4.3 Future prospects

As mentioned before, the intrinsic difficulties of woody plants as experimental models have limited the application of molecular techniques until recently. Nevertheless, considerable knowledge about the biology of woody plants has been gained in the recent years by the application of the new genomic technologies to study tree

growth and development as well as the response of trees to biotic and abiotic stresses. Functional genomic approaches, such as large-scale EST sequencing, gene expression profiling, and proteomic approaches, are allowing the identification of genes/proteins involved in several important processes in trees. Although these new developments in functional genomics are restricted to a few tree models, the rapid extension to other woody plants of great commercial interest can be expected. A relevant milestone in the near future will be the availability of the first complete genome sequence from a tree species. Due to the large genome size of conifers, poplar has recently been proposed as the model tree to accomplish this goal. An international consortium has been constituted to develop this project that is expected to have the first draft by the end of 2003 (http://bahama.jgi-psf.org/prod/bin/populus/whitepaper.populus.cgi; http://genome.jgi-psf.org/poplar0/poplar0.home.html). The imminent availability of the poplar genome information as well as the characterization of a large number of full-length cDNAs in conifer models will permit the functional characterization of proteins/enzymes encoded by the tree genomes, and the determination of their physiological roles. Thus, the expression, purification and molecular characterization of a large number of recombinantly expressed proteins will be possible. Specific antibodies and protein arrays will be generated and used as important tools for the functional characterization of woody plant systems. The precise localization of mRNAs and proteins in cells and tissues will be determined providing new insights on how metabolic pathways are organized in different cell types. Furthermore, functional studies in transgenic trees are now possible because routine transformation protocols *via Agrobacterium* are available for poplar and rapid progress in this technology has been reported in the last few years for conifers. Results obtained with woody plants will be compared with the information derived from functional genomic studies currently carried out in *Arabidopsis*. The biological information derived from ongoing research efforts will greatly enhance our understanding of the molecular basis of tree structure and function. This new knowledge will have a clear impact on the future of forestry practices and management.

5 Proteomics of plant-microbe interactions

Plant-microbe interactions range from mutualistic to pathogenic ones. Regarding the beneficial interactions, mycorrhizal symbioses concern most vascular plant species while rhizobial symbioses are essentially established between legumes and nitrogen-fixing bacteria of the *Rhizobiaceae* family. Although occurring to a lesser extent, plant-pathogen interactions are of great agronomic and economic concern. Besides molecular approaches, proteomics, based on the recent developments of 2-DE, MS and bioinformatics, offers a complementary insight into protein expression and regulation within plant-microbe interactions. After reviewing past and recent studies, the potential of proteomics for an integrated understanding of the processes involved in plant-microbe interactions will be discussed. Attempts to identify the main bottlenecks will be presented.

5.1 Proteomics for studying plant pathogen interactions

Although there is an increasing amount of literature dealing with genes involved in bacterial and fungal plant pathogenesis [77], very few reports have addressed proteome modifications associated with such interactions. Changes in proteins from leaf blades of rice plants infected with the blast fungus *Magnaporthe grisea* were recently investigated following 2-DE and *N*-terminal/internal amino acid sequencing. Forty-five proteins out of 63 were identified [78]. A targeted proteomic approach was used to identify the most abundant proteins from tomato xylem sap upon infection with *Fusarium oxysporum*. This not only confirmed the presence of known pathogenesis-related (PR) proteins in the tomato sap, but also the identification of a new PR-5 isoform [79]. Beside the analyses of the whole plant-pathogen proteome, several reports have focused on elicitation/perception signalling mechanisms. Rakwal [80] studied the role of exogenous jasmonic acid (JA) in defence mechanisms of rice (*Oryza sativa* L.) using proteome analysis. Proteins appearing or modified were identified through *N*-terminal and/or internal amino acid sequencing. Proteomics followed by immunological studies indicated that JA affects defense-related gene expression in rice seedlings, as evidenced by *de novo* synthesis of novel proteins with potential roles in plant defense. In the field of plant perception signalling, protein kinases play a central role during pathogen recognition and subsequent activation of plant defence mechanisms. The first bases for dissecting the upstream protein phosphorylation pathways involved in the signal transduction of cryptogein, an elicitor of defence reactions in *Nicotiana tabacum* cells, were obtained following 2-DE analyses. Although this study clearly highlighted the role of protein kinases and/or constitutive phosphatases in the cryptogein signal transduction, only two microsequences, without homology with known proteins, were obtained for the phosphoproteins [81]. More recently, in a remarkable approach, *Arabidopsis* suspension-cultured cells were pulse-labeled with radioactive orthophosphate during treatment with the flagellin elicitor. This allowed, in conjunction with 2-DE and MS, to set up the first phos-

phoproteome of flagellin- and chitin-treated *Arabidopsis* cells. The availability of the *Arabidopsis* genome sequence was of great advantage, as it allowed subsequent protein identification using peptide mass fingerprinting. One of these proteins, AtPhos43, was identified as being a novel protein specifically phosphorylated in response to the bacterial elicitor [82]. Directed proteomics was shown to be a viable method for analyzing signal transduction in plants.

The effects of leaf extracts from a susceptible host-plant, a resistant and a non-host plant on the bacterial proteome of *Xanthomonas axonopodis* pv. *citri* were compared by 2-DE. A few differentially expressed proteins were *N*-terminally sequenced but putative functions assigned following homology searches in databases were uncertain since few Xanthomonad genes have been described [83]. Recently, proteomics of the plant pathogen *Xyllela fastidiosa* revealed major cellular and extracellular bacterial proteins, including toxins, adhesion-related proteins, antioxidant enzymes, and proteases [84].

5.2 Proteomics for studying plant symbioses

5.2.1 Model plants

Many agronomically important crop plants are legumes, such as soybean, pea, and alfalfa. However, the size and complexity of their genomes make them unwieldy and have slowed progress on the genetic characterization of these crops. Recently, genomics have been initiated in *Lotus japonicus* [85] and *Medicago truncatula* [86–88]. These model legume species have received increasing attention during recent years, due to their simple diploid genome, short life cycle, ease of transformation, and regeneration [89, 90]. However, by comparison to *Arabidopsis thaliana*, *L. japonicus* and *M. truncatula* possess relatively large genomes (~500 Mb). Although an international genome sequencing project has recently been launched (http://www.noble.org/press_release/medicago/NewsConference2001/MedicagoSequencingProject.htm) for *M. truncatula*, current genomic information is essentially in the form of expressed sequence tags (ESTs) [91, 92]. A regularly updated database containing more than 160 000 *M. truncatula* ESTs, representing roughly 29 000 clustered sequences, is held by the TIGR (The Institute for Genomic Research, Rockville, Maryland) [93]. In France, a joint program between the Genoscope (Evry), the CNRS-INRA LBMRPM (Toulouse) and INRA-Université de Bourgogne PME (Dijon) laboratories, produced 21'473 5'-3'-ESTs (expressed sequence tags) from control plants, nodules and mycorrhiza, yielding 6'359 EST clusters, corresponding to distinct virtual genes [86, 94]. Major advances have come from the use of plant mutants, isolated following ethyl methane sulfonate treatment, γ-ray irradiation, insertion mutagenesis, or screening for natural variants [95–97]. The mutants studied are defective in defined steps of the symbiosis, helping to dissect the sequence of events leading to a nodule or a mycorrhiza [98–100].

5.2.2 Nitrogen-fixing symbiosis

The rhizobial symbiosis involves specific host-microorganism recognition mechanisms mediated by chitooligosaccharide molecules (Nod factors) produced by the bacteria. Bacteria colonize roots within infection threads and induce a meristematic activity in inner cortical cells, leading to the formation of nodules. Once inside the nodules, they differentiate into bacteroids, which are active in the fixation and reduction of atmospheric nitrogen into ammonium [101]. Due to the agronomic importance of this symbiosis, the genetic program of plant/microbe interaction has been extensively dissected [88, 102]. With the recent completion of the genome of two rhizobial species, *Mesorhizobium loti* and *Sinorhizobium meliloti*, and the hundreds of thousands of expressed sequence tags from three major legume species: soybeans, *M. truncatula*, and *L. japonicus* deposited in the public domain [87, 88, 103–105], functional genomics is now a reality.

Since Govers's publication [106], in which the expression of plant genes involved in the pea-*Rhizobium* symbiosis was studied by 2-DE separation of *in vitro* translated mRNA root nodule products, several reports were published with extensive protein identifications of rhizobial symbioses. Recently, 2-DE was used to identify proteins differentially expressed during the symbiotic interaction between the bacterium *S. meliloti* strain 1021 and the legume *Melilotus alba* (white sweetclover) [107]. By identifying gene products that are differentially present between symbiotic and nonsymbiotic states, the aims were to characterize novel symbiosis proteins and to determine how the two symbiotic partners alter their respective metabolisms during the interaction. Proteome maps from control *M. alba* roots, wild-type nodules, and cultured *S. meliloti* and *S. meliloti* bacteroids were generated and compared. More than 250 proteins were induced or upregulated in the nodule compared with the root, and over 350 proteins were downregulated in the bacteroid form, compared to cultured cells. *N*-Terminal amino acid sequencing and MALDI-TOF-MS peptide mass fingerprint analysis, in conjunction with database searching, were used to assign putative identity to nearly 100 nodule, bacterial, and bacteroid proteins. This work clearly demonstrated that proteomics is a useful strategy to link sequence information and functional genomics.

Similarly, differential proteomics was used to identify proteins involved in the early stages of nodulation between the subterranean clover cv. Woogenellup and the *Rhizobium leguminosarum* bv. *trifolii* strains ANU843 and ANU794 [108]. Strain ANU843 induces nitrogen-fixing nodules whereas strain ANU794 forms aberrant nodules on the cv. Woogenellup roots that fail to develop beyond an early stage. The aim was to identify proteins that might be involved in the early stages of nodulation over a 48 h period and those differentially displayed during the interactions between the host and the two microbes. Proteome maps from control and inoculated roots were generated and compared at 24 and 48 h post inoculation. Of the 16 protein spots that were differentially displayed or developmentally regulated, 10 were assigned with putative identities. Of the 22 constitutively expressed proteins spots examined, through *N*-terminal sequencing, 8 spots were assigned a putative protein homology, including several pathogenesis and stress-related proteins.

Recently, a proteome reference map for *M. truncatula* root proteins was established using 2-DE combined with peptide mass fingerprinting to aid the dissection of nodulation and root developmental pathways [109]. It will be updated continuously (http://semele.anu.edu.au/2d/2d.html). Over 2500 root proteins were reproducibly displayed. 485 proteins were analyzed by peptide mass fingerprinting, among which 179 were identified by matching against the current *M. truncatula* ESTs database. The majority of identified proteins were metabolic enzymes and stress response proteins, and 44% of proteins occurred as isoforms, a result that could not have been predicted from sequencing data alone.

Nevertheless, due to the intrinsic limitation of the 2-DE techniques, it has soon appeared necessary to focus on restricted cell tissue compartments. Indeed, from the symbiosome soybean membrane, Winzer *et al.* [110] isolated a nodule-specific 53-kDa protein (GmNOD53b) allowing then its functional analysis to be achieved. Later, soybean peribacteroid membrane (PBM) proteins were isolated from nitrogen-fixing root nodules and subjected to *N*-terminal sequencing [111]. Sequence data from 17 putative PBM proteins were obtained. The identification of homologues of HSP70 and HSP60 associated with the PBM was the first evidence that the molecular machinery for co- or post-translational import of cytoplasmic proteins is present in symbiosomes. More recently, Saalbach *et al.* [112] examined pea root nodules and identified 46 proteins from the PBM and from the space between this membrane and the bacteroid one. Concerning the actinorhizal symbiosis, only one case study reported on the modifications of the protein expression pattern in the nitrogen-fixing Frankia sp. strain ACN14a-tsr induced by root exudates of its symbiotic host *Alnus glutinosa* [113].

5.2.3 Ectomycorrhizal symbiosis

Ectomycorrhiza are mutualistic associations formed between roots (mainly trees) and a wide range of soil ascomycetes and basidiomycetes. Pioneering proteomic studies were achieved with ectomycorrhiza in the early nineties [114–117] allowing the detection of symbiosis-related (SR) proteins, upregulated or newly induced in mycorrhizal roots, as well as downregulated proteins, by comparison to control roots and mycelium extracts. Some "targeted" proteomics studies were also successfully achieved allowing the identification of proteins of interest [118–122]. Only very recently, some SR proteins as well as mycelial proteins were identified by MS and *N*-terminal sequencing [122–124]. A large-scale proteomic project is now planned in order to match with EST data obtained for *Pinus sylvestris* (Martin F., personal communication).

5.2.4 Endomycorrhizal symbiosis

Most vascular flowering plants, including many agriculturally important crop species, form symbiotic associations with arbuscular mycorrhizal (AM) fungi [125]. The resulting association plays an essential role in the acquisition of mineral nutrients and leads, among other benefits, to enhanced plant growth and health. In spite of the difficult biological system with regard to the obligatory biotrophic status of AM fungi and their asynchronous infection process, a similar progression was followed for AM interactions, when compared to other root symbioses, going from descriptive studies [126–131] to the first identifications [132, 133].

Proteome analysis is now being used as a powerful tool to reveal more and more proteins involved in AM symbiosis [134]. Bestel-Corre *et al.* [135], in a time-course analysis, compared two *M. truncatula* symbioses, for which roots were inoculated with either the AM fungus *Glomus mosseae* or the nitrogen-fixing bacterium *S. meliloti*. No common plant protein was found to be induced by both symbionts, although this was expected from molecular data [98, 136–142]. However, numerous proteins were characterized as up- and downregulated, or newly synthesized, among which a few were identified by MS/MS. Other symbioses-related proteins were recently identified by MALDI-TOF and peptide mass fingerprinting [143].

Additionally, although protein patterns of dormant and germinated spores of several AM fungi were established quite early [144], the lack of databases on fungi (and particularly on AM fungi) has slowed down the use of proteomics. Only a few identifications were recently obtained

following MS/MS [145]. The use of Ri T-DNA transformed roots of *Daucus carota* [145] or *M. truncatula* ([146]; Dumas-Gaudot E., unpublished data) will undoubtedly help to collect arbuscular extraradical fungal material. Very recently, proteomics was also employed to study the impact of xenobiotics, revealing thus a potential use of this technology for environmental studies [143, 147, 148]. Similarly, variations in the pea (*Pisum sativum* L.) root proteome were identified in response to cadmium stress, during the symbiotic interaction with *G. mosseae* [149].

In conclusion, despite the complexity of the biological material when working with two interacting genomes (plant and microorganism), proteomics appears as a powerful tool to gain a global picture of plant-microbe interactions. In the near future, further protein identifications with peptide mass fingerprinting will be facilitated by the use of model plants and/or microbes. Protein identifications may also be obtained for nonsequenced organisms (plants/pathogens/beneficial microbes), assuming that MS/MS analyses can be performed. In addition, attempts should be made to (i) identify proteins corresponding to defined steps of the studied plant/microbe interaction processes and, whenever possible, (ii) to relate proteome and transcriptome analyses. To this end, a protocol has recently been set up for carrying out on the same root proteomics and transcriptomics analyses of the early steps of the AM symbiosis (Dumas-Gaudot *et al.*, this issue). In addition, subcellular fractionation could be used to enrich plant extracts with specific proteins. In particular, membrane proteins may be addressed, since significant advances have been achieved for their extraction and solubilization (this article and [46, 150–153]). Special attention may be paid to the perisymbiosome membrane, surrounding the arbuscules or the bacteroids. This approach, recently used for *L. japonicus* nodules [154], is currently being developed to the arbuscular symbiosome (Dumas-Gaudot *et al.*, unpublished results). The availability of more and more reference proteome maps including the plant phosphorilome [155] will provide a basis for future proteome comparisons of biotically and abiotically challenged plants [156].

6 Conceptual and technological perspectives for plant proteomics

Plant proteomics is sharing the technical limitations of proteome analysis in general but will benefit from rapid developments in current methodology. Proteins are physically and chemically much more diverse than nucleic acids, which hinders the quantitative analysis of complex samples of proteins. In addition, due to different RNA splicing and post-translational modifications, it is expected that for a given organism the number of protein species exceeds severalfold the number of genes. Another level of complexity arises when considering the potential number of protein-protein interactions in an organism modified by developmental events and physiological constraints. In the following we will address only a few selected topics most likely representing key steps in future proteomic studies and the reader is referred to recent reviews on general technological developments in proteomics [3, 157–159].

Although not comprehensive, the overview on the literature published in the field of plant proteomics given in the previous sections showed that to date separation of complex mixtures mostly relied on 2-DE. To overcome the limitations of this technology, specific tissues or sub-proteomes have been used as a starting material. With the aim of increasing the resolution of 2-DE of plant tissue proteins by improving the protein extraction procedure, Giavalisco *et al.* [160] reported a protocol comprising sequences of steps for tissue desintegration, protein solubilization, and removal of insoluble material by ultracentrifugation, leading to three different fractions. The first one (obtained by mechanical cell wall disruption in the presence of protease inhibitors) contains only buffer-soluble, mainly cytosolic proteins, the second one includes membrane, hydrophobic, structure-associated proteins (with chemical disruption of membranes by using detergents, chaotropes, and further mechanical grinding), and the third one nucleic acid-associated proteins (by using DNAse). The authors conclude that the fractionation-based extraction, compared with the precipitation one, avoids protein losses and decreases the complexity of the protein pattern, leading to higher resolution; additional advantages are reduced protein comigration, and higher protein-loading capacity.

Current technical developments are directed towards improved resolution and better quantification and include novel techniques for sample prefractionation such as by free-flow electrophoresis or with a chromatofocusing device such as the multicompartment electrolyzer [161], the use of a set of narrow pH gradients and larger gel-formats during 2-DE [162], and the introduction of new staining and labelling methods [163–165]. Difference gel electrophoresis (DIGE) is a prelabelling technique using separate Cy dyes for different samples which then can be analyzed on one gel avoiding shifts in gel patterns normally occurring when samples conventionally separated on two gels are compared [166]. This technique promises to make quantitative analysis based on 2-DE much more reliable [167]. Whereas the labelling

procedure itself appears relatively easy to use, dedicated equipment and dedicated software are necessary for evaluation of the gel(s).

Liquid chromatography (LC) is an emerging alternative to 2-DE for protein separation, although both can be combined in a single experiment, when dealing with low-abundant proteins [168], particularly in the case of *n*-dimensional chromatography [169]. LC has the advantage of using liquid phase compatible with different infusion into a mass spectrometer (LC-MS), although unfortunately, proteins are not generically amenable to high resolution by RP-HPLC. This has been solved by digesting the proteins to peptides with a protease, normally trypsin. The digestion process creates more complexity, but it can be simplified by the use of isotope-coded affinity tags (ICATs) for generic protein expression analysis [170]. ICAT-based LC-MS offers the ability of selecting certain peptides by derivatization of specific amino acids (typically cysteine residues), with an affinity tag. Use of stable isotope-labelled affinity tag also permits quantitation and discrimination on the mass spectrometer itself. Ficarro *et al.* [171] report a similar method for enriching phosphoproteins by affinity columns of affinity tag-derived phosphoproteins.

Another interesting approach is the multidimensional protein identification technology (MudPIT) approach introduced in Yates laboratory [172]. MudPIT is a technique for the separation and identification of complex protein and peptide mixtures. Rather than using traditional 2-D gel electrophoresis, MudPIT separates peptides by 2-D LC. In this way, the separation can be interfaced directly with the ion source of a mass spectrometer. This technique has recently been applied to the study of plant proteomes [11, 173].

The identification of multiprotein complexes components and the analysis of protein-protein interaction are essential to understand most cellular processes. Protein interaction studies are carried out by using two-hybrid systems, developed by Fields [174], protein chips, and the large-scale approach of tandem-affinity purification (TAP)-MS. This last strategy has been used in the analysis of the yeast proteome, allowing the purification of more than 589 multiprotein assemblies [175] and can be applied to higher eukaryotes, avoiding the problem of the competition from corresponding endogenous proteins [176]. A more recent general method involves the use of fluorescence resonance energy transfer (FRET) between fluorescent tags on interacting proteins, by using green, cyan, and yellow fluorescent protein [159]. The great advantage of the approach is its application to *in vivo* analysis by microscopy.

The authors gratefully acknowledge Dr. Richard Thompson (Dijon) for critical reading of the manuscript. J. Jorrin acknowledges financial support from the Spanish MCyT (AGL2002-03248) and the Consejería de Educación y Ciencia, Junta de Andalucía. H.-P. Mock thanks the DFG for funding (Mo 479/4-1; Mo 479/4-2).

Received April 25, 2003
Revised July 22, 2003
Accepted July 28, 2003

7 References

[1] Thiellement, H., Bahrman, N., Damerval, C., Plomion, C. *et al.*, *Electrophoresis* 1999, *20*, 2013–2026.

[2] Zivy, M., de Vienne, D., *Plant Mol. Biol.* 2000, *44*, 575–580.

[3] van Wijk, K. J., *Plant Physiol.* 2001, *126*, 501–508.

[4] Rossignol, M., *Curr. Opin. Biotechnol.* 2001, *12*, 131–134.

[5] Porubleva, L., Vander Velden, K., Kothari, S., Oliver, D. J. *et al.*, *Electrophoresis* 2001, *22*, 1724–1738.

[6] Gallardo, K., Job, C., Groot, S. P. C., Puype, M. *et al.*, *Plant Physiol.* 2001, *126*, 835–848.

[7] Gallardo, K., Job, C., Groot, S. P. C., Puype, M. *et al.*, *Plant Physiol.* 2002, *129*, 823–837.

[8] Finnie, C., Melchior, S., Roepstorff, P., Svensson, B., *Plant Physiol.* 2002, *129*, 1308–1319.

[9] Maltman, D. J., Simon, W. J., Wheeler, C. H., Dunn, M. J. *et al.*, *Electrophoresis* 2002, *23*, 626–639.

[10] Goff, S. A., Ricke, D., Lan, T. H., Presting, G. *et al.*, *Science* 2002, *296*, 92–100.

[11] Koller, A., Washburn, M. P., Lange, B. M., Andon, N. L. *et al.*, *Proc. Natl. Acad. Sci. USA* 2002, *99*, 11969–11974.

[12] Woo, S. H., Fukuda, M., Islam, N., Takaoka, M. *et al.*, *Electrophoresis* 2002, *23*, 647–654.

[13] Hajduch, M., Rakwal, R., Agrawal, G. K., Yonekura, M., *et al.*, *Electrophoresis* 2001, *22*, 2824–2831.

[14] Santoni, V., Bellini, C., Caboche, M., *Planta* 1994, *192*, 557–566.

[15] Herbik, A., Giritch, A., Horstmann, C., Becker, R. *et al.*, *Plant Physiol.* 1996, *111*, 533–540.

[16] Moseley, J. L., Allinger, T., Herzog, S., Hoerth, P. *et al.*, *EMBO J* 2002, *21*, 6709–6720.

[17] Santoni, V., Delarue, M., Caboche, M., Bellini, C., *Planta* 1997, *202*, 62–69.

[18] Faure, J. D., Vittorioso, P., Santoni, V., Fraisier, V. *et al.*, *Development* 1998, *125*, 909–918.

[19] Kasten, B., Buck, F., Nuske, J., Reski, R., *Planta* 1997, *201*, 261–272.

[20] Tacchini, P., Fink, A., Thiellement, H., Greppin, H., *Physiol. Plant.* 1995, *93*, 312–316.

[21] Damerval, C., De Vienne, D., *Heredity* 1993, *70*, 38–51.

[22] Damerval, C., Le Guilloux, M., *Mol. Gen. Genet.* 1998, *257*, 354–361.

[23] Thiellement, H. P., C. Zivy, M., Pennington, M. J. (Ed.), in: *Proteomics*, BIOS Scientific Publishers, Oxford 2001, pp. 289–309.

[24] Burstin, J., Devienne, D., Dubreuil, P., Damerval, C., *Theor. Appl. Genet.* 1994, *89*, 943–950.

[25] Riccardi, F., Gazeau, P., de Vienne, D., Zivy, M., *Plant Physiol.* 1998, *117*, 1253–1263.

[26] Kristoffersen, H. E., Flengsrud, R., *Electrophoresis* 2000, *21*, 3693–3700.

[27] Ostergaard, O., Melchior, S., Roepstorff, P., Svensson, B., *Proteomics* 2002, *2*, 733–739.

[28] Skylas, D. J., Copeland, L., Rathmell, W. G., Wrigley, C. W., *Proteomics* 2001, *1*, 1542–1546.

[29] De Vienne, D. B. J., Gerber, S., Leonardi, A., Le Guilloux, M., *et al.*, *Heredity* 1996, *76*, 166–176.

[30] Touzet, P., Morin, C., Damerval, C., Leguilloux, M. *et al.*, *Electrophoresis* 1995, *16*, 1289–1294.

[31] Plomion, C. C., P. Bahrmann, N., Frigerio, J. M., *Silvae Genetica* 1997, *46*, 161–165.

[32] Damerval, C., Maurice, A., Josse, J. M., Devienne, D., *Genetics* 1994, *137*, 289–301.

[33] Jansen, R. C., Nap, J. P., *Trends Genet.* 2001, *17*, 388–391.

[34] Consoli, L., Lefevre, A., Zivy, M., de Vienne, D. *et al.*, *Plant Mol. Biol.* 2002, *48*, 575–581.

[35] Girke, T., Todd, J., Ruuska, S., White, J. *et al.*, *Plant Physiol.* 2000, *124*, 1570–1581.

[36] Seki, M., Narusaka, M., Ishida, J., Nanjo, T. *et al.*, *Plant J.* 2002, *31*, 279–292.

[37] Robertson, D., Mitchell, G. P., Gilroy, J. S., Gerrish, C. *et al.*, *J. Biol. Chem.* 1997, *272*, 15841–15848.

[38] Santoni, V., Rouquie, D., Doumas, P., Mansion, M. *et al.*, *Plant J.* 1998, *16*, 633–641.

[39] Prime, T. A., Sherrier, D. J., Mahon, P., Packman, L. C. *et al.*, *Electrophoresis* 2000, *21*, 3488–3499.

[40] Santoni, V., Rabilloud, T., Doumas, P., Rouquie, D. *et al.*, *Electrophoresis* 1999, *20*, 705–711.

[41] Peltier, J. B., Friso, G., Kalume, D. E., Roepstorff, P. *et al.*, *Plant Cell* 2000, *12*, 319–341.

[42] Schubert, M., Petersson, U. A., Haas, B. J., Funk, C. *et al.*, *J. Biol. Chem.* 2002, *277*, 8354–8365.

[43] Hippler, M., Klein, J., Fink, A., Allinger, T. *et al.*, *Plant J.* 2001, *28*, 595–606.

[44] Gomez, S. M., Nishio, J. N., Faull, K. F., Whitelegge, J. P., *Mol. Cell. Proteomics* 2002, *1*, 46–59.

[45] Seigneurin-Berny, D., Rolland, N., Garin, J., Joyard, J., *Plant J.* 1999, *19*, 217–228.

[46] Ferro, M., Seigneurin-Berny, D., Rolland, N., Chapel, A. *et al.*, *Electrophoresis* 2000, *21*, 3517–3526.

[47] Ferro, M., Salvi, D., Riviere-Rolland, H., Vermat, T. *et al.*, *Proc. Natl. Acad. Sci. USA* 2002, *99*, 11487–11492.

[48] Yamaguchi, K., von Knoblauch, K., Subramanian, A. R., *J. Biol. Chem.* 2000, *275*, 28455–28465.

[49] Yamaguchi, K., Subramanian, A. R., *J. Biol. Chem.* 2000, *275*, 28466–28482.

[50] Peltier, J. B., Ytterberg, J., Liberles, D. A., Roepstorff, P. *et al.*, *J. Biol. Chem.* 2001, *276*, 16318–16327.

[51] Kruft, V., Eubel, H., Jansch, L., Werhahn, W. *et al.*, *Plant Physiol.* 2001, *127*, 1694–1710.

[52] Millar, A. H., Sweetlove, L. J., Giege, P., Leaver, C. J., *Plant Physiol.* 2001, *127*, 1711–1727.

[53] Werhahn, W., Braun, H. P., *Electrophoresis* 2002, *23*, 640–646.

[54] Millar, A. H., Heazlewood, J. L., *Plant Physiol.* 2003, *131*, 443–453.

[55] Miras, S., Salvi, D., Ferro, M., Grunwald, D. *et al.*, *J. Biol. Chem.* 2002, *277*, 47770–47778.

[56] Whetten, R., Sun, Y. H., Zhang, Y., Sederoff, R., *Plant Mol. Biol.* 2001, *47*, 275–291.

[57] Groome, M. C. A. S., Giffort, D. J., *Physiol. Plant.* 1991, *83*, 99–106.

[58] Ekramoddoullah, A. K. M., Tan, Y., *Can. J. Plant Pathol. Rev. Can. Phytopathol.* 1998, *20*, 308–318.

[59] Costa, P., Bahrman, N., Frigerio, J. M., Kremer, A. *et al.*, *Plant Mol. Biol.* 1998, *38*, 587–596.

[60] Avila, C., Garcia-Gutierrez, A., Crespillo, R., Canovas, F. M., *Plant Physiol. Biochem.* 1998, *36*, 857–863.

[61] Bahrman, N., Petit, R. J., *J. Mol. Evol.* 1995, *41*, 231–237.

[62] Plomion, C., Bahrman, N., Durel, C. E., Omalley, D. M., *Heredity* 1995, *74*, 661–668.

[63] Saez, C. A., Munoz-Chapuli, R., Plomion, C., Frigerio, J. M. *et al.*, *FEBS Lett.* 2000, *477*, 237–243.

[64] Bradshaw, H. D., Ceulemans, R., Davis, J., Stettler, R., *J. Plant Growth Regul.* 2000, *19*, 306–313.

[65] Lev-Yadun, S., Sederoff, R., *J. Plant Growth Regul.* 2000, *19*, 290–305.

[66] Sterky, F., Regan, S., Karlsson, J., Hertzberg, M. *et al.*, *Proc. Natl. Acad. Sci. USA* 1998, *95*, 13330–13335.

[67] Allona, I., Quinn, M., Shoop, E., Swope, K. *et al.*, *Proc. Natl. Acad. Sci. USA* 1998, *95*, 9693–9698.

[68] Cantón, F. R., Garcia, V., Barré, A., Frigério, J.-M. *et al.*, in: Ritter, E., Espinel, G., Barrade, Y. (Eds.), *Sustainable Forestry, Woods products and Biotechnology*, DFA, Vitoria (in press).

[69] Hertzberg, M., Aspeborg, H., Schrader, J., Andersson, A. *et al.*, *Proc. Natl. Acad. Sci. USA* 2001, *98*, 14732–14737.

[70] Stasolla, C., van Zyl, L., Egertsdotter, U., Craig, D. *et al.*, *Plant Physiol.* 2003, *131*, 49–60.

[71] van Zyl, L., von Arnold, S., Bozhkov, P., Chen, Y. Z. *et al.*, *Compar. Funct. Genom.* 2002, *3*, 306–318.

[72] Heath, L. S., Ramakrishnan, N., Sederoff, R. R., Whetten, R. W. *et al.*, *Compar. Funct. Genom.* 2002, *3*, 226–243.

[73] Costa, P., Pionneau, C., Bauw, G., Dubos, C. *et al.*, *Electrophoresis* 1999, *20*, 1098–1108.

[74] Plomion, C., Pionneau, C., Brach, J., Costa, P. *et al.*, *Plant Physiol.* 2000, *123*, 959–969.

[75] Vander Mijnsbrugge, K., Meyermans, H., Van Montagu, M., Bauw, G. *et al.*, *Planta* 2000, *210*, 589–598.

[76] Chaffey, N., Cholewa, E., Regan, S., Sundberg, B., *Physiol. Plant.* 2002, *114*, 594–600.

[77] Hensel, M., Holden, D. W., *Microbiol.* 1996, *142*, 1049–1058.

[78] Konishi, H., Ishiguro, K., Komatsu, S., *Proteomics* 2001, *1*, 1162–1171.

[79] Rep, M., Dekker, H. L., Vossen, J. H., de Boer, A. D. *et al.*, *Plant Physiol.* 2002, *130*, 904–17.

[80] Rakwal, R., Komatsu, S., *Electrophoresis* 2000, *21*, 2492–2500.

[81] Lecourieux-Ouaked, F., Pugin, A., Lebrun-Garcia, A., *Mol. Plant Microbe Interact.* 2000, *13*, 821–829.

[82] Peck, S. C., Nuhse, T. S., Hess, D., Iglesias, A. *et al.*, *Plant Cell* 2001, *13*, 1467–1475.

[83] Mehta, A., Rosato, Y. B., *Proteomics* 2001, *1*, 1111–1118.

[84] Smolka, M. B., Martins, D., Winck, F. V., Santoro, C. E. *et al.*, *Proteomics* 2003, *3*, 224–237.

[85] Cyranoski, D., *Nature* 2001, *409*, 272.

[86] Journet, E. P., Carreau, V., Gouzy, J., Thoquet, P. *et al.*, *OCL–Oleagineux Corps Gras Lipides* 2001, *8*, 478–484.

[87] Frugoli, J., Harris, J., *The Plant Cell* 2001, *13*, 458–462.

[88] Oldroyd, G. E. D., Geurts, R., *Trends Plant Sci.* 2001, *6*, 552–553.

[89] Barker, D. G., Bianchi, S., Blondon, F., Dattée, Y. *et al.*, *Plant Molecular Biology Reporter* 1990, *8*, 40–49.

[90] Handberg, K., Stougaard, J., *Plant J.* 1992, *2*, 487–496.

[91] Kawasaki, S., Murakami, Y., *J. Plant Res.* 2000, *113*, 497–506.

[92] Bell, C., Dixon, R. A., Farmer, A. D., Flores, R. *et al.*, *Nucleic Acids Res.* 2001, *29*, 114–117.

[93] Quackenbush, J., Liang, F., Holt, I., Pertea, G. *et al.*, *Nucleic Acids Res.* 2000, *8*, 141–145.

[94] Journet, E. P., van Tuinen, D., Gouzy, J., Crespeau, H. *et al.*, *Nucleic Acids Res.* 2002, *30*, 5579–5592.

[95] Penmetsa, R. V., Cook, R. D., *Plant Physiol.* 2000, *123*, 1387–1397.

[96] Sagan, M., Morandi, D., Tarenghi, E., Duc, G., *Plant Sci.* 1995, *111*, 63–71.

[97] Borisov, A. Y., Jakobi, L., Lebsky, V., Morzhina, E. V. *et al.*, *Pisum Genet.* 1999, *3*, 40–44.

[98] Catoira, R., Galera, C., de Billy, F., Penmetsa, R. V. *et al.*, *Plant Cell* 2000, *12*, 1647–1665.

[99] Walker, S. A., Viprey, V., Downie, J. A., *Proc. Natl. Acad. Sci. USA* 2000, *97*, 13413–13418.

[100] Marsh, J. F., Schultze, M., *New Phytol.* 2001, *150*, 525–532.

[101] Van Rhijn, P., Vanderleyden, J., *Microbiol. Rev.* 1995, *59*, 124–142.

[102] Lhuissier, F. G. P., DeRuijter, N. C. A., Sieberer, B. J., Esseling, J. J. *et al.*, *Ann. Bot.* 2001, *87*, 289–302.

[103] Cook, D. R., *Curr. Opin. Plant Biol.* 1999, *2*, 301–304.

[104] Colebatch, G., Trevaskis, B., Udvardi, M., *New Phytol.* 2002, *153*, 37–42.

[105] Stougaard, J., *Curr. Opin. Plant Biol.* 2001, *4*, 328–35.

[106] Govers, F., Gloudemans, T., Moerman, M., van Kammen, A. *et al.*, *EMBO J.* 1985, *4*, 861–867.

[107] Natera, S. H. A., Guerreiro, N., Djordjevic, N. A., *Mol. Plant Microbe Interact.* 2000, *13*, 995–1009.

[108] Morris, A. C., Djordjevic, M. A., *Electrophoresis* 2001, *22*, 586–98.

[109] Mathesius, U., Keijzers, G., Natera, S. H. A., Weinman, J. J. *et al.*, *Proteomics* 2001, *1*, 1424–1440.

[110] Winzer, T., Bairl, A., Linder, M., Linder, D. *et al.*, *Mol. Plant Microbe Interact.* 1999, *12*, 218–226.

[111] Panter, S., Thomson, R., deBruxelles, G., Laver, D. *et al.*, *Mol. Plant Microbe Interact.* 2000, *13*, 325–333.

[112] Saalbach, G., Erik, P., Wienkoop, S., *Proteomics* 2002, *2*, 325–337.

[113] Hammad, Y., Marechal, J., Cournoyer, B., Normand, P. *et al.*, *Can. J. Bot.* 2001, *47*, 541–547.

[114] Hilbert, J. L., Costa, G., Martin, F., *Plant Physiol.* 1991, *97*, 977–984.

[115] Simoneau, P., Viemont, J. D., Moreau, J. C., Strulu, D. J., *New Phytol.* 1993, *124*, 495–504.

[116] Burgess, T., Laurent, P., Dell, B., Malajczuk, N. *et al.*, *Planta* 1995, *195*, 408–417.

[117] Burgess, T., Dell, B., *Can. J. Bot.* 1996, *74*, 553–560.

[118] Laurent, P., Voiblet, C., Tagu, D., de Carvalho, D. *et al.*, *Mol. Plant Microbe Interact.* 1999, *12*, 862–871.

[119] Niini, S. S., Tarkka, M. T., Raudaskoski, M., *Physiol. Plant.* 1996, *96*, 186–192.

[120] Timonen, S., Finlay, R., Söderström, B., Raudaskoski, M., *The New Phytologist* 1993, *124*, 83–92.

[121] Timonen, S., Söderström, B., Raudaskoski, M., *Mycorrhiza* 1996, *6*, 423–429.

[122] Tarkka, M. T., Nyman, T. A., Kalkinen, N., Raudaskoski, M., *Eur. J. Biochem.* 2000, *267*, 1–8.

[123] Tarkka, M. T., Nyman, T. A., Kalkinen, N., Raudaskoski, M., *Eur. J. Biochem.* 2001, *268*, 86–92.

[124] Vallorani, L., Bernardini, F., Sacconi, C., Pierleoni, R. *et al.*, *Electrophoresis* 2000, *21*, 3710–3716.

[125] Smith, S. E., Read, D. J., in: Smith, S. E., Read, D. J. (Eds.), *Mycorrhizal Symbiosis*, Academic Press, New York 1997.

[126] Garcia-Garrido, J. M., Toro, N., Ocampo, J. A., *Mycorrhiza* 1993, *2*, 175–177.

[127] Dumas-Gaudot, E., Guillaume, P., Tahiri-Alaoui, A., Gianinazzi-Pearson, V. *et al.*, *Mycorrhiza* 1994, *4*, 215–221.

[128] Simoneau, P., Louisy-Lois, N., Plenchette, C., Strullu, D. G., *AEM* 1994, *60*, 1810–1813.

[129] Samra, A., Dumas-Gaudot, E., Gianinazzi, S., *New Phytol.* 1997, *135*, 711–722.

[130] Dassi, B., Samra, A., Dumas-Gaudot, E., Gianinazzi, S., *Symbiosis* 1999, *26*, 65–77.

[131] Benabdellah, K., Azcon-Aguilar, C., Ferrol, N., *New Phytol.* 1998, *140*, 135–143.

[132] Benabdellah, K., AzconAguilar, C., Ferrol, N., *J. Exp. Bot.* 2000, *51*, 747–754.

[133] Fester, T., Kiess, M., Strack, D., *Mycorrhiza* 2002.

[134] Dumas-Gaudot, E., Bestel-Corre, G., Gianinazzi, S., in: Pandalai, S. G. (Ed.), *Recent Research Developments in Plant Biology*, Research Signpost: Trivandrum, India 2001, pp. 95–104.

[135] Bestel-Corre, G., Dumas-Gaudot, E., Poinsot, V., Dieu, M. *et al.*, *Electrophoresis* 2002, *23*, 122–137.

[136] Gianinazzi-Pearson, V., Dénarié, J., *Trends Plant Sci.* 1997, *2*, 371–372.

[137] Harrison, M. J., *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 1999, *50*, 361–389.

[138] Frühling, M., Roussel, H., Gianinazzi-Pearson, V., Puhler, A. *et al.*, *Mol. Plant Microbe Interact.* 1997, *10*, 124–131.

[139] Balestrini, R., Perotto, S., Gasverde, E., Dahiya, P. *et al.*, *Mol. Plant Microbe Interact.* 1999, *12*, 785–791.

[140] Bonfante, P., Genre, A., Faccio, A., Martini, I. *et al.*, *Mol. Plant Microbe Interact.* 2000, *13*, 1109–1120.

[141] Journet, E. P., ElGachtouli, N., Vernoud, V., deBilly, F. *et al.*, *Mol. Plant Microbe Interact.* 2001, *14*, 737–748.

[142] Werner, D., Barea, J. M., Brewin, N. J., Cooper, J. E. *et al.*, *Symbiosis* 2002, *32*, 83–104.

[143] Bestel-Corre, G., Dumas-Gaudot, E., Gianinazzi, S., *Proteomics* 2003, *3*, in press.

[144] Samra, A., Dumas-Gaudot, E., Gianinazzi-Pearson, V., Gianinazzi, S., *Agronomie* 1996, *16*, 709–719.

[145] Bestel-Corre, G., Dumas-Gaudot, E., St-Arnaud, M., Fontaine, B. *et al.*, *Mol Plant Microbe Interact.* 2003, in press.

[146] Boisson-Dernier, A., Chabaud, M., Garcia, F., Becard, G. *et al.*, *Mol. Plant Microbe Interact.* 2001, *14*, 695–700.

[147] Bestel-Corre, G., Dumas-Gaudot, E., Hartmann, A., Gianinazzi, S., *AEM* 2003, in press.

[148] Bestel-Corre, G., Dumas-Gaudot, E., Gianinazzi, S., *Agronomie* 2003, in press.

[149] Repetto, O., Bestel-Corre, G., Dumas-Gaudot, E., Berta, G. *et al.*, *New Phytol.* 2002, 555–567.

[150] Molloy, M. P., *Anal. Biochem.* 2000, *280*, 1–10.

[151] Rabilloud, T., *Electrophoresis* 1998, *19*, 758–760.

[152] Santoni, V., Kieffer, S., Desclaux, D., Masson, F. *et al.*, *Electrophoresis* 2000, *21*, 3329–3344.

[153] Santoni, V., Molloy, M., Rabilloud, T., *Electrophoresis* 2000, *21*, 1054–1070.

[154] Wienkoop, S., Saalbach, G., *Plant Physiol.* 2003, *131*, 1080–90.

[155] Holtorf, H., Guitton, M. -C., Reski, R., *Naturwissenschaften* 2002, *89*, 235–249.

[156] Watson, B. S., Asirvatham, V. S., Wang, L., Sumner, L. W., *Plant Physiol.* 2003, *131*, 1104–1123.

[157] Tyers, M., Mann, M., *Nature* 2003, *422*, 193–197.

[158] Aebersold, R., Mann, M., *Nature* 2003, *422*, 198–207.

[159] Phizicky, E., Bastiaens, P. I. H., Zhu, H., Snyder, M. *et al.*, *Nature* 2003, *422*, 208–215.

[160] Giavalisco, P., Nordhoff, E., Lehrach, H., Gobom, J. *et al.*, *Electrophoresis* 2003, *24*, 207–216.

[161] Herbert, B., Righetti, P. G., *Electrophoresis* 2000, *21*, 3639–3648.

[162] Wildgruber, R., Harder, A., Obermaier, C., Boguth, G. *et al.*, *Electrophoresis* 2000, *21*, 2610–2616.

[163] Gorg, A., Boguth, G., Kopf, A., Reil, G. *et al.*, *Proteomics* 2002, *2*, 1652–1657.

[164] Hoving, S., Gerrits, B., Voshol, H., Müller, D. *et al.*, *Proteomics* 2002, *2*, 127–134.

[165] Patton, W. F., *J. Chromatogr. B* 2002, *771*, 3–31.

[166] Tonge, R., Shaw, J., Middleton, B., Rowlinson, R. *et al.*, *Proteomics* 2001, *1*, 377–396.

[167] Alban, A., David, S. O., Bjorkesten, L., Andersson, C. *et al.*, *Proteomics* 2003, *3*, 36–44.

[168] Oda, Y., Huang, K., Cross, F. R., Cowburn, D. *et al.*, *Proc. Natl. Acad. Sci. USA* 1999, *96*, 6591–6596.

[169] Butt, A., Davison, M. D., Smith, G. J., Young, J. A. *et al.*, *Proteomics* 2001, *1*, 42–53.

[170] Gygi, S. P., Rist, B., Aebersold, R., *Curr. Opin. Biotechnol.* 2000, *11*, 396–401.

[171] Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J. *et al.*, *Nat. Biotechnol.* 2002, *20*, 301–305.

[172] Link, A. J., Eng, J., Schieltz, D. M., Carmack, E. *et al.*, *Nat. Biotechnol.* 1999, *17*, 676–682.

[173] Islam, N., Tsujimoto, H., Hirano, H., *Proteomics* 2003, *3*, 307–316.

[174] Fields, S., Song, O., *Nature* 1989, *340*, 235–246.

[175] Gavin, A. C., Bosche, M., Krause, R., Grandi, P. *et al.*, *Nature* 2002, *415*, 141–147.

[176] Forler, D., Kocher, T., Rode, M., Gentzel, M. *et al.*, *Nat. Biotechnol.* 2003, *21*, 89–92.

