Eur J Plant Pathol (2011) 129:539–554
DOI 10.1007/s10658-010-9718-6

Using metabolic profiling to assess plant-pathogen interactions: an example using rice (*Oryza sativa)* and the blast pathogen *Magnaporthe grisea*

Oliver A. H. Jones · Mahon L. Maguire · Julian L. Griffin · Young-Ho Jung · Junko Shibato · Randeep Rakwal · Ganesh K. Agrawal · Nam-Soo Jwa

Accepted: 15 November 2010 / Published online: 30 November 2010

Abstract A metabolomics based approach has been used to study the infection of the Hwacheong rice cultivar (*Oryza sativa* L. cv. Hwacheong) with compatible (KJ201) and incompatible (KJ401) strains of the rice blast fungal pathogen *Magnaporthe grisea*. The metabolic response of the rice plants to each strain was assessed 0, 6, 12, 24, 36, and 48 h post inoculation. Nuclear Magnetic Resonance (NMR) spectroscopy and Gas and Liquid Chromatography Tandem Mass spectrometry (GC/LC-MS/MS) were used to study both aqueous and organic phase metabolites, collectively resulting in the identification of 93 compounds. Clear metabolic profiles were observed at each time point but there were no significant differences in the metabolic response elicited by each pathogen strain until 24 h post inoculation. The largest change was found to be in alanine, which was ~30% (±9%) higher in the leaves from the compatible, compared to the resistant, plants. Together with several other metabolites (malate,

O. A. H. Jones · J. L. Griffin
The Hopkins Building, Department of Biochemistry,
University of Cambridge,
Tennis Court Road,
Cambridge CB2 1QW, UK

M. L. Maguire
BHF Magnetic Resonance Unit,
Wellcome Trust Centre for Human Genetics,
University of Oxford,
Roosevelt Drive,
Oxford OX3 7BN, UK

Y.-H. Jung · N.-S. Jwa
Department of Molecular Biology,
College of Life Sciences, Sejong University,
Seoul 143-747, South Korea

Y.-H. Jung
Analysis Department, General Analysis Division,
Biotoxtech Co., Ltd.,
Ochang Scientific Industrial Complex,
686-2 Yangcheong-ri,
Ochang-eup, Chungcheongbuk-do 363-883, South Korea

J. Shibato · R. Rakwal
Health Technology Research Centre (HTRC),
National Institute of Advanced Industrial Science
and Technology (AIST),
16-1 Onogawa,
Tsukuba 305-8569 Ibaraki, Japan

R. Rakwal · G. K. Agrawal
Research Laboratory for Biotechnology and Biochemistry
(RLABB),
GPO Box 8207, Kathmandu, Nepal

Present Address:
O. A. H. Jones (✉)
The Christopherson Building,
School of Engineering & Computing Sciences,
University of Durham,
South Road,
Durham DH1 3LE, UK
e-mail: oliver.jones@durham.ac.uk

glutamine, proline, cinnamate and an unknown sugar) alanine exhibited a good correlation between time of fungal penetration into the leaf and the divergence of metabolite profiles in each interaction. The results indicate both that a wide range of metabolites can be identified in rice leaves and that metabolomics has potential for the study of biochemical changes in plant-pathogen interactions.

Keywords Metabolomics · Metabonomics · NMR spectroscopy · Chromatography · Mass spectrometry

Introduction

Rice blast (caused by the fungal pathogen *Magnaporthe grisea*) is an extremely virulent plant disease which causes tremendous production losses in all rice growing areas of the world (Talbot 1995). *M. grisea* is one of the most damaging pathogens of cultivated rice and is capable of infecting many related cereals including; barley, wheat, rye and millet (as well as over fifty other species of grasses and sedges). It therefore poses a considerable threat to many of the world's most important food crops and has emerged as a model system for investigation of foliar pathogenicity. Interestingly however, despite its apparently broad range, most strains of *M. grisea* are only able to infect a few host species and the majority of strains isolated from rice are only able to infect a limited number of cultivars (Talbot 2003).

In addition to their economic importance the recent release of the complete draft genome sequences of both the pathogen and host makes this system an excellent model for the investigation of plant-pathogen interactions (Goff et al. 2002; Yu et al. 2002; Dean et al. 2005). During infection of the host plant the fungus undergoes a series of well-characterised developmental steps, including the formation of elaborate penetration structures (appressoria). Recent studies have focussed on studying this process at the transcriptional, proteomic and even the biophysical level (Bechinger et al. 1999). However, to our knowledge, the infection process has yet to be investigated using a metabolomics based approach.

Metabolomics itself may be defined as the study of the small, endogenous molecules (such as sugars, organic acids, amino acids and nucleotides) (Fiehn 2002). These compounds are the substrates and by-products of cell processes (such as enzymatic reactions) and, as such have a direct effect on phenotype. The sub-discipline of plant metabolomics is a growing research area. Indeed, some of the most significant developments in metabolic profiling have been made in this field and metabolomics based work has been conducted on a diverse array of important food plant species; for a review see Schauer and Fernie (2006). Metabolomics is useful in plant studies since it offers the ability to identify biochemical changes relatively quickly, usually before any overt phenotypic changes become apparent. This increases the likelihood of successful intervention (e.g. early treatment of infected plants with herbicides). In addition, in those cases where measurable phenotypic changes are observed, metabolomics can be used to decipher, possibly even elucidate, the biochemical causes and/or consequences of such changes (Griffin and Shore 2007).

Pathogen-host relationships are extremely interesting in terms of both biological importance and metabolite richness and thus an ideal area for exploration using metabolomics and related -omic techniques. Indeed, considering the prominence of research into this area within plant biology as a whole, it is surprising that (to date) so few metabolomic studies have focussed in this area. Those studies which have been conducted are well reviewed in Allwood et al. (2008). In the present work metabolomics has been utilized to study the response of primary metabolites in relation to the infection process of rice (*Oryza sativa* L. cv. Hwacheong) with the KJ201 (compatible) and KJ401 (incompatible) strains of *M. grisea*. Infection of rice with the KJ201 *M. grisea* strain leads to the eventual death of the plant, whereas plants infected with the KJ401 strain survive with lesions on the leaves but are otherwise in general good health.

Clear differences in metabolic profiles were observed at each time point in the aqueous (but not the organic) phase metabolites and there were also notable differences in the response elicited by each pathogen strain 24 h post inoculation. Interestingly this time point corresponds with the time the fungus first successfully penetrates the leaf tissue during infection.

Materials and methods

Rice cultivation

Factors such as cultivar and growth conditions (e.g. light and temperature) can have a large effect on metabolite

levels in plants. Thus extreme care was taken to reproduce natural (but controlled) conditions as far as possible and to ensure the differences between sample groups were limited to only the *M. grisea* strain used. All rice plants were of the Hwacheong cultivar. Fresh seeds were planted into pots measuring 24 by 34 by 24 cm. The pots were then placed in a growth chamber and maintained at 28°C at 70% humidity with a 16/8 h light dark cycle for three weeks. After three weeks the cultivars were inoculated with either the KJ201 (compatible interaction) or the KJ401 (incompatible interaction) strain of *M. grisea* at a concentration 5×10^5 pathogen spores/mL. Samples of leaf (4th) tissue (the main point of infection for rice blast) were taken at 0 h (pre inoculation) and 6, 12, 24, 36 and 48 h post inoculation. These time points were specifically chosen in order to give a good spread of data points over the infection period of this particular pathogen; different plant/pathogen interactions may require an alternative set up to reflect the different infection profile. There were three replicates at each point for both the compatible and incompatible pathogens. After harvest each leaf was quickly (<30 s) placed in liquid nitrogen and ground to a fine powder with a mortar and pestle. All samples were stored at −80°C prior to analysis.

Metabolite extraction

Metabolites were extracted from a 100 mg sub sample of the crushed leaf tissue using a methanol-water-chloroform extraction (Le Belle et al. 2002). The choice of extraction method is an important factor in any metabolomics study. A detailed comparison of extraction methods was conducted recently and the methanol-water-chloroform extraction was found to be the most suitable approach in the majority of cases (Lin et al. 2007). A 100 mg sample of tissue has also been shown to be sufficient to ensure good coverage of the metabolome.

Powdered tissue samples (100 mg) were first mixed with 600 μl of a methanol-chloroform mixture (2:1 *v*:*v*) and sonicated for 15 min. After this time 300 μl of chloroform and 300 μl of ultrapure (18.2 MΩ) water (Millipore, Watford, UK) were added. The samples were then centrifuged for 20 min at 13,200 rpm. The resultant organic and aqueous fractions, as well as the protein pellet, were then transferred to separate Eppendorf tubes. Both fractions were analysed by NMR and GC\LC-MS.

^{1}H NMR analysis

All NMR based experiments were performed using a Bruker NMR spectrometer at 11.7 T (^{1}H frequency of 500.3 MHz) using a 5 mm ATMA TXI probe and an Avance II+console (Bruker BioSpin, Rheinstetten, Germany). Spectra were acquired using a 1D NOESY pulse sequence with water and chloroform presaturation used for the aqueous and organic phases respectively.

Aqueous extracts from the metabolite extraction step were dried down in a Concentrator 5301 evacuated centrifuge (Eppendorf, Histon, UK). They were then dissolved in 600 μL of D_2O and buffered in 0.24 M sodium phosphate (pH 7.0) containing 1 mM sodium-3-(trimethylsilyl)-2,2,3,3-tetradeuteriopropionate (TSP; Cambridge Isotope Laboratories, Andover, MA) which provided a chemical shift reference (0 ppm) for the resulting spectra. Organic extracts were dried down in air in a fume hood and then made up in 600 μl of a 10/90% CDCl3/CHCl3 mix (Goss Scientific Instruments, Great Baddow, UK), with solvent suppression on the chloroform peak which again acted as a chemical shift reference (7.26 ppm) for the resulting spectra.

For all samples 128 transients were collected into 64 k data points using a spectral width of 8.00 kHz (20 ppm) and an acquisition time of 1.95 s per FID. The resulting one-dimensional spectra were processed using a macro written within the ACD NMR manager software (Advanced Chemistry Development Inc., Toronto, Canada). FIDs were multiplied by an exponential weighting function equivalent to 1 Hz line broadening and Fourier-transformed from the time to frequency domain. They were then manually phased, baseline-corrected, and referenced as stated above. The region around the major solvent peak in each phase was excluded from further analysis to avoid the residual H_2O/chloroform signals (which represented highly variable regions in the spectra). Areas of the spectrum less than 0 and greater than 10 ppm were also removed since they contained no useful data.

The spectra were processed with NMR Suite Professional, version 5.1 (Chenomx, Alberta, Canada) which is capable of identifying and quantifying individual compounds based on their respective spectral signatures. Metabolites were identified using the library and peak fitting routine contained within the Chenomx software in conjunction with two dimensional NMR spectroscopy as well as with reference to previous literature (Sato et al. 2004;

Sobolev et al. 2005) and the online Madison Biological Magnetic Resonance Data Bank (http://www.bmrb.wisc.edu/metabolomics/).

GC-MS analysis

Metabolites in the aqueous phase were derivatised via methoximation and trimethylsilylation to enable the analysis of compounds such as amino acids, sugars, sugar alcohols, and aromatic acids (Gullberg et al. 2004). A 150 µl sample of the aqueous extract previously analysed by ^{1}H NMR spectroscopy was first evaporated to dryness in an evacuated centrifuge. A 30 µl aliquot of methoxyamine hydrochloride (20 mg per ml in pyridine, Sigma-Aldrich, Gillingham, UK) was then added. The sample was then vortex mixed for 30 s and then left for 17 h at room temperature. After this time, 30 µl of N-methyl-N-trimethylsilyltrifluoroacetamide (MSTFA—Sigma-Aldrich) was added. The samples were again vortex mixed for 30 s and then left to react for 1 h at room temperature. All samples were made up to a volume of 600 µl with hexane prior to analysis via GC-MS. GC analysis was carried out using a ZB-5MS column (Phenomenex, Cheshire, UK—30 m×0.25 mm ID×0.25 µm). The oven temperature program started at 70°C and was then increased at 10°C per minute to 130°C, then by 5°C a minute to a temperature of 230°C, and then by 20°C a minute to a final temperature of 310°C which was held for 5 min. The total run time was ~34 min per sample.

In the lower phase the concentration of plant pigments was too high to enable the analysis of all samples without damage to the GC-MS and using column chromatography to remove the pigments was also found to significantly affect the type and relative abundance of the metabolites observed in the final extract. However, 3 test samples from both the incompatible and compatible interactions were run with the aim of helping to identify some of the peaks observed in the NMR spectra. For this whole (600 µl) samples of the organic phase from the NMR analysis were derivatised via acid-catalysed esterification using Boron Trifloride (BF_3). This allowed the analysis of compounds such as fatty acids, sterols, and aliphatic compounds (Morrison and Smith 1964). The organic layer from the extraction process was first dried down in a fume hood and reconstituted in 750 µl of a chloroform/methanol mixture (1:1 *v*/*v*). A 250 µl aliquot of 10% BF_3 in methanol (Sigma-Aldrich) was then added to each sample and the vials sealed and incubated at 80°C for 90 min. Each vial was then allowed to cool to room temperature before 300 µl of water (ultrapure 18.2 mΩ) and 500 µl of hexane were added and the sample vortex mixed for 1 min. The lower aqueous layer was discarded and the remaining organic layer was evaporated to dryness and reconstituted in 200 µl of hexane. GC analysis was carried out using a TR-FAME column (Thermo Fisher Scientific—30 m×0.25 mm ID×0.25 µm). The GC oven temperature program was 60°C held for 2 min, then increased at, 10°C per minute to 150°C and then 4°C a minute to a final temperature of 230°C which was held for 7 min. The total run time was ~35 min per sample.

In both cases the GC column eluent was introduced into the Thermo Fisher Scientific DSQ quadrupole mass spectrometer (transfer line temperature 340°C for aqueous metabolites, 240°C for FAMEs), ion source temperature 250°C, electron beam=70 eV, source current=100 µA in all cases. The detector was turned on after a solvent delay of 4 min and data were collected in full scan mode using 3 scans s^{-1} across a mass range of 50–650 m/z.

All GC-MS chromatograms were analysed using the Xcalibur processing software package, version 1.4 (Thermo Fisher Corporation). Compound identification was performed by matching the mass spectra and retention time of each compound with those obtained from standards run in house and using data from the Max Planck Institute of Molecular Plant Physiology online database (http://www-en.mpimp-golm.mpg.de/02-instUeberInstitut/04-instRessources/webbasedRsrc/metaboliteMSL/index.html) and the National Institute of Standards and Technology (NIST) database (2002 edition) using a similarity score cut off of >60%; retention time matching to known standards was also undertaken. A 30 s threshold window was used for the deviation of peaks away from the predicted retention time.

LC-MS analysis

For the LC-MS analysis of the organic phase was dried down and the lipid extracts reconstituted in 600 µl of a 2:1 methanol/chloroform mix. Analysis was carried out using a UPLC Acquity™ (Waters, Milford, USA) with a C8 column (C8 2.1×100 mm, 1.7 µm particle size, Waters). The mobile phase consisted of water, with 0.1% formic acid and 10 mM ammonium acetate (phase A)

and acetonitrile:isopropanol (5:2) with 0.1% formic acid and 10 mM ammonium acetate (phase B). The analytes were eluted from the column over 12 min. The gradient used was 60 to 100% B followed by 2 min at 100% B and 2 min of column re-equilibration at the starting condition (60% B). In all experiments the eluent was then injected onto a Q-ToF micro (Waters) with a 1:1 split operated in the positive mode. In order to ensure consistent performance of the system a pooled sample followed by a blank were injected after every 15 sample injections. MassLynx software (Waters) was used to process the resulting data on retention time and mass. All variables with missing values in more than 10 spectra were discarded along with all signals of low intensity in the pooled sample spectra, since relative quantification becomes unreliable with low intensity peaks. The remaining data were then normalized and exported to SIMCA-P (Umetrics) for further analysis. However, since no significant differences between sample groups were found these data were not analyzed any further.

Multivariate statistical analysis

PCA analysis Datasets from both analytical methods were imported into SIMCA-P version 11.0 (Umetrics, Umeå, Sweden) for processing. GC-MS data was scaled to unit variance by dividing each variable by $1/(S_d)$ where S_d is the standard deviation of the variable. Since it contained more noise than the GC-MS data the ^{1}H NMR spectroscopy data was Pareto scaled so that each variable was centred and multiplied by $1/(S_d)^{1/2}$. This increases the influence of low concentration metabolites without significant amplification of noise and hence is a more suitable type of scaling for this type of data. Samples were analysed primarily using Principal Components Analysis (PCA) and Partial Least Squares-Discriminate analysis. Identification of the major metabolic perturbations within the pattern recognition models was achieved by analysis of corresponding loadings plots, which showed the metabolites considered to have a significant contribution to a particular PC ($p<0.05$). Additionally, R^2 and Q^2 values calculated by the SIMCA software were used as measures for the robustness of a pattern recognition model. R^2 is the fraction of variance explained by a component and Q^2 is an indication of the predictive power of the model (Eriksson et al. 2001). Both values may vary between 0 and 1, with values closer to 1 indicating a more reliable model. All models were also validated via random permutation of the X matrix.

O-PLS STOCSY

Statistical total correlation spectroscopy (STOCSY) takes advantage of the multi-co-linearity of the intensity variables in a dataset (in this case ^{1}H-NMR spectra and GC-MS chromatograms). A combination of STOCSY with supervised pattern recognition, particularly orthogonal projection on latent structure analysis (O-PLS), offers a powerful framework for analysis of metabonomic data (Weljie et al. 2007). O-PLS is used to extract data related to discrimination between groups and this information is then cross-combined with the STOCSY results to give a graphical display which is coloured to represent the value of the correlation coefficient between the peak and the factor regressed against (in this case time). This in turn helps to identify the molecules responsible for the variation. We applied this technique to the NMR and to GC-MS data (the first time this has been tried in the latter case). In the case of NMR data, the processed spectra were first exported to the Chenomx NMR Suite software where they were reprocessed into 0.001 ppm bins (an option not available in the ACD software). The intensity of the spectra in each bin was used as a variable; these data were used to build O-PLS models and also for the STOCSY analysis using custom written MATLAB code based on the methods outlined in Cloarec et al. (2005). For the GC-MS analysis each chromatogram was exported as a text file containing time and intensity information. These data was then used for further analysis in the same way as the bucketed NMR data.

Batch processing

The application of batch processing focused on the differences in plant responses to the infectious and non-infections strains of a pathogen over time. It was undertaken using the data matrices from the PCA of the NMR and GC-MS aqueous phase datasets. These were incorporated into SIMCA as before but using time as the batch variable. The mean of the three samples from each group at each time point was taken and samples from the incompatible reactions were

labelled as the control/normal batch. Confidence limits (±2 standard deviations) were used to detect deviation from the control model data from the compatible\infectious interaction when these data were mapped onto the incompatible\non infectious models.

Results

Experimental strategy

The experimental strategy for this study is presented in Fig. 1. Briefly, rice plants were exposed to each strain of *M. grisea*. Inoculation was confirmed using the defence-related molecular gene markers *OsPR1b* and *OsPR10a* and the infection process of both strains was observed via the use of a GFP tagged blast pathogen. The metabolic response of the rice plants to each strain was assessed 0, 6, 12, 24, 36, and 48 h post inoculation. In this study the use of the molecular gene markers was limited to confirming successful/unsuccessful infection. No attempt was made in this particular work to assign a detailed function to these genes (although further work is being undertaken in this area).

Previous plant-pathogen studies have mainly used ^{1}H Nuclear Magnetic Resonance (NMR) spectroscopy as

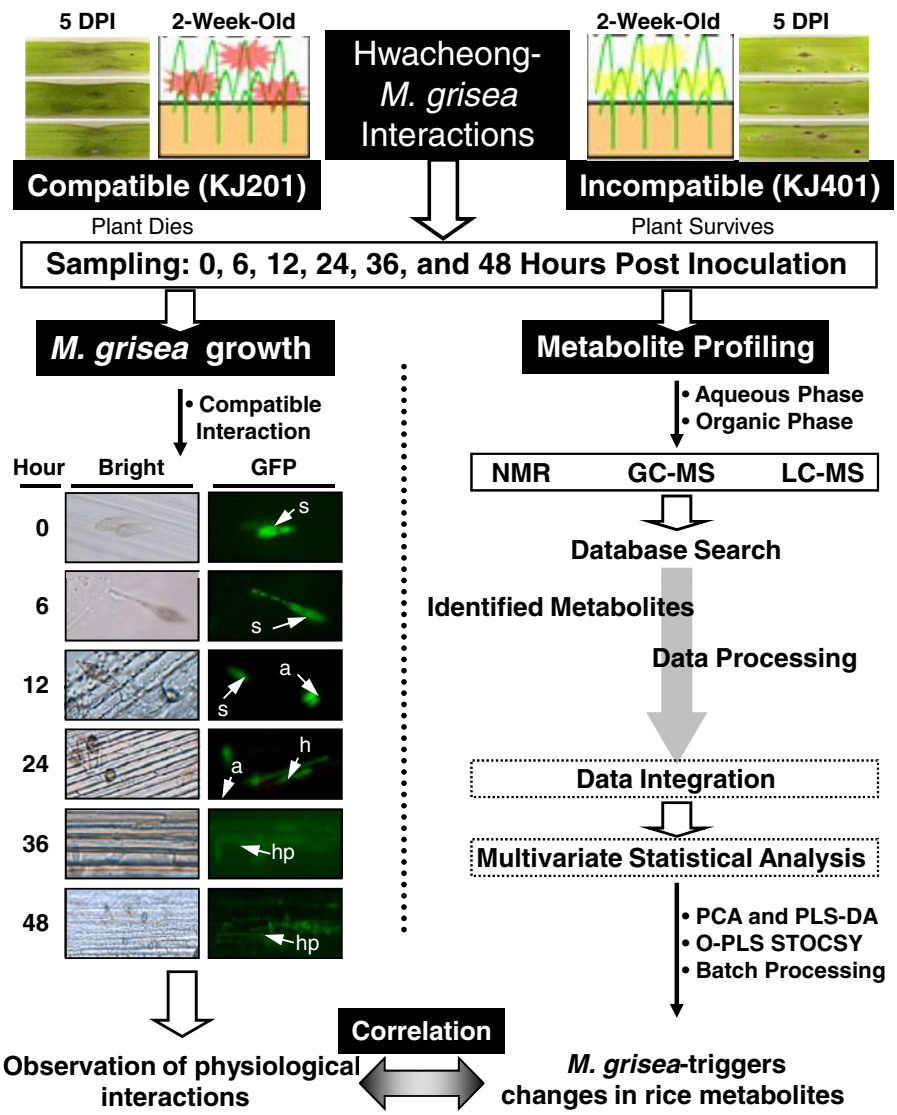

Fig. 1 The experimental strategy used in this study. Leaf segments are shown 5 days post inoculation (DPI). Metabolite profiling (right-hand side) of aqueous and organic phases using three techniques (NMR, GC-MS and LC-MS) were used to assess the metabolic changes in each case. Acquired data were searched, integrated and analyzed as discussed in the experimental section. Abbreviations used are as follows: s, spore / a, appressoria / h, haustoria / hp, hyphae; all images were taken using fluorescence microscopy

a diagnostic tool. However, while useful for the rapid analysis and investigation of metabolic variation, NMR is relatively insensitive compared to mass spectrometry based methods. Consequently a combined approach is needed to obtain as full a picture as possible. To take this into account NMR and Gas and Liquid Chromatography Tandem Mass spectrometry (GC/LC–MS/MS) were all used to study both aqueous and organic phase metabolites in leaf extracts in this study.

Metabolite identification—NMR

Figure 2 shows high resolution ^{1}H NMR spectra (acquired in ~7 min) of the aqueous phase metabolites extracted from *O. sativa*.

NMR analysis did not identify as many metabolites in the aqueous or organic phases as GC-MS, although each method did identify a slightly different set of compounds. The NMR organic spectra were dominated by lipid side chains, mainly from linoleic and hexanoic acid. There were also numerous small peaks between 3 and 5 ppm and 8 and 10 ppm from low concentration metabolites which could not be identified. In total, 56 metabolites were identified using NMR and these are listed in Table 1. All the metabolites listed were found in both the compatible and incompatible interactions. For the interested reader a good example of NMR analysis of organic phase metabolites (using lettuce leaves) is given in Sobolev et al. (2005).

Metabolite identification—GC-MS

Figure 3 shows GC-MS chromatograms of the aqueous (panel A) and organic phase (panel B) of rice extracts respectively. The GC-MS analysis

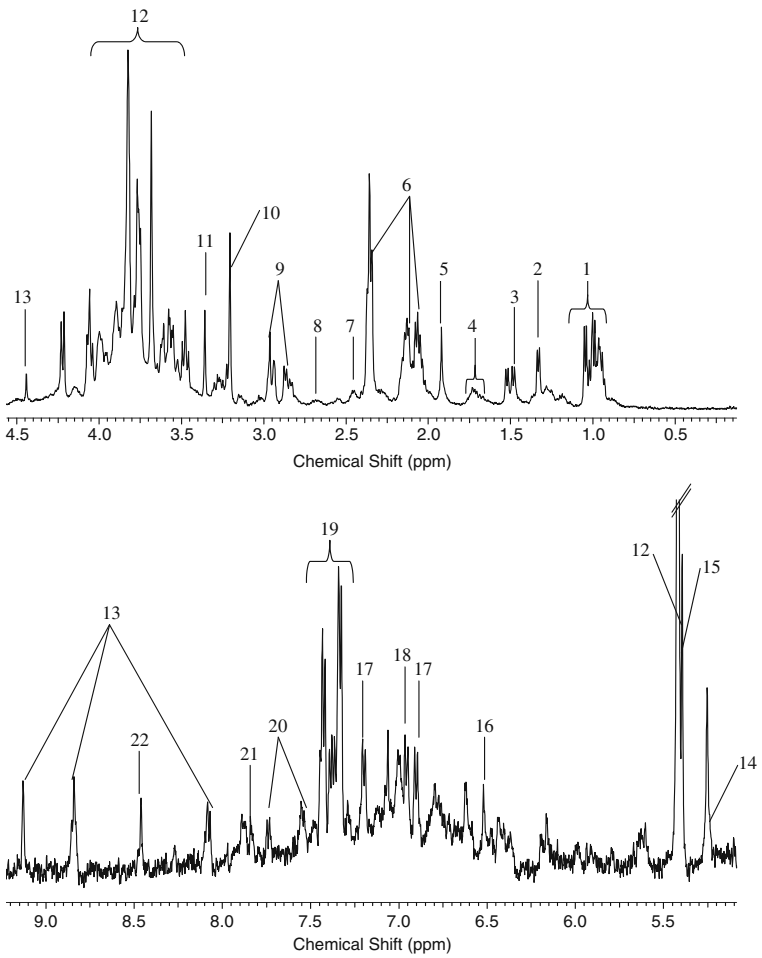

Fig. 2 NMR spectra of aqueous phase metabolites. The vertical scale on the lower panel has been increased relative to the upper panel to aid visual interpretation. Key: 1) Overlapping resonances from Leucine, Isoleucine and Valine, 2) Lactate, 3) Alanine, 4) Lysine, 5) Acetate, 6) Glutamate, 7) Glutamine, 8) Methionine, 9) Asparagine, 10) O-Acetyl carnosine, 11) GlyceroPhosphoCholine, 12) Overlapping sugar resonances (mostly Sucrose, Fructose and Glucose), 13) Trigonelline, 14) Alpha glucose, 15) Allantoin, 16) Fumarate, 17) Tyrosine, 18) NADPH, 19) Phenyl alanine, 20) Tryptophan, 21) Cytidine, 22) Formate

Table 1 List of identified metabolites from rice leaf extracts

Metabolite name	Identified with	
	NMR	GC-MS
Aqueous Phase		
Amino acids, amines and purine bases		
Alanine	Y	Y
Asparagine	Y	Y
Adenine	Y	N
Adenosine	Y	N
Beta-Alanine	Y	Y
Carnosine	Y	N
Creatine	Y	N
Cytidine	Y	N
Citrulline	Y	N
Cystathionine	N	Y
Cysteine	Y	Y
Glutamate	Y	Y
Glutamine	Y	Y
Glycine	Y	Y
Guanosine	Y	N
Histidine	Y	Y
Isoleucine	Y	Y
Leucine	Y	Y
Lysine	Y	Y
Methionine	Y	N
Phenylalanine	Y	Y
Proline	Y	Y
Serine	Y	Y
Threonine	Y	Y
Tryptophan	Y	N
Tyrosine	Y	Y
Valine	Y	Y
O-Acetylcarnitine	Y	N
Nucleotides and Coenzymes		
ADP	Y	N
AMP	Y	N
GTP	Y	N
NADP+	Y	N
NADPH	Y	N
Organic and carboxylic acids		
Acetate	Y	N
Allantoin	Y	N
Citrate	Y	N
Cyclohexanecarboxylic acid	N	Y
Formate	Y	N
Fumarate	Y	Y
Galactaric acid	N	Y
Galactonic acid	N	Y
Gluconic acid	N	Y
Glyceric acid	N	Y
Isocitrate	Y	Y
Lactate	Y	Y
Malate	Y	Y
N-Acetylglutamate	Y	N
Oxalic acid	N	Y
Succinate	Y	Y
Sugars		
Alpha-glycopyranose	N	Y
Betaglucopyranose	N	Y
Fructose	Y	Y
Galactose	Y	N
Glucitol (possibly from *M. grisea*)	N	Y
Glucose	Y	Y
Mannose	Y	N
Melibiose	N	Y
Ribose	Y	Y
Ribitol	N	Y
Sucrose	Y	Y
Xylitol	N	Y
1,3-Dihydroxyacetone	Y	N
Others		
Betaine	Y	N
Myo-Inositol	Y	Y
Threonate	Y	Y
Trigonelline	Y	N
Trimethylamine N-oxide	Y	N
Organic Phase		
Fatty Acids		
11-Eicosenoic (Gondoic) acid	N	Y
8,11 octadecadienoic (Linoleic) acid	Y	Y
11 octadecanoic (Vaccenic) acid	N	Y
9-hexadecanoic (Palmitoleic) acid	N	Y
5,8,11,14-Eicosatetraenoic (Arachidonic) acid	N	Y
Benzenepropanoic (Rosmarinic) acid	N	Y
Butanoic (Butyric) acid	N	Y
Benzoic (Dracylic) acid	N	Y
Hexanoic (Caproic) acid	Y	Y
Docosanoic (Behenic) acid	N	Y

Table 1 (continued)

Metabolite name	Identified with	
	NMR	GC-MS
Dodecanoic (Lauric) acid	N	Y
Eicosadienoic (Gamma linoleic) acid	Y	Y
Eicosanoic (Arachidic) acid	N	Y
Heptadecanoic (Margaric) acid	N	Y
Hexadecanoic (Palmitic) acid	N	Y
9,12,15-octadecatrienoic acid (Alpha linolenic) acid	N	Y
Tetradecanoic (Myristic) acid	N	Y
Octadecanoic (Steric) acid	N	Y
9-Octadecenoic acid (Oleic) acid	N	Y
Pentadecanoic (Isopalmitic) acid	N	Y
Stearidonic (Moroctic) acid	N	Y
Plant Sterols		
Alpha Sitosterol	N	Y
Beta Sitosterol	N	Y
Stigmasterol	N	Y
Alpha Tocopherol (Vitamin E)	N	Y
Mono-stearin	N	Y

identified a total of 67 metabolites with some overlap with the NMR.

Metabolite identification—LC-MS

NMR analysis of the organic phase metabolites indicated no significant differences between sample groups. However, NMR is not routinely used to study organic phase metabolites and is not as sensitive as mass spectrometry based approaches. Therefore LC-MS analysis was also used to study the organic phase metabolites. LC-MS was not used for identification purposes however, since this is difficult using such techniques and the GC-MS and NMR results had already provided such data.

Comparison with previous studies

The combination of NMR and GC-MS identified a total of 93 metabolites (see Table 1 for details). In comparison, a study by Sato et al. (2004) reported the use of CE-MS and capillary electrophoresis diode array detection (CE-DAD) to detect 88 water soluble intracellular metabolites in rice leaves. The authors stated that NMR and LC-MS required a larger sample quantity for analysis than was necessary for CE-MS while sample derivatisation, a prerequisite for analysis via GC-MS, significantly increased the sample preparation time and cost. In fact, as demonstrated here, NMR analysis detects a large number of compounds extremely rapidly, using a sample size of only 100 mg (the same sample size used in the Sato et al. study). Both GC-MS and LC-MS can also give detailed results with substantially smaller sample sizes and all three techniques are highly reproducible. While GC-MS does require some form of sample derivatisation this is not prohibitive to analysis. Additionally, newer techniques such as GC-TOF-(Time Of Flight)-MS analysis can theoretically identify even greater numbers of metabolites (Bölling and Fiehn 2005). Using a combined analytical approach meant that more metabolites were measured than the Sato et al. study which did not look at organic phase metabolites due to the difficulties of analysing these compounds via CE-MS.

Metabolic Dynamics—PCA analysis of the aqueous phase

The Principle Component Analysis (PCA) of the aqueous metabolites for the incompatible and compatible interactions using the NMR dataset is shown in Fig. 4. Each time point in each of the two experiments was treated as a separate group for this analysis and the Q^2 for the models were >0.5 indicating good separation between groups. However, the time points from both interactions tend to cluster together indicating a strong time trend in the data. The GC-MS data from the aqueous phase metabolites (data not shown) followed a similar trend to the NMR in that separation could be seen between each time point (with high Q^2 values) but it was harder to find separation between the compatible and incompatible interactions at the same time point.

The observation that plant metabolite levels fluctuate with time is perhaps not surprising. Nevertheless this result means that it is hard to distinguish time related changes from those caused by infection with *M. grisea* using PCA analysis alone. A notable exception is at the 24 h time point. Here, a pairwise comparison between the incompatible and compatible interactions using both NMR and GC-MS data did produce PCA models with good predictability and a high Q^2 values of >0.6. Interestingly, 24 h is the time

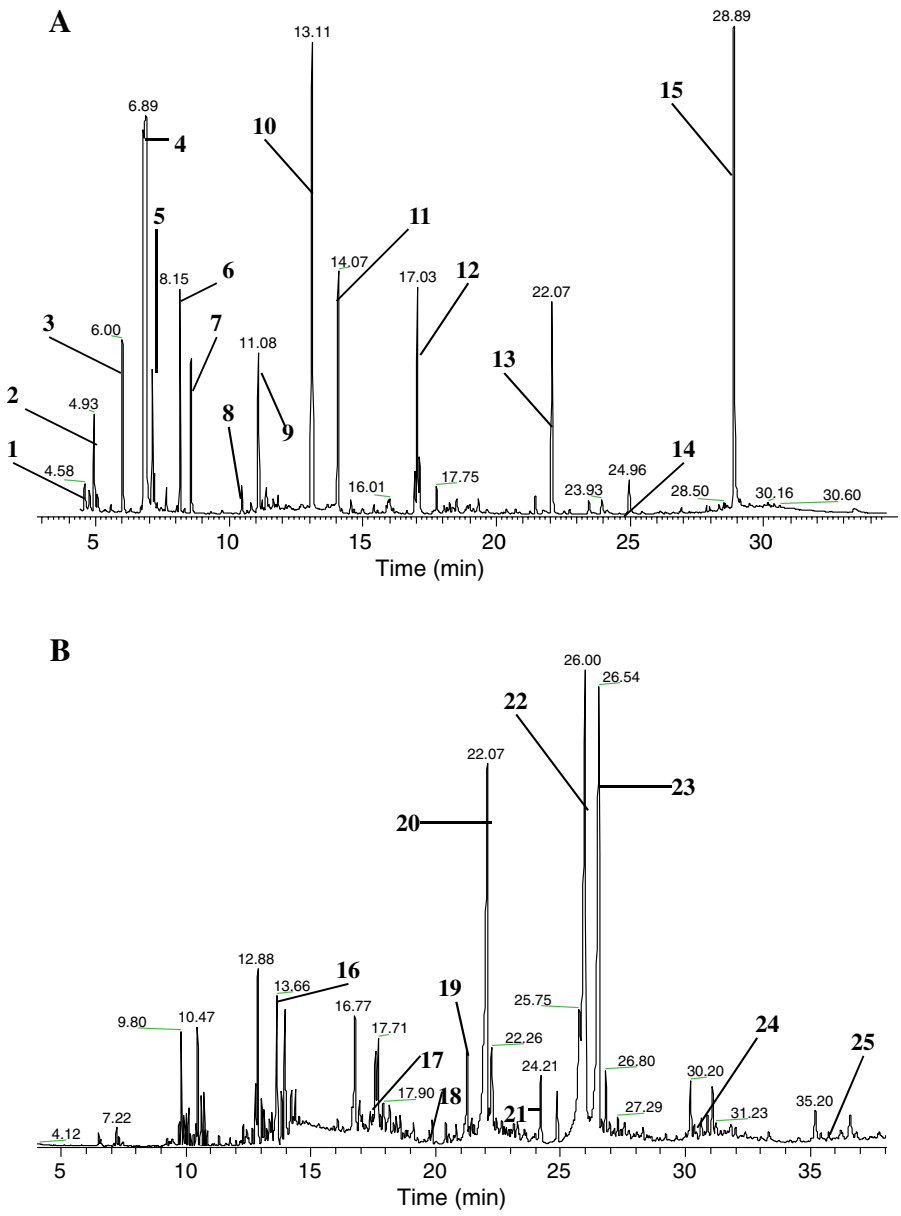

Fig. 3 GC-MS chromatograms of the aqueous (panel **a**) and organic phase (panel **b**) metabolites (X axis=time, Y axis= response). Due to space restrictions and differences in dynamic range only the major metabolites are labelled in this figure. For a full list of identified metabolites please see Table 1. Key: 1) Alanine, 2) Oxalic acid, 3) Valine 4) Phosphate (from NMR analysis), 5) Isoleucine, 6) Serine, 7) Threonine, 8) Malic acid, 9) Proline, 10) Glutamine, 11) Asparagine, 12) Isocitrate, 13) Inositol, 14) Glucose, 15) Unidentified sugar. Organic Phase—16) Lauric Acid, 17) Tetradecanoic acid, 18) Pentadecanoic acid, 19) 9-hexadecanoic acid, 20) Hexadecanoic acid, 21) Linolenic acid, 22) Heptadecanoic acid, 23) Steric acid, 24) Eicosadienoic acid, 25) Behenic acid

when the blast pathogen in the compatible interaction successfully penetrates the leaf. The fact that a metabolic change can be detected at this time indicates that this invasion invokes a rapid metabolic response from the plant.

Metabolic Dynamics—PCA analysis of the organic phase

Neither PCA nor the supervised extension PLS-DA could separate out groups using the organic phase NMR, GC-MS or LC-MS data with all models having low Q^2 values indicating poor predictive power. Hence it would appear that the organic phase metabolites did not differ significantly between the compatible/incompatible interaction groups and thus metabolites, such as lipids and fatty acids seem unlikely to be involved in the response of rice plants to pathogen infection. This may, at first, seem unexpected since organic phase metabolites such as phosphatidic and phosphatidyl glycerol phospholipids have previously been shown to be the major discriminatory non-polar metabolites in

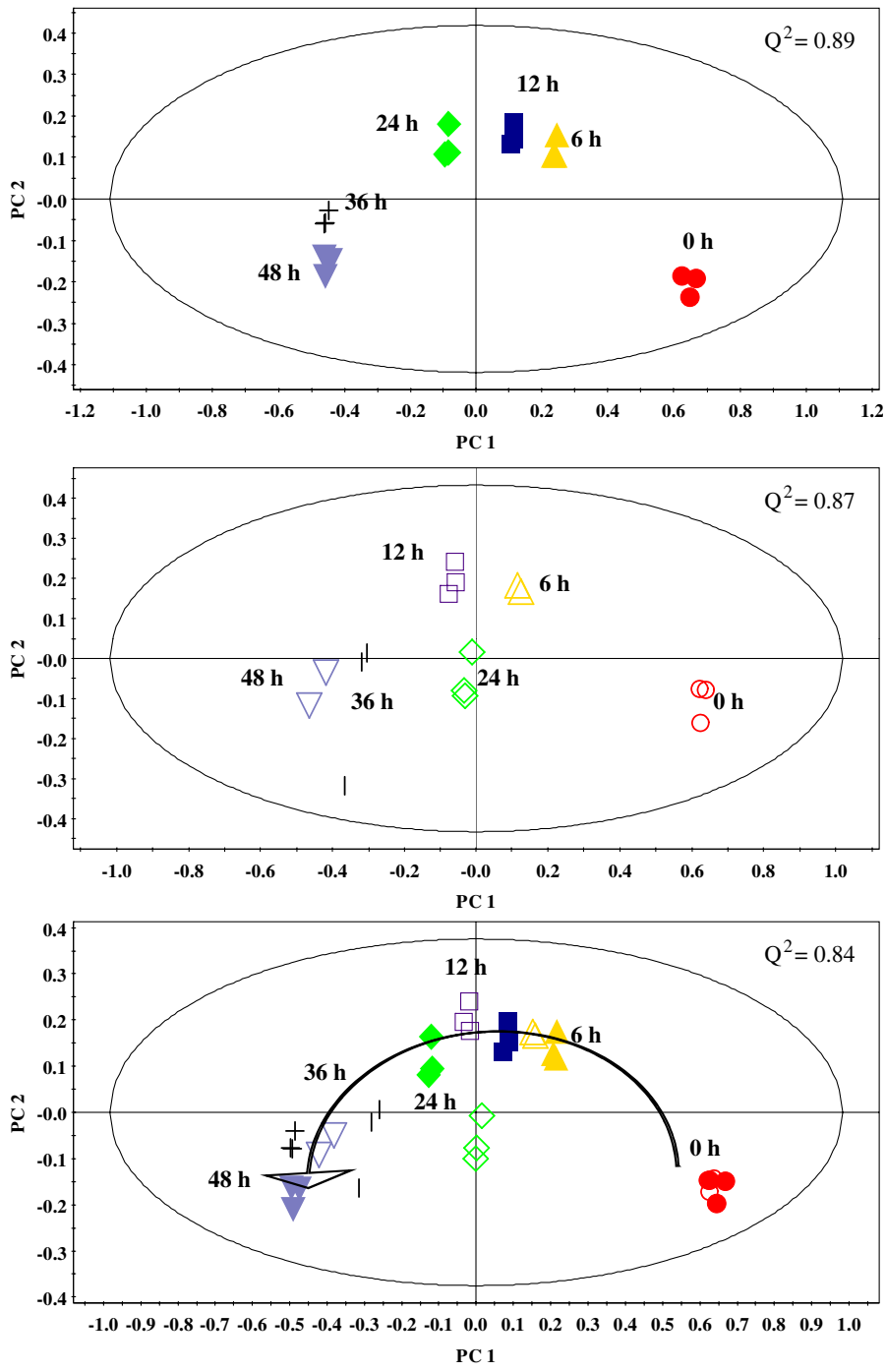

Fig. 4 PCA analysis of the aqueous metabolites for the incompatible (*filled symbols*-upper panel) and compatible interactions (*open symbols*—middle panel). NMR data is used in each case (although GC and LC-MS data show equivalent results in each phase). The bottom panel shows both groups plotted together. In the organic phase data the two groups do not separate

the response of the grass *Brachypodium distachyon* to challenge by *M. grisea* (Allwood et al. 2006) while jasmonic acid and its derivatives have also been demonstrated to be highly involved in plant defence systems (Tamogami et al. 2008). However, one must consider the huge variance in the biochemical defences of plants against pathogens or grazers. Indeed rice blast fungus is known to infect *Arabidopsis* via a mechanism which is totally distinct from that required for the infection of rice (Park et al. 2009) and thus infection response may also vary significantly between both plants and individual pathogen strains.

O-PLS STOCSY

Since a strong time trend was observed, an orthogonal partial least squares (O-PLS) model, regressing against time, was used to further analyse the data. This enabled any metabolic responses over time which differed between the compatible and incompatible reactions to be highlighted. An orthogonal partial least squares (O-PLS), Statistical Total Correlation Spectroscopy (STOCSY) based approach was used for this purpose. This was originally designed by Cloarec et al (2005) to be applied to NMR data. Here, we have also used it to analyse the GC-MS data, the first time (to the best of our knowledge) that this has been attempted. O-PLS STOCSY of the NMR data from both the compatible and incompatible reactions using the aqueous phase data are shown in Fig. 5 and the same data from the GC-MS are shown in Fig. 6. Peaks in each diagram which point up represent metabolites which increase with time while those that point down indicate those that decreased with time. Each peak is coloured (blue to red) according to its increasing correlation with time.

In the NMR data it can be seen that the compatible and incompatible interactions looks very similar. Although there are some differences between the groups in each case (highlighted with red circles), these are not highly correlated with time. The GC-MS data shown in Fig. 6 paints a more encouraging picture. Several metabolites increase with time and, crucially, these are different between each group. There were significant increases in malate, alanine, proline, glutamine, cinnamate and an unknown compound (most likely a sugar) in the compatible interaction group relative to then incompatible interaction group

Batch processing

To build on the STOCSY results and to further separate out differences between the compatible and incompatible interactions the technique of batch processing (BP) was used. The purpose of the method is to diagnose an evolving batch process as normal or abnormal and ascertain information on the variables or parameters responsible for any observed devia-

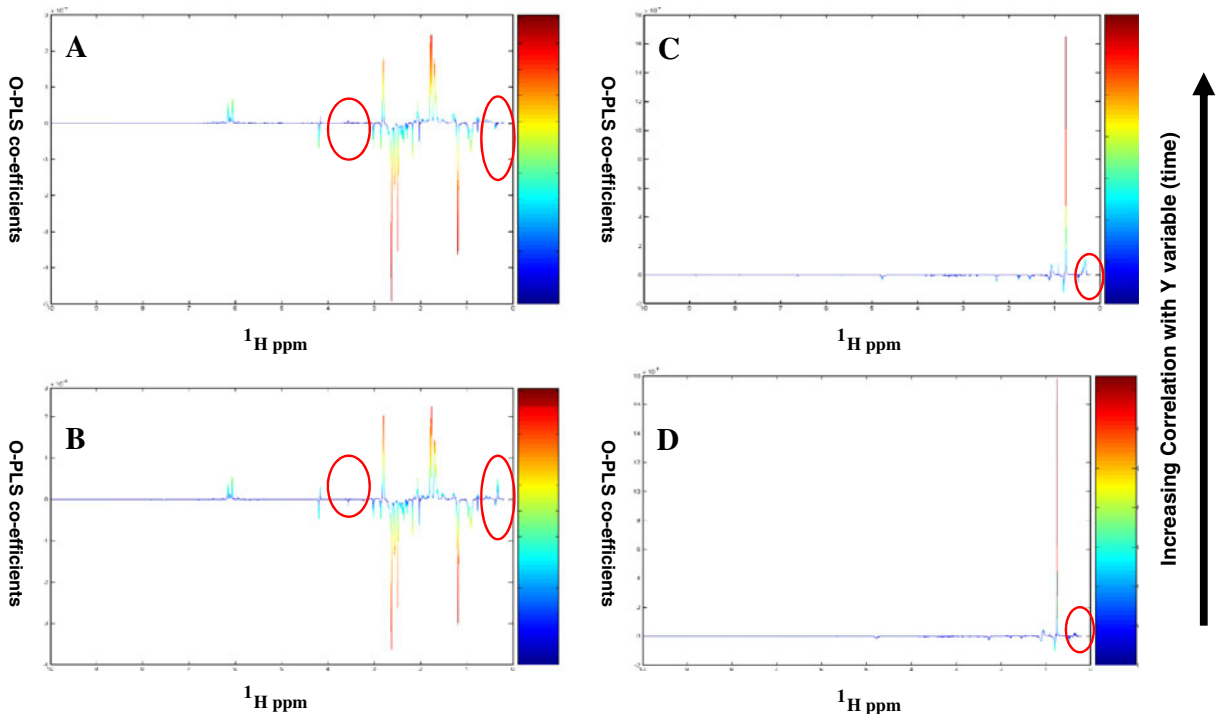

Fig. 5 O-PLS STOCSY of the NMR data from both the compatible and incompatible interactions. Panels **a** and **b** show the results of the incompatible and compatible blast strains (using the aqueous phase data) respectively, while panels **c** and **d** show the corresponding organic phase data. Small changes are highlighted within the *red circles*

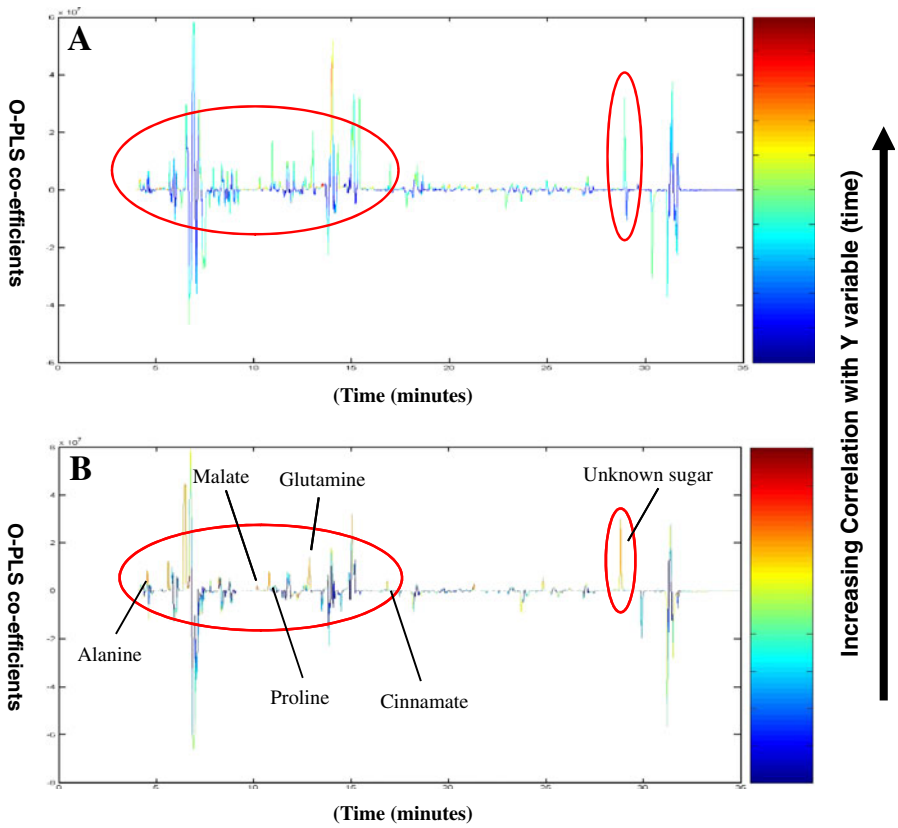

Fig. 6 O-PLS STOCSY of the GC-MS data from both the compatible and incompatible interactions. Panel **a** shows the incompatible interaction using the aqueous phase data while panel **b** shows the corresponding compatible interaction data

tions. Analysis of batch processes requires that all batches have equal duration and are synchronised according to sample collection time, thus the technique has primarily been used to monitor industrial applications evolving with time (such as brewing or baking). However, if one considers each animal or plant in a study as a batch then the demands for BP are fulfilled. Recent work has shown that using this approach can be useful in both toxicology (Antti et al. 2002) and plant growth studies (Wiklund et al. 2005).

BP was undertaken using the normal metabolic variation from control samples over time as the "good batch" as well as to map the dynamic response of the pathogen treated plants in relation to the control in both the compatible and incompatible interactions. In this way treatment related metabolic responses over time could be detected and classified as being disease related or not. The results are shown in Fig. 7.

Mapping of the compatible interaction onto the incompatible model again shows that the two models are virtually identical until the 24 h time-point. After this the compatible interaction samples are outside the limits of control model, as defined by±2 times the standard deviation (SD) boundary. This occurs in both the NMR and GC-MS data from the aqueous phase data from the 1st principal component and this divergence was again, primarily induced by changes in amino acid and antioxidant resonances. The 2nd and 3rd PLS components did not provide any extra information on the metabolic differences between the two groups and again, there were no changes in the organic phase in the NMR, GC-MS or LC-MS data.

The data from the 24, 36 and 48 h time points from each interaction were then combined into a new PCA model with nine members in each group. This showed good separation with a high Q^2 value in both data sets. These changes were increases in levels of the following metabolites in the compatible interaction group relative to then incompatible interaction group; malate, alanine, proline, glutamine, cinnamate and the same sugar peak as in the GC-MS STOCSY. The BP approach outlined here is thus demonstrated to be a useful framework for defining the evolution of metabolic fluctuations associated with plant disease/infection.

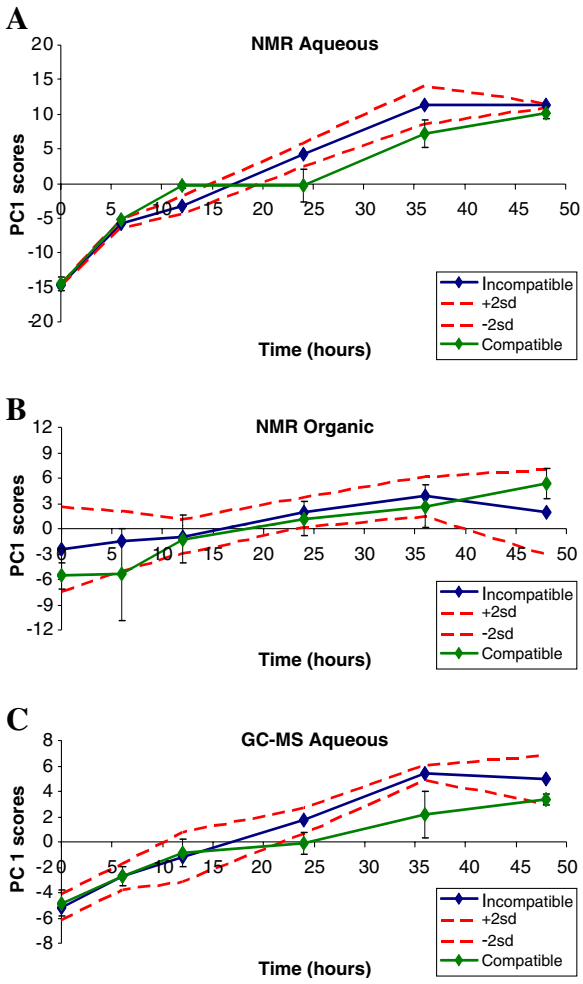

Fig. 7 Batch modelling of **a** NMR aqueous and **b** organic phase together with **c** the GC-MS aqueous phase data. All models were built using the incompatible interaction as a "good batch" onto which the compatible reaction was mapped . The red dashed lines show the limits of 2 times the standard deviation from the normal model in each case. The average trajectory deviates from these limits at the 24 h time point in the aqueous phase data, indicating that from this point on the samples are different. In the organic phase data the two groups do not separate further indicating that there is no significant change in the organic phase metabolites over the time course of the experiment

Compatible interactions indicate alanine as a possible marker of pathogen infection

The change in concentration was derived by looking at the difference in mean normalized intensity value for the metabolites between each time point in each data set for each compound. Of the 6 metabolites found to change in response to successful infection of the pathogen in the compatible interaction, alanine exhibited by far the largest relative change in concentration. At the 24 h time point it was found to be ~30% (± 9%) higher in the leaves from the compatible compared to the incompatible interaction. Levels of the other compounds which increased concentration all exhibited a less than 10% change with standard deviations of <5%. However, it should be noted that production of alanine and similar metabolites can also be induced via a range of stresses; for example anoxia and thus further work is needed before any firm conclusions can be drawn on the role of this compound in this case.

Discussion

Plants have evolved efficient mechanisms to combat pathogen attack. The critical differences between resistant and susceptible plant strains are the timely recognition of an invading pathogen in conjunction with the rapid and effective activation of host defences (Del Pozo et al. 2004; Pedley and Martin 2004). One of the earliest responses is often the generation of oxidative burst that can trigger hypersensitive cell death. This is called the hypersensitive response (HR) and is one of the earliest cellular responses following successful pathogen recognition in plants and a major element of plant disease resistance (Govrin and Levine 2000). However, some pathogens may induce production of reactive oxygen species (ROS) for their own advantage. For example, necrotrophs (such as *M. grisea*) may stimulate ROS production in infected tissue to induce cell death, facilitating subsequent infection and utilisation by the invader (Govrin and Levine 2000). Some studies also indicate ROS production being linked with increased levels of ROS detoxification enzymes, during compatible interactions involving viruses (Allan et al. 2001). This indicates that ROS is produced as part of a complex network of signals elicited as a response to pathogen attack with the potential to lead to opposite effects in different contexts or in response to different pathogens. Detection of induced intracellular ROS by the host may thus be the key difference in the successful defence against pathogens.

Interestingly, alanine has been shown to be involved with the activation of a programmed cell

death (PCD) response in suspension cultures of *Vitis labrusca*. This was accompanied by the coordinated activation of cinnamic acid 4-hydroxylase and several other related genes (Chen et al. 2006). It is possible that successful penetration of pathogen here triggers an increase in alanine levels to try and promote cell death of the infected tissue which *M. grisea* then exploits to facilitate invasion. The successful negation of this response in the case of the incompatible interaction therefore stops the infection early on and thus only small lesions are formed on the leaves around the site of infection. However, further studies would be needed before this hypothesis could be confirmed. The integration of multiple -omic techniques such as proteomics and transcriptomics would also help further elucidate the biochemical and common metabolic mechanisms used by plants to respond to pathogen infection.

An important feature of metabolomic data is that it simultaneously measures a large number of metabolites in a quantitatively reliable manner and therefore allows for discovery of correlations between pairs of metabolites, even when the biological connections between them are not obvious. In this study, which involves metabolites collected from whole leaves, few strong correlations were found. However, it was observed that several of the responsive metabolites had substantial positive or negative correlations with the infection process in PCA along PC1. Accordingly, the abundance of several of these metabolites also varied in a concerted fashion. For example, the metabolites that were decreased in susceptible compared to resistant plants following infection are involved in energy metabolism. These changes may be a result of the pathogen usurping some of the plant's resources for its own growth and reproductive needs.

The present study demonstrates the strength of metabolomics as a tool to assess the overall effects of pathogen infection on plants. It also emphasizes the utility of using multiple parallel metabolomic techniques for comprehensive identification of metabolites. Together with sample based multivariate analysis (pattern recognition), this allows scientists to visualize the overall metabolic effects of stress by comparison to the untreated control, and forms a quantitative and functional basis for a discussion of the overall progression of the effects of infection that would have been difficult to establish by other means. Although the experimental design provides information about changes in relative concentrations of metabolites rather than metabolic pathways, the results show that it can provide important information about the mechanisms of the infection response in plants as well as help to provide a more meaningful understanding of the plant/pathogen interaction of this important disease

While the metabolomics approach offers considerable potential for the rapid assessment of the metabolic status of virtually any organism or biological system it does have potential limitations of which it is important to be aware. These include its sensitivity to external influences such as temperature, light levels, nutritional status etc. However, providing studies are carefully constructed and monitored, and the results interpreted with respect to the experimental conditions, metabolomics holds great potential in plant science. If metabolomic data is integrated within a systems biology framework along with other -omics based technologies such as gene, transcript and protein expression profiling, it will become possible to obtain a more holistic understanding of the biochemistry underpinning plant-pathogen interactions in general.

Acknowledgements Oliver A.H. Jones was financially supported by the European Union (European Commission, FP6 Contract No. 003956. The study was in part funded by a grant from the Plant Signalling Network Research Centre from the Korean Science and Engineering Foundation. The authors also thank Dr. Young Hae Choi, Professor Robert Verpoorte and Professor George Ratcliffe for help with the NMR analysis and Drs. Aalim Welji and Denis Rubtsov for help with the Matlab code. The authors also thank the three anonymous reviewers whose thoughtful comments greatly improved the final manuscript.

References

Allan, A. C., Lapidot, M., Culver, J. N., & Fluhr, R. (2001). An early tobacco mosaic virus-induced oxidative burst in tobacco indicates extracellular perception of the virus coat protein. *Plant Physiology, 126*, 97–108.

Allwood, J. W., Ellis, D. I., & Goodacre, R. (2008). Metabolomic technologies and their application to the study of plants and plant-host interactions. *Physologia Plantarum, 132*, 117–135.

Allwood, J. W., Ellis, D. I., Heald, J. K., Goodacre, R., & Mur, L. A. J. (2006). Metabolomic approaches reveal that phosphatidic and phosphatidyl glycerol phospholipids are major discriminatory non-polar metabolites in responses by *Brachypodium distachyon* to challenge

by *Magnaporthe grisea*. *The Plant Journal, 46*, 351–368.

Antti, H., Bollard, M. E., Ebbels, T., Keun, H., Lindon, J. C., Nicholson, J. K., et al. (2002). Batch statistical processing of ^{1}H NMR-derived urinary spectral data. *Journal of Chemometrics, 16*, 461–468.

Bechinger, C., Giebel, K.-F., Schnell, M., Leiderer, P., Deising, H. B., & Bastmeyer, M. (1999). Optical measurements of invasive forces exerted by appressoria of a plant pathogenic fungus. *Science, 285*, 1896–1899.

Bölling, C., & Fiehn, O. (2005). Metabolite profiling of *Chlamydomonas reinhardtii* under nutrient deprivation. *Plant Physiology, 139*, 1995–2005.

Chen, J., Hall, D. E., Murata, J., & De Luca, V. (2006). l-Alanine induces programmed cell death in *V. labrusca* cell suspension cultures. *Plant Science, 171*, 734–744.

Cloarec, O., Dumas, M.-E., Craig, A., Barton, R. H., Trygg, J., Hudson, J., et al. (2005). Statistical total correlation spectroscopy: An exploratory approach for latent biomarker identification from metabolic ^{1}H NMR data sets. *Analytical Chemistry, 77*, 1282–1289.

Dean, R. A., Talbot, N. J., Ebbole, D. J., Farman, M. L., Mitchell, T. K., Orbach, M. J., et al. (2005). The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature, 434*, 980–986.

Del Pozo, O., Pedley, K. F., & Martin, G. B. (2004). MAPKKK alpha is a positive regulator of cell death associated with both plant immunity and disease. *EMBO Journal, 23*, 3072–3082.

Eriksson, L., Johansson, E., Kettaneh-Wold, N., & Wold, S. (2001). *Multi- and megavariate data analysis: principles and applications*. Umeå, Sweden: Umetrics.

Fiehn, O. (2002). Metabolomics – the link between genotypes and phenotypes. *Plant Molecular Biology, 48*(1), 155–171.

Goff, S. A., Ricke, D., Lan, T.-H., Presting, G., Wang, R., Dunn, M., et al. (2002). A draft sequence of the rice genome (*Oryza sativa L. ssp. japonica*). *Science, 92*, 100.

Govrin, E. M., & Levine, A. (2000). The hypersensitive response facilitates plant infection by the necrotrophic pathogen *Botrytis cinerea*. *Current Biology, 10*, 751–757.

Griffin, J. L., & Shore, R. F. (2007). Applications of metabonomics within environmental toxicology. In J. C. Lindon, J. K. Nicholson, & E. Holmes (Eds.), *The Handbook of Metabonomics and Metabolomics* (pp. 517–532). Kidlington: Elsevier Press.

Gullberg, J., Jonsson, P., Nordstrom, A., Sjostrom, M., & Moritz, T. (2004). Design of experiments: an efficient strategy to identify factors influencing extraction and derivatization of *Arabidopsis thaliana* samples in metabolomic studies with gas chromatography/mass spectrometry. *Analytical Biochemistry, 331*, 283–295.

Le Belle, J., Harris, N., Williams, S., & Bhakoo, K. (2002). A comparison of cell and tissue extraction techniques using high-resolution ^{1}H-NMR spectroscopy. *NMR in Biomedicine, 15*, 37–44.

Lin, C. U., Wu, H., Tjeerdema, R. S., & Viant, M. R. (2007). Evaluation of metabolite extraction strategies from tissue samples using NMR metabolomics. *Metabolomics, 3*, 55–67.

Morrison, W. R., & Smith, L. M. (1964). Preparation of fatty acid methyl esters and dimethylacetals from lipids with boron fluoride-methanol. *Journal of Lipid Research, 5*, 600–608.

Park, J.-Y., Jin, J., Lee, Y.-W., Kang, S., & Lee, Y.-H. (2009). Rice blast fungus (*Magnaporthe oryzae*) infects *Arabidopsis* via a mechanism distinct from that required for the infection of rice. *Plant Physiology, 149*, 474–486.

Pedley, K. F., & Martin, G. B. (2004). Identification of MAPKs and their possible MAPK kinase activatiors involved in the Pto-mediated defence response of tomato. *Journal of Biological Chemistry, 279*, 49229–49235.

Sato, S., Soga, T., Nishioka, T., & Tomita, M. (2004). Simultaneous determination of the main metabolites in rice leaves using capillary electrophoresis mass spectrometry and capillary electrophoresis diode array detection. *The Plant Journal, 40*, 151–163.

Schauer, N., & Fernie, A. R. (2006). Plant metabolomics: towards biological function and mechanism. *Trends in Plant Science, 11*, 508–516.

Sobolev, A. P., Brosio, E., Gianferri, R., & Segre, A. L. (2005). Metabolic profile of lettuce leaves by high-field NMR spectra. *Magnetic Resonance in Chemistry, 43*, 625–638.

Talbot, N. J. (1995). Having a blast: exploring the pathogenicity of *Magnaporthe grisea*. *Trends in Microbiology, 3*, 9–16.

Talbot, N. J. (2003). On the trail of a cereal killer: Exploring the biology of *Magnaporthe grisea*. *Annual Review of Microbiology, 57*, 177–202.

Tamogami, S., Rakwal, R., & Agrawal, G. K. (2008). Interplant communication: Airborne methyl jasmonate is essentially converted into JA and JA-Ile activating jasmonate signaling pathway and VOCs emission. *Biochemical and Biophysical Research Communications, 376*, 723–727.

Weljie, A. M., Dowlatabadi, R., Miller, B. J., Vogel, H. J., & Jirik, F. R. (2007). An inflammatory arthritis-associated metabolite biomarker pattern revealed by ^{1}H NMR spectroscopy. *Journal of Proteome Research, 6*, 3456–3464.

Wiklund, S., Karlsson, M., Antti, H., Johnels, D., Sjöström, M., Wingsle, G., et al. (2005). A new metabonomic strategy for analysing the growth process of the poplar tree. *Plant Biotechnology Journal, 3*, 353–362.

Yu, J., Hu, S., Wang, J., Wong, G. K.-S., Li, S., Liu, B., et al. (2002). A draft sequence of the rice genome (*Oryza sativa L. ssp. indica*). *Science, 296*, 79–92.