

OPEN

Unraveling Asian Soybean Rust metabolomics using mass spectrometry and Molecular Networking approach

Evandro Silva¹, José Perez da Graça⁴, Carla Porto^{1,2}, Rodolpho Martin do Prado^{1,3}, Clara Beatriz Hoffmann-Campo⁴, Mauricio Conrado Meyer⁴, Estela de Oliveira Nunes^{4,5} & Eduardo Jorge Pilau^{1*}

Asian Soybean Rust (ASR), caused by the biotrophic fungus *Phakopsora pachyrhizi*, is a devastating disease with an estimated crop yield loss of up to 90%. Yet, there is a lack of information on the metabolic response of soybean plants to the pathogen. Untargeted metabolomics and Global Natural Products Social Molecular Networking platform approach was used to explore soybean metabolome modulation to *P. pachyrhizi* infection. Soybean plants susceptible to ASR were inoculated with *P. pachyrhizi* spore suspension and non-inoculated plants were used as controls. Leaves from both groups were collected 14 days post-inoculation and extracted using different extractor solvent mixtures. The extracts were analyzed on an ultra-high performance liquid chromatography system coupled to high-definition electrospray ionization-mass spectrometry. There was a significant production of defense secondary metabolites (phenylpropanoids, terpenoids and flavonoids) when *P. pachyrhizi* infected soybean plants, such as putatively identified liquiritigenin, coumestrol, formononetin, pisatin, medicarpin, biochanin A, glycoellidin I, glycoellidin II, glycoellidin III, glycoellidin IV, glycoellidin V, glycoellidin VI. Primary metabolites (amino acids, peptides and lipids) also were putatively identified. This is the first report using untargeted metabolomics and GNPS-Molecular Networking approach to explore ASR in soybean plants. Our data provide insights into the potential role of some metabolites in the plant resistance to ASR, which could result in the development of resistant genotypes of soybean to *P. pachyrhizi*, and effective and specific products against the pathogen.

The soybean (*Glycine max* (L.) Merrill) is prominent among crops due to its agro-economic and nutritional value, used mainly as a source of proteins and oils for human consumption, in animal feeds and for biofuel production^{1,2}. However, estimates indicate a necessity to double global agricultural output by 2050 to feed a rising population, hence soybean productivity needs to be increased by 2.4% per annum³. Yet, soybean diseases were estimated to reduce crop yield by 11% in the United States⁴ and 50% or greater in the southeastern United States⁵.

The Asian Soybean Rust (ASR) was registered for the first time in Brazil during the 2001/2002 harvest⁶. Recently, Brazilian soybean farmers spent US\$ 2.16 billion with fungicides during the 2016/2017 harvest, and 96% of the sum was used for ASR control. Another US\$ 1.62 billion was invested on insecticides, totaling US\$ 3.78 billion. This amount represented 12.4% of the production costs for the harvest⁷. Despite increasing costs, disease management practices are needed as yield losses of up to 90% have already been reported when control measures were absent⁴.

¹Laboratory of Biomolecules and Mass Spectrometry, Department of Chemistry, State University of Maringá, 5790, Colombo Av, Maringá, PR, 87020-080, Brazil. ²Master in Science, Technology and Food Safety, Cesumar Institute of Science, Technology and Innovation – ICETI, University Center of Maringá - UNICESUMAR, 1610, Guedner Av, Maringá, PR, 87050-900, Brazil. ³Department of Animal Science, State University of Maringá, 5790, Colombo Av, Maringá, PR, 87020-080, Brazil. ⁴Brazilian Agricultural Research Corporation Soybean, Carlos João Strass Rd, Londrina, PR, 86001-970, Brazil. ⁵Brazilian Agricultural Research Corporation Swine and Poultry, BR-153, Km 110 Distrito de Tamanduá, SC, 89715-899, Brazil. *email: ejpilau@uem.br

The fungus *Phakopsora pachyrhizi* is the causal agent of the ASR. The phytopathogen *P. pachyrhizi* survives all-year-round whenever a soybean host plant is available and when conditions favor soybean development, thus supporting epidemic outbreaks⁸. Fungicides, such as demethylation and quinone outside inhibitors are largely used to prevent infection by *P. pachyrhizi*⁹. However, the Fungicide Resistance Action Committee (FRAC) recently reported regions in Brazil with intensive use of chemical fungicides where the efficiency of those products was reduced¹⁰, which can be caused by intensive use of a single chemical and excessive application. Management strategies to address ASR can range from planting early ripening varieties, field monitoring, eliminating secondary hosts, and even imposing soybean-free growth periods (60–90 days) in the threatened areas^{4,9}. Thus, no single solution is available to address *P. pachyrhizi* infection. Therefore, the use of several management strategies must be associated to reduce yield losses and to ensure the crop sustainability, which is time and resource consuming.

Perception of plants by invaders rapidly activates ion flow mechanisms; modulates the production of reactive oxygen species (ROS), primary metabolite modification, and biosynthesis of secondary metabolites; and favors expression of defense genes^{11,12}. The signaling cascade can be mediated by plant hormones such as salicylic and jasmonic acid, and ethylene. These hormones can then generate secondary metabolites with antimicrobial properties and modulate plant defenses against pathogens^{13–16}. During a pathogen attack, plants recognize pathogens or microbe-associated molecular patterns (PAMPs) and increase its immune defense systems¹⁷. Nonetheless, some pathogens can reduce plant immunity¹⁸, which triggers a second immune system known as effector-triggered (ETI). The ETI is activated by the recognition of microbial molecules from cell-surface-membrane proteins¹⁹. However, *P. pachyrhizi* can bypass both mechanisms and alter the host metabolism²⁰.

The vast majority of studies exploring plant metabolites have focused on the prospection of bioactive metabolites. However, with recent advances in analytical methods and annotation tools, metabolic modulation in plants as a result of abiotic and biotic stresses can now be systematically addressed²¹. Untargeted metabolomics has already been employed to improve our understanding of the metabolism of inter-kingdom interactions, including plant-pathogen interactions²². The analysis of the whole metabolic profile of a raw plant extract is challenging due to the diversity of chemical classes²³. Therefore, high efficiency liquid chromatography coupled to mass spectrometry can be used to explore complex matrix, allowing separation and detection of metabolites in a single chromatographic analysis²⁴. In addition, tandem mass spectrometry, such as the Q-TOF (Quadrupole - Time of Flight) geometry, enables sequential experiments (MS/MS) to provide more discriminating data for the metabolic identification process²⁵.

Metabolomics analysis of complex samples by liquid chromatography - mass spectrometry (LC-MS/MS) generate a large amount of data and the identification process is challenging^{26,27}. Therefore, the association of analytical techniques with molecular annotation tools is essential during the identification stage of the vast metabolic chemical diversity present in biological systems. Recently, Dorrestein and colleagues introduced a tandem mass spectrometry (MS/MS) data organizational approach, the Molecular Networking^{28,29}. Chemical entities can be arranged, visualized, and compared to databases available in the literature^{28–30}. Furthermore, fragmentation patterns similarities can be explored, thus enabling the discovery of novel compounds. Molecular Networking has led to the development of Global Natural Products Social Molecular Networking (GNPS) (<http://gnps.ucsd.edu>), a data-sharing platform that allows the research community to perform data-driven, crowd-sourced analysis and, it has been widely used for the annotation of MS-based chemical signatures³⁰.

In this context, the objective of this study was to use untargeted metabolomics approach to evaluate differences in the metabolic profile of soybean leaves when *P. pachyrhizi* were inoculated or not, and to identify metabolites actively produced by the soybean plant to cope with pathogen infection.

Materials and Methods

Plant materials and growth conditions. The genotype PI636463 (susceptible to ASR), provided by the Germplasm Active Bank of Embrapa Soybean was used. The get seeds (daughters seeds) having homogeneous physiological characteristics were obtained from a single seed (mother seed). The experiment using the get seeds was conducted in a greenhouse under the following growth conditions: temperature = 28 ± 2 °C; ultraviolet radiation = 70%, $\pm 10\%$; and 12 h/12 h light/dark photoperiod, up to vegetative stage V6-V7³¹. Two daughter seeds were sown per pot (5 L) containing substrate (soil, sand and compost; 3:2:2).

Inoculum preparation. The spores of *P. pachyrhizi* were supplied by the Embrapa Soybean Phytopathology Laboratory (Londrina, Parana, Brazil), obtained from cultivar BRS 284 (susceptible standard), with 92% of germination viability. A total of 500 mL of a spore suspension (3.6×10^4 U mL⁻¹) was prepared with sterile water as a vehicle to be used as an inoculant. 250 μ L of 70% ethanol solution was added into the suspension as the dispersing agent. For the control group (false-inoculated plants), inoculation with the same excipients was carried out, however, without the presence of spores.

Experimental design and metabolites extraction. The plants were divided into two groups: (I) control (non-inoculated) and (II) inoculated with *P. pachyrhizi*. The inoculum suspension containing *P. pachyrhizi* spores was distributed in each plant individually using a spray bottle for group II, and without spores for group I. Both groups were submitted to automatic nebulization (40 seconds) every two hours for a period of 14 hours to ensure increased levels of humidity. Increased humidity and absence of light simulates the dew or night breeze conditions of the field, which favors spore germination. After nebulization, control plants were transferred to another greenhouse kept under similar conditions of temperature, relative humidity and photoperiod to avoid contamination. After 14 days, it was possible to visually observe the sporulation of the pathogen and formation of lesions on soybean leaves (data not shown). Thus, leaves from groups I and II were collected, wrapped in aluminum foil,

immediately immersed in liquid nitrogen and transported to the laboratory. The samples were stored at -80°C until metabolites extraction.

Metabolite extraction procedures. The solvent mixtures ternary 1 (3:1:1; chloroform:methanol:water; v-v), ternary 2 (4:4:2; methanol:acetonitrile:water; v-v) and binary (8:2; methanol:water; v-v) were used to evaluate the fingerprint of soybean leaves from both groups³². Control and inoculated soybean leaves (5 whole leaves) were macerated separately under liquid nitrogen. Then, about 100 mg of plant material was transferred to 2 ml glass vials and 1.5 ml of the solvent mixture was added. The solution was shaken in vortex for 10 seconds (Vortex-Genie 2, Scientific Industries, New York, USA) and submitted to ultrasound (USC-1400, Unique, Brazil) for 25 minutes. The solution was centrifuged at 12,000 rpm for 10 minutes at 4°C . The supernatant was collected and filtered using a $0.22\ \mu\text{m}$ PVDF syringe filter (Millex-GV Durapore). The extraction procedures were performed in three replicates for each solvent system and for each group (I and II). The samples were concentrated in nitrogen flow and stored at -80°C until further analysis.

LC-MS/MS analysis. Extracts were resuspended in $600\ \mu\text{l}$ water:acetonitrile (1:1; v-v) and $3\ \mu\text{L}$ of each extract were injected and analyzed ultra-high performance liquid chromatograph (Shimadzu, Nexera X2, Japan) coupled to a hybrid quadrupole time-of-flight high resolution mass spectrometer (Impact II, Bruker Daltonics Corporation, Germany) equipped with an electrospray ionization source. Chromatographic separation was performed using an Acquity UPLC HSS T3 C18 column (Waters, USA, $1.7\ \mu\text{m}$, $2.1 \times 100\ \text{mm}$) at flow rate of $0.250\ \text{mL min}^{-1}$ ³³. The gradient mixture of solvents A (H_2O with 0.1% formic acid, v-v) and B (acetonitrile with 0.1% formic acid, v-v) was as follow: 5% B 0–1 min, 70% B 1–10 min, 98% B 12–20 min and maintained at 5% B 20–25 min at 40°C , the final five minutes being intended for reconstitution of the column for the next analysis. The instrument was calibrated using a solution of sodium formate ($10\ \text{mmol L}^{-1}$; isopropanol:water; 1:1; v-v) containing $50\ \mu\text{L}$ concentrated formic acid. The ionization source was operated in the positive ionization mode and adjusted to 4500 V, with a potential plate end of $-500\ \text{V}$. The dry gas parameters were set to $8\ \text{L min}^{-1}$ at 180°C with a nebulization gas pressure of 4 bar. The data were obtained in a range of m/z 50 to 1800 with an acquisition rate of 5 Hz. The 5 most intense ions were selected for automatic fragmentation (AutoMS/MS). The data were acquired by the Hystar Application software version 3.2 and Otof Control (Bruker Daltonics Corporation, Germany) and were converted to the mzXML file format³⁴.

Molecular networking. The MS/MS data was transferred to the GNPS Molecular Networking server to generate the chemical map (ID = e26e17c3ffca42dc9968f260c2b01cee), according to the GNPS documentation³⁰. Solvent and injection blanks were subtracted to generate the chemical map. The molecular network (MN) was generated so that the mass tolerance of precursor ions was 0.02 Da, as this value influences the clustering of almost identical fragmentation spectra (MS/MS). The mass variances of ion fragments for each group of acquired MS/MS spectra were stipulated as $\pm 0.02\ \text{Da}$ for clustering (consensus spectrum creation). The lines (connections between nodes) were formed only if the cosine score was above 0.7 and with a minimum correspondence of 4 peaks in the fragmentation spectrum. The MN spectra were then compared to the spectra of the GNPS spectral libraries, such as ReSpec, Massbank and HMDB^{35–37} where the same data parameters were applied to the sample spectra. MN data were visualized in the Cytoscape software³⁸. Ion fragmentation spectra with similarities to the mass spectral libraries had the fragmentation spectra manually verified and the mass error calculated, using a mass error tolerance of less than 10 ppm. Area-proportional Venn diagrams from exported molecular features data using GNPS were constructed using meta-chart (<https://www.meta-chart.com/>).

Results

Solvent effect on metabolites extraction. Extraction procedures varying in solvents were used to obtain abundant diversity of metabolites from both groups. Quantities of chemical entities extracted from control (group I) soybean leaves or when soybean leaves were inoculated with *P. pachyrhizi* (group II) using extraction solvent mixtures ternary 1, ternary 2, and binary were compiled in Venn diagrams (Figs. 1 and 2, respectively).

A total of 1101 chemical entities were observed in group I (Fig. 1). Of those, 360 chemical entities were extracted using the three extractor solvent mixtures, which represents 32.70% of the total (Fig. 1). The ternary 2 and binary solvent mixtures resulted in the highest quantities of detected chemical entities, 804 for each solvent mixture (73.02% of total). The ternary 1 solvent mixture resulted in the lowest quantity of detected chemical entities, 638 (57.95% of total). However, ternary 1 solvent mixture had the highest amount of unique chemical entities, 172 (15.62% of total; Fig. 1).

When soybean leaves were inoculated with *P. pachyrhizi* (group II), a total of 1282 chemical entities were observed (Fig. 2). This amount represented an increase of 14.12% when compared to group I. Of the total, 421 (32.84% of total) were extracted using the three extractor solvent mixtures (Fig. 2). The ternary 2 and binary solvent mixtures extracted the highest quantity of chemical entities, 968 (75.51% of total) and 992 (77.38% of total) respectively. The binary, ternary 2 and ternary 1 solvent mixture extracted 184 (14.35% of total), 139 (10.84% of total) and 81 (6.32% of total) unique chemical entities, respectively (Fig. 2).

Molecular networking and metabolites identification. The manual evaluation of each chemical entity detected by a mass spectrometer using an untargeted metabolomics approach is a challenging task. Thus, Molecular Networking tool was used to organize the MS/MS data (Fig. 3). Each node represents a chemical entity, which can be connected to other nodes according to chemical structure similarities. A node identified using the GNPS or open access libraries can then be used to identify another node in the same cluster by extrapolating gain or loss of a chemical group.

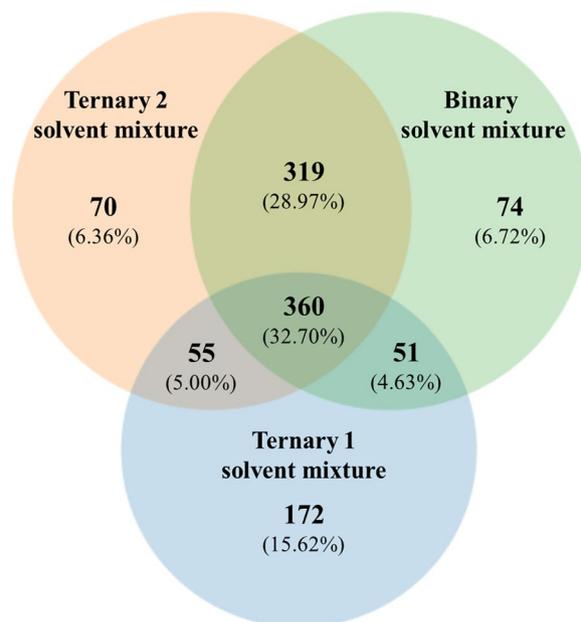


Figure 1. Venn diagram showing the distribution of chemical entities extracted from soybean control leaves using ternary solvent mixture 1 (3:1:1 chloroform: methanol: water; v-v), ternary solvent mixture 2 (4:4:2 methanol: acetonitrile: water; v-v) and binary solvent mixture (8:2 methanol: water; v-v). Chemical entities were detected using an UHPLC-ESI(+)-MS/MS.

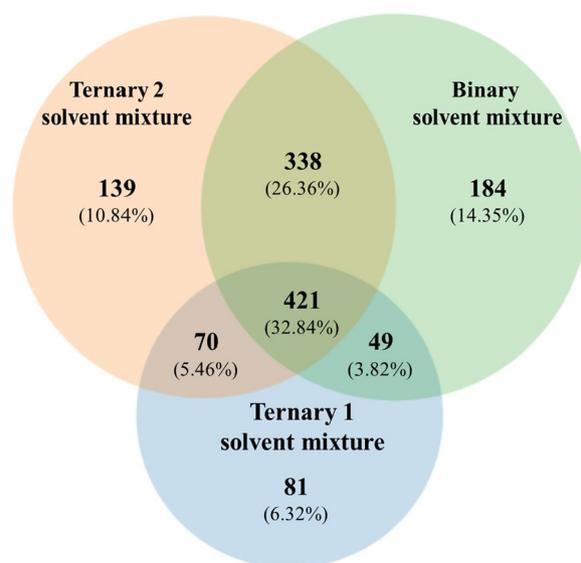


Figure 2. Venn diagram showing the distribution of chemical entities extracted from soybean leaves inoculated with *P. pachyrhizi*, using ternary solvent mixture 1 (3:1:1 chloroform: methanol: water; v-v), ternary solvent mixture 2 (4:4:2 methanol: acetonitrile: water; v-v) and binary solvent mixture (8:2 methanol: water; v-v). Chemical entities were detected using an UHPLC-ESI(+)-MS/MS.

When chemical entities of groups I and II were combined, a total of 1917 nodes (consensus spectra) were obtained in the MN (Fig. 3). Of these, 625 consensus spectra (red nodes) were obtained exclusively when *P. pachyrhizi* was inoculated on soybean leaves; 406 consensus spectra (green nodes) were obtained from the control group; and 886 consensus spectra (gray nodes) were obtained by samples of both groups. Fifty-two consensus spectra (black bold border nodes) had hits with the GNPS library (2.71% of the total). Furthermore, the MN had 97 clusters of nodes interconnected. Clusters containing at least one flavonoid/isoflavonoid, saponins, lipid, amino acid, chlorophyll, carboxylic acids or peptides were identified (Fig. 3A–J).

Metabolites putatively identification was confirmed manually by verification of fragmentation spectra (data not shown). Mass errors ranged from 0 to 3.9 ppm. Part of one cluster (Fig. 3A) containing the MS/MS spectra of $[M + H]^+ m/z$ 447.125, $[M + H]^+ m/z$ 463.121, $[M + H]^+ m/z$ 503.116 and $[M + H]^+ m/z$ 519.112, putatively

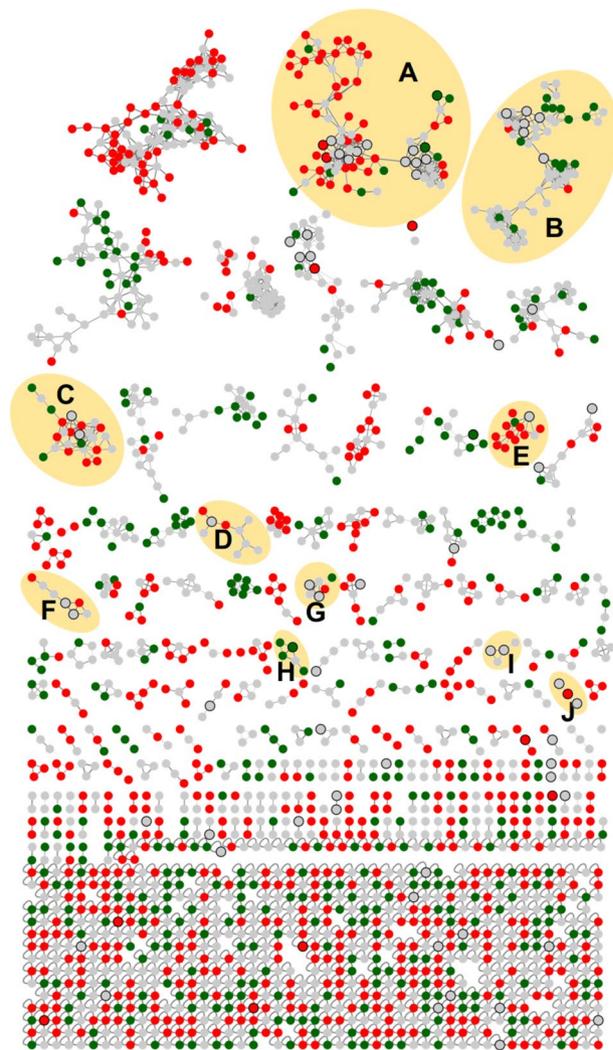


Figure 3. Molecular Network of the MS/MS spectra obtained by the analysis of the soybean control leaves, or inoculated with *P. pachyrhizi*, using solvent mixtures ternary 1 (3:1:1 chloroform: methanol: water; v-v), ternary 2 (4:4:2 methanol: acetonitrile: water; v-v) and binary (8:2 methanol: water; v-v). Green nodes correspond to the MS/MS spectra of soybean control leaves samples. Red nodes correspond to the MS/MS spectra of soybean leaves inoculated with *P. pachyrhizi*. Gray nodes correspond to the MS/MS spectra shared spectra in both samples. The edge width represents the cosine score (0.7 to 1.0). The black bold borders nodes represent the MS/MS spectra that had hits with the spectra of the GNPS libraries. (A) flavonoids/isoflavonoids, (B) saponins, (C) lipids, (D) amino acids, (E) flavonoids/isoflavonoids, (F) flavonoids (G) lipids, (H) chlorophyll, (I) carboxylic acids, (J) peptides.

identified as biochanin A 7-O-D-glucoside, 6-methoxyluteolin-7-rhamnoside, malonyldaidzin and malonylgenistin, respectively, was explored to highlight the potential of the GNPS approach (Fig. 4). These compounds are classified as isoflavone and flavones. Analogues of some of these compounds were separated by 14.016 Da, 15.999 Da and 162.052 Da, which are attributed to CH_2 , O and $\text{C}_6\text{H}_{10}\text{O}_5$ differences, respectively. All the nodes had higher spectral similarity since they were clustered with relatively high cosine scores (0.88–0.99). Using this approach, other ions belonging to the same cluster could be putatively identified. Furthermore, a pie chart layout³⁹ was generated using the peak ion area in each sample group (control and inoculated) for qualitative evaluation.

Using the GNPS approach and manual verification of fragmentation spectra, another 94 metabolites were putatively identified (Supplementary Table S1), which represented 4.90% of total. These metabolites belong to the classes of amino acids, nucleoside, vitamin, nucleotide, peptide, phenylpropanoids, flavonoids, isoflavonoids, organic acids, lipids, terpenoids and chlorophyll. From the total, 67 metabolites were identified in both groups (control and inoculated), 26 metabolites were putatively identified exclusively when soybean leaves were inoculated with *P. pachyrhizi*, and one metabolite was putatively identified in the group I. Of the ions exclusively putatively identified in group II, liquiritigenin, coumestrol, formononetin, pisatin, medicarpin, biochanin A (Fig. 5), glycoellidin I, glycoellidin II, glycoellidin I, glycoellidin II, glycoellidin III, glycoellidin IV, glycoellidin VI (Fig. 6) were identified.

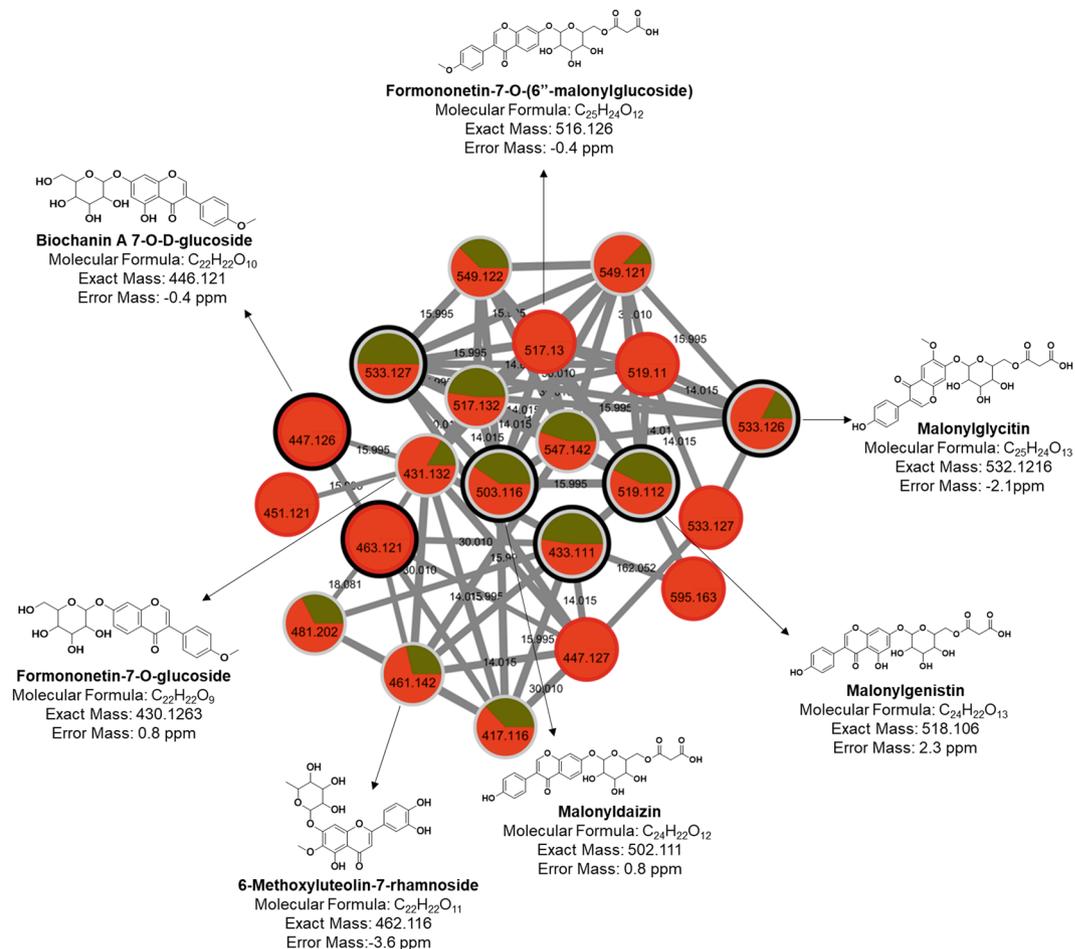


Figure 4. Cluster of isoflavonoids and flavonoids putatively characterized by molecular network obtained from MS/MS data from leaves of control and inoculated with *P. pachyrhizi*. The edge width represents the cosine score (0.88 to 0.99). The edge label represents the mass difference between nodes (14.016 Da, 15.999 Da and 162.052 Da). The black bold borders nodes represent the MS/MS that had hits with the spectra of the GNPS libraries. The pie chart within each node corresponds to the percentage relative of the metabolite in the sample, green indicates soybean control leaves, and red when soybean leaves were inoculated with *P. pachyrhizi*.

Discussion

As expected, each solvent mixture extracted diverse quantities of unique chemical entities (Figs. 1 and 2). Ternary 2 and binary solvent mixtures, which tends to favor the extraction of polar metabolites, extracted more of the exclusive chemical entities when *P. pachyrhizi* was inoculated (Fig. 2). On the other hand, the ternary 1 solvent mixture, which was used to favor non-polar metabolites, resulted on lowest amount of chemical entities (Fig. 2).

The whole metabolome of a given matrix can be complex, containing compounds with a large of range polarity⁴⁰. Solvent mixtures used as extractors are likely to impact the outcome of a study and should promote reproducible and comprehensive data, thus combinations of solvents were used in this study (ternary 1, ternary 2 and binary solvent mixtures). Using one solvent, or a solvent mixture can be a useful to explore a given sample when there is prior knowledge of the chemical class to be prospected, but with limitations. Conversely, the use of different solvent mixtures is needed to comprehensively explore the chemical ecology of a given matrix. Metabolites which were not known previously can be detected (Figs. 3 and 4), which expands the knowledge and provides novel insights about specific questions. For example, the soybean metabolome modulation was expected when the fungus *P. pachyrhizi*, which is known to cause decreased soybean crop yield⁷, was inoculated in the soybean plant. Indeed, phenylpropanoid, glyceollin and glyceollidin were observed only when *P. pachyrhizi* was inoculated (Figs. 5 and 6). These metabolites are involved in plant metabolism caused by biotic stress, where precursor molecules of biochemical pathways undergo chemical transformations by enzymes, leading to the production of secondary defense metabolites⁴¹. Hence, the combination of untargeted metabolomics and with the GNPS was used to prospect possible metabolites involved in the disease modulation.

Several of the identified metabolites are used by plants as defense mechanisms to counter microbial infections, such as terpenes, phenolics and N and S containing compounds⁴². Likewise, amino acids were observed (Supplementary Table S1), and these are important precursors of relevant secondary metabolites pathways that

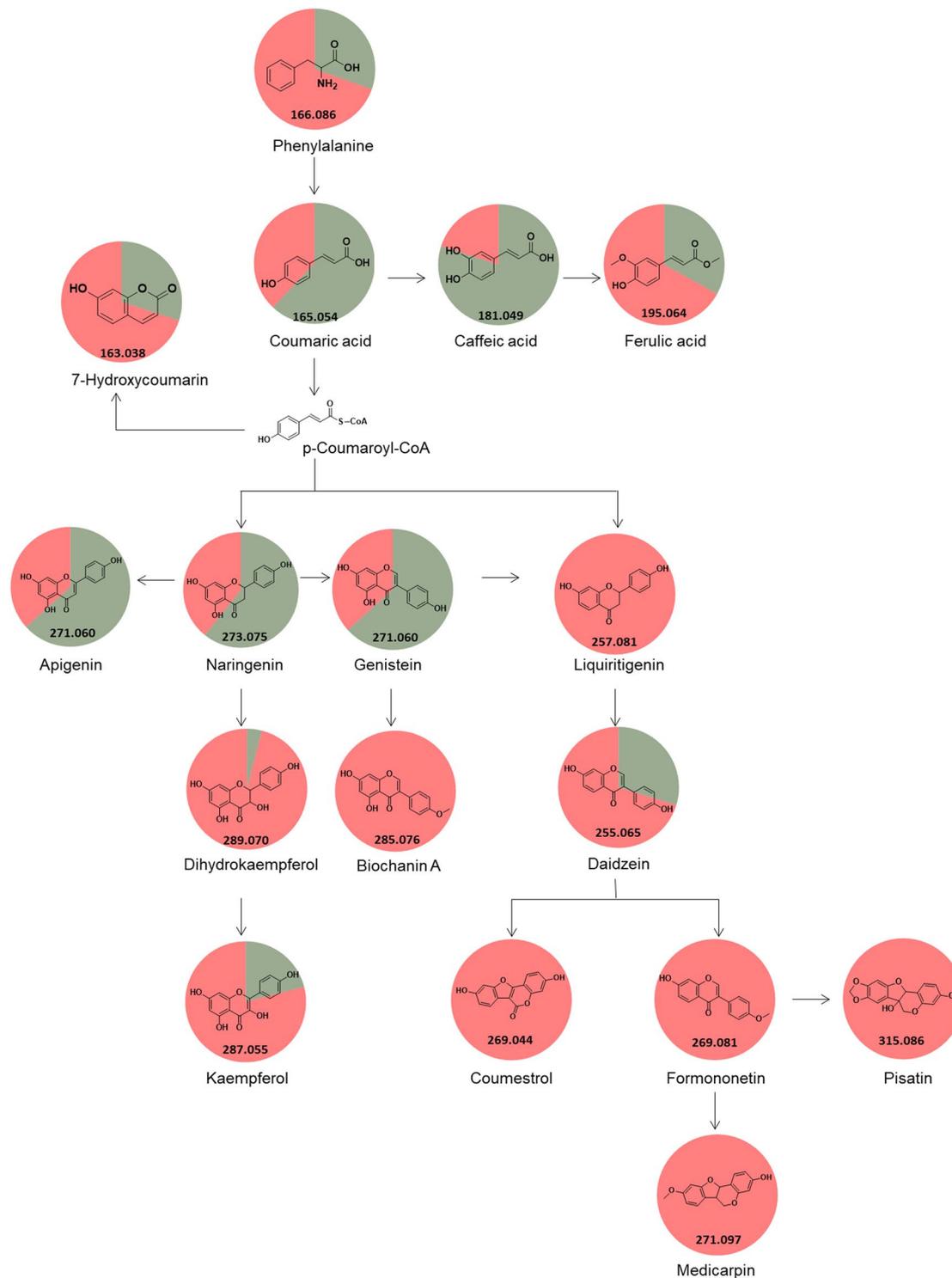


Figure 5. Summary pathway of phenylpropanoid biosynthesis. The pie chart corresponds to the percentage relative of the metabolite in the sample, green indicates soybean control leaves, and red when soybean leaves were inoculated with *P. pachyrhizi*. Adapted from Kanehisa⁵⁷.

participate in plant defense mechanisms such as salicylic acid, polyamines, alkaloids, phenolic compounds, phenylpropanoids⁴³.

In this study, we putatively identified four amino acids present in both groups: proline, phenylalanine, tyrosine and tryptophan. Phenylalanine is one of the main precursors of plant defense metabolites⁴⁴. Phenylalanine is converted to trans-cinnamic acid and ammonia by the enzyme phenylalanine ammonia-lyase (PAL). Trans-cinnamic acid can then be incorporated into several phenolic compounds (4-coumaric acid, caffeic acid, ferulic acid and sinapic acid), which are present in the formation of esters, coumarins, flavonoids and lignins⁴⁵.

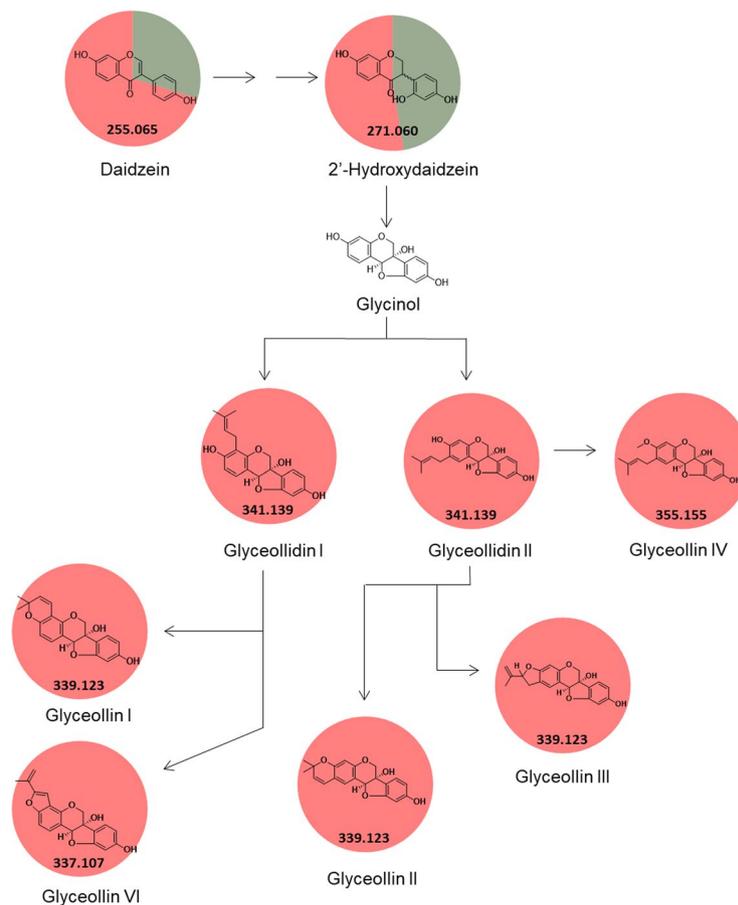


Figure 6. Glyceollin and glyceollidin biosynthesis pathway. The pie chart corresponds to the percentage relative of the metabolite in the sample, green indicates soybean control leaves, and red when soybean leaves were inoculated with *P. pachyrhizi*. Adapted from Simons⁵⁸.

The phenylpropanoids are plant secondary metabolites derived from aromatic amino acids. The induction of phenylpropanoid, flavonoid, and isoflavonoid metabolic pathway genes as a defense response to *P. pachyrhizi* was reported in transcriptome studies of susceptible and resistant soybean genotypes^{46–48}. We putatively identified 16 phenylpropanoids (Supplementary Table S1). Of those, liquiritigenin, coumestrol, formononetin, pisatin, medicarpin, biochanin A were presented, which are in the pathway of phenylpropanoid biosynthesis, were observed only when *P. pachyrhizi* was inoculated (Fig. 5).

Flavonoids and isoflavonoids are metabolites known as phytoalexins, antimicrobial substances that are synthesized by plants primarily after invasion or contact pathogens occur. Thus, the accumulation of these substances when *P. pachyrhizi* were inoculated soybean leaves may be indicate active plant defense against the pathogen⁴⁹.

Coumestrol, glyceollin and glyceollidin are isoflavonoids, belonging to subclasses of coumestans and pterocarpan, respectively. Both compounds are derived from daidzein (precursor), one of the major isoflavones belonging to soybean grains and leaves. We putatively identified glyceollidin I, glyceollidin II, glyceollin I, glyceollidin II, glyceollidin III, glyceollidin IV, glyceollidin VI when *P. pachyrhizi* were inoculated soybean leaves (Fig. 6). Simons and co-workers⁵⁰ reported the accumulation of coumestrol, glyceollin and glyceollidin in soybeans grown in the presence of the pathogen *Rhizopus spp.* Likewise, the phytoalexin glyceollin was also only detected when soybean leaves inoculated with *P. pachyrhizi*⁵¹.

Lipids and fatty acids are essential constituents of plant cells, providing structural integrity and energy for various metabolic processes. They can act as signal transduction mediators, playing both defense and signaling roles in response to biotic and abiotic stresses⁵². Some polyunsaturated fatty acids, such as linoleic and linolenic acids are precursors of oxylipins in plants⁵³. Oxylipins correspond to a wide range of products generated by self-oxidation or enzymatic oxidation of polyunsaturated fatty acids. Jasmonic acid was putatively identified for both groups in this study. Considered as a signaling substance, jasmonic acid is able to recognize plant infection by the pathogen and activate biosynthesis of defensive metabolites such as phytoalexins, terpenoids and flavonoids¹⁹.

Other putatively identified compounds were saponins, which are part of the terpenes class. Saponins can regulate plant growth and defense against insects and pathogens. They are present in high concentration in healthy plants, as they act as barriers to chemical reactions from fungi attack⁵⁴. Indeed, *Avena strigosa* saponin-deficient genotypes had their resistance compromised to a variety of fungal pathogens⁵⁵. The medicagenic acid saponins have been reported to have antifungal activity for *Sclerotium rolfsii*, *Rhizoctonia solani*, *Trichoderma viride*,

Aspergillus niger and *Fusarium oxysporum*⁵⁶. In this study, this class of compounds was putatively identified in control and inoculated leaves (Supplementary Table S1). However, the importance of saponins when soybean plants are infected by *P. pachyrhizi* is still unknown.

Conclusions

Asian Soybean Rust is an important disease and efficient pathogen control is challenging. Using untargeted metabolomics, we proposed a novel approach to evaluate the soybean metabolome modulation when affected by the Asian Soybean Rust disease. The combination of mass spectrometry and the Global Natural Products Social Molecular network was successful to identify exclusive metabolites present when soybean leaves were inoculated with *P. pachyrhizi*. Identification of metabolites associated with plant defenses is powerful information on the development of resistant genotype to resist *P. pachyrhizi*, as well as in the development of more effective and specific products against the disease.

Received: 24 September 2019; Accepted: 13 December 2019;

Published online: 10 January 2020

References

- Hill, J., Nelson, E., Tilman, D., Polasky, S. & Tiffany, D. Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proc. Natl. Acad. Sci.* **103**, 11206–11210, <https://doi.org/10.1073/pnas.0604600103> (2006).
- Fritz, H. *et al.* Soy, Red Clover, and Isoflavones and Breast Cancer: A Systematic Review. *PLoS ONE* **8**, e81968, <https://doi.org/10.1371/journal.pone.0081968> (2013).
- Ray, D. K., Mueller, N. D., West, P. C. & Foley, J. A. Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLoS ONE* **8**, e66428, <https://doi.org/10.1371/journal.pone.0066428> (2013).
- Hartman, G. L., *et al.* 2015. Compendium of Soybean Diseases and Pests, 5th Ed. (American Phytopathological Society Press, St. Paul, MN, 2016).
- Sikora, E. J. *et al.* A Coordinated Effort to Manage Soybean Rust in North America: A Success Story in Soybean Disease Monitoring. *Plant Dis.* **98**, 864–875, <https://doi.org/10.1094/PDIS-02-14-0121-FE> (2014).
- Yorinori, J. T., Paiva, W. M., Frederick, R. D. & Fernandez, P. F. T. Ferrugem da soja (*Phakopsora pachyrhizi*) no Brasil e no Paraguai, nas safras 2000/01 e 2001/02. In Congresso Brasileiro de Soja (p. 94). Londrina: Embrapa Soja (2002).
- CEPEA - Centro de Estudos Avançados em Economia Aplicada Mensuração econômica da incidência de pragas e doenças no Brasil: uma aplicação para as culturas de soja, milho e algodão - Parte 1, Maio de 2019, 12 p. (2019).
- Helfer, S. Rust fungi and global change. *New Phytol.* **201**, 770–780, <https://doi.org/10.1111/nph.12570> (2014).
- Godoy, C. V. *et al.* Asian soybean rust in Brazil: past, present, and future. *Pesqui. Agropecuária Bras.* **51**, 407–421, <https://doi.org/10.1590/S0100-204X2016000500002> (2016).
- FRAC. Fungicide Resistance Action Committee. Available at: www.frac.info/about-frac?v=46576. (Accessed: 22nd August 2019).
- Dat, J. *et al.* Dual action of the active oxygen species during plant stress responses. *Cell. Mol. Life Sci.* **57**, 779–795, <https://doi.org/10.1007/s000180050041> (2000).
- Yang, L. *et al.* Response of Plant Secondary Metabolites to Environmental Factors. *Molecules* **23**, 762, <https://doi.org/10.3390/molecules23040762> (2018).
- Hasegawa, M. *et al.* Phytoalexin Accumulation in the Interaction Between Rice and the Blast Fungus. *Mol. Plant-Microbe Interact.* **23**, 1000–1011, <https://doi.org/10.1094/MPMI-23-8-1000> (2010).
- Meng, X. & Zhang, S. MAPK Cascades in Plant Disease Resistance Signaling. *Annu. Rev. Phytopathol.* **51**, 245–266, <https://doi.org/10.1146/annurev-phyto-082712-102314> (2013).
- Couto, D. & Zipfel, C. Regulation of pattern recognition receptor signalling in plants. *Nat. Rev. Immunol.* **16**, 537–552, <https://doi.org/10.1038/nri.2016.77> (2016).
- Mhlongo, M. I., Piater, L. A., Madala, N. E., Labuschagne, N. & Dubery, I. A. The Chemistry of Plant–Microbe Interactions in the Rhizosphere and the Potential for Metabolomics to Reveal Signaling Related to Defense Priming and Induced Systemic Resistance. *Front. Plant Sci.* **9**, 112, <https://doi.org/10.3389/fpls.2018.00112> (2018).
- Nrnberger, T. & Kemmerling, B. Pathogen-Associated Molecular Patterns (PAMP) and PAMP-Triggered Immunity. In *Molecular Aspects of Plant Disease Resistance* 16–47, <https://doi.org/10.1002/9781444301441.ch2> (2009).
- De Coninck, B., Timmermans, P., Vos, C., Cammue, B. P. A. & Kazan, K. What lies beneath: belowground defense strategies in plants. *Trends Plant Sci.* **20**, 91–101, <https://doi.org/10.1016/j.tplants.2014.09.007> (2015).
- Chagas, F. O., Pessotti, R. D. C., Caraballo-Rodríguez, A. M. & Pupo, M. T. Chemical signaling involved in plant–microbe interactions. *Chem. Soc. Rev.* **47**, 1652–1704, <https://doi.org/10.1039/C7CS00343A> (2018).
- Ökmen, B. & Doehlemann, G. Inside plant: biotrophic strategies to modulate host immunity and metabolism. *Curr. Opin. Plant Biol.* **20**, 19–25, <https://doi.org/10.1016/j.pbi.2014.03.011> (2014).
- Tugizimana, F., Steenkamp, P. A., Piater, L. A. & Dubery, I. A. Multi-Platform Metabolomic Analyses of Ergosterol-Induced Dynamic Changes in *Nicotiana tabacum* Cells. *PLoS ONE* **9**, e87846, <https://doi.org/10.1371/journal.pone.0087846> (2014).
- Cox, D. G., Oh, J., Keasling, A., Colson, K. L. & Hamann, M. T. The utility of metabolomics in natural product and biomarker characterization. *Biochim. Biophys. Acta - Gen. Subj.* **1840**, 3460–3474, <https://doi.org/10.1016/j.bbagen.2014.08.007> (2014).
- Wolfender, J.-L., Marti, G., Thomas, A. & Bertrand, S. Current approaches and challenges for the metabolite profiling of complex natural extracts. *J. Chromatogr. A* **1382**, 136–164, <https://doi.org/10.1016/j.chroma.2014.10.091> (2015).
- Dias, D. A., Urban, S. & Roessner, U. A Historical Overview of Natural Products in Drug Discovery. *Metabolites* **2**, 303–336, <https://doi.org/10.3390/metabo2020303> (2012).
- Kind, T. & Fiehn, O. Advances in structure elucidation of small molecules using mass spectrometry. *Bioanal. Rev.* **2**, 23–60, <https://doi.org/10.1007/s12566-010-0015-9> (2010).
- Garg, N. *et al.* Mass spectral similarity for untargeted metabolomics data analysis of complex mixtures. *Int. J. Mass Spectrom.* **377**, 719–727, <https://doi.org/10.1016/j.ijms.2014.06.005> (2015).
- Scheubert, K. *et al.* Significance estimation for large scale metabolomics annotations by spectral matching. *Nat. Commun.* **8**, 1494, <https://doi.org/10.1038/s41467-017-01318-5> (2017).
- Watrous, J. *et al.* Mass spectral molecular networking of living microbial colonies. *Proc. Natl Acad. Sci.* **109**, E1743–E1752, <https://doi.org/10.1073/pnas.1203689109> (2012).
- Yang, J. Y. *et al.* Molecular Networking as a Dereplication Strategy. *J. Nat. Prod.* **76**, 1686–1699, <https://doi.org/10.1021/np400413s> (2013).
- Wang, M. *et al.* Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **34**, 828–837, <https://doi.org/10.1038/nbt.3597> (2016).
- Fehr, W. R. & Caviness, C. E. *Stages of Soybean Development. Special report* (1977).

32. Kim, H. K. & Verpoorte, R. Sample preparation for plant metabolomics. *Phytochem. Anal.* **21**, 4–13, <https://doi.org/10.1002/pca.1188> (2010).
33. t'Kindt, R., De Veylder, L., Storme, M., Deforce, D. & Van Bocxlaer, J. LC-MS metabolic profiling of Arabidopsis thaliana plant leaves and cell cultures: Optimization of pre-LC-MS procedure parameters. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.*, <https://doi.org/10.1016/j.jchromb.2008.06.039> (2008).
34. de Almeida, R. T. R. *et al.* Exploring the rumen fluid metabolome using liquid chromatography-high-resolution mass spectrometry and Molecular Networking. *Sci. Rep.* **8**, 17971, <https://doi.org/10.1038/s41598-018-36196-4> (2018).
35. Forsythe, I. J. & Wishart, D. S. Exploring Human Metabolites Using the Human Metabolome Database. In *Current Protocols in Bioinformatics* (John Wiley & Sons, Inc., 2009).
36. Horai, H. *et al.* MassBank: a public repository for sharing mass spectral data for life sciences. *J. Mass Spectrom.* **45**, 703–714, <https://doi.org/10.1002/jms.1777> (2010).
37. Sawada, Y. *et al.* RIKEN tandem mass spectral database (ReSpect) for phytochemicals: A plant-specific MS/MS-based data resource and database. *Phytochemistry* **82**, 38–45, <https://doi.org/10.1016/j.phytochem.2012.07.007> (2012).
38. Shannon, P. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **13**, 2498–2504, <https://doi.org/10.1101/gr.1239303> (2003).
39. Olivin, F., Roussi, F., Litaudon, M. & Touboul, D. Optimized experimental workflow for tandem mass spectrometry molecular networking in metabolomics. *Anal. Bioanal. Chem.* **24**, 5767–5778, <https://doi.org/10.1007/s00216-017-0523-3> (2017).
40. Mushtaq, M. Y., Choi, Y. H., Verpoorte, R. & Wilson, E. G. Extraction for Metabolomics: Access to The Metabolome. *Phytochem. Anal.* **25**, 291–306, <https://doi.org/10.1002/pca.2505> (2014).
41. Wang, S., Aseekh, S., Fernie, A. R. & Luo, J. The Structure and Function of Major Plant Metabolite Modifications. *Mol. Plant* **12**, 899–919, <https://doi.org/10.1016/j.molp.2019.06.001> (2019).
42. Mazid, M., Khan, T. A. & Mohammad, F. Role of secondary metabolites in defense mechanisms of plants. *Biology and Medicine* **3**, 232–249 (2011).
43. Gamir, J., Sánchez-Bel, P. & Flors, V. Molecular and physiological stages of priming: how plants prepare for environmental challenges. *Plant Cell Rep.* **33**, 1935–1949, <https://doi.org/10.1007/s00299-014-1665-9> (2014).
44. Kim, D. S. & Hwang, B. K. An important role of the pepper phenylalanine ammonia-lyase gene (PAL1) in salicylic acid-dependent signalling of the defence response to microbial pathogens. *J. Exp. Bot.* **65**, 2295–2306, <https://doi.org/10.1093/jxb/eru109> (2014).
45. Hyun, M. W., Yun, Y. H., Kim, J. Y. & Kim, S. H. Fungal and Plant Phenylalanine Ammonia-lyase. *Mycobiology* **39**, 257–265, <https://doi.org/10.5941/MYCO.2011.39.4.257> (2011).
46. Aoyagi, L. N. *et al.* Genomic and transcriptomic characterization of the transcription factor family R2R3-MYB in soybean and its involvement in the resistance responses to *Phakopsora pachyrhizi*. *Plant Sci.* **229**, 32–42, <https://doi.org/10.1016/j.plantsci.2014.08.005> (2014).
47. Ishiga, Y. *et al.* Transcriptomic and metabolomic analyses identify a role for chlorophyll catabolism and phytoalexin during *Medicago* nonhost resistance against Asian soybean rust. *Sci. Rep.* **5**, 13061 (2015).
48. Hossain, M. Z., Ishiga, Y., Yamanaka, N., Ogiso-Tanaka, E. & Yamaoka, Y. Soybean leaves transcriptomic data dissects the phenylpropanoid pathway genes as a defence response against *Phakopsora pachyrhizi*. *Plant Physiol. Biochem.* **132**, 424–433, <https://doi.org/10.1016/j.plaphy.2018.09.020> (2018).
49. Jeet, P. *et al.* Deciphering the role of phytoalexins in plant-microorganism interactions and human health. *Molecules*. **19**, 18033–18056, <https://doi.org/10.3390/molecules191118033> (2014).
50. Simons, R. *et al.* Identification of prenylated pterocarpan and other isoflavonoids in *Rhizopus spp.* elicited soya bean seedlings by electrospray ionisation mass spectrometry. *Rapid Commun. Mass Spectrom.* **25**, 55–65, <https://doi.org/10.1002/rcm.4826> (2011).
51. Lygin, A. V. *et al.* The Importance of Phenolic Metabolism to Limit the Growth of *Phakopsora pachyrhizi*. *Phytopathology* **99**, 1412–1420, <https://doi.org/10.1094/PHYTO-99-12-1412> (2009).
52. Wasternack, C. & Feussner, I. The Oxylipin Pathways: Biochemistry and Function. *Annu. Rev. Plant Biol.* **69**, 363–386, <https://doi.org/10.1146/annurev-arplant-042817-040440> (2018).
53. Nishiguchi, S. *et al.* Accumulation of 9- and 13-KODEs in response to jasmonic acid treatment and pathogenic infection in rice. *J. Pestic. Sci.* **43**, 191–197, <https://doi.org/10.1584/jpestics.D18-022> (2018).
54. Singh, B. & Kaur, A. Control of insect pests in crop plants and stored food grains using plant saponins: A review. *LWT* **87**, 93–101, <https://doi.org/10.1016/j.lwt.2017.08.077> (2018).
55. Papadopoulou, K., Melton, R. E., Leggett, M., Daniels, M. J. & Osbourn, A. E. Compromised disease resistance in saponin-deficient plants. *Proc. Natl. Acad. Sci.* **96**, 12923–12928, <https://doi.org/10.1073/pnas.96.22.12923> (1999).
56. Zehavi, U. *et al.* Synthesis and antifungal activity of medicagenic acid saponins on plant pathogens: modification of the saccharide moiety and the 23 α substitution. *Carbohydr. Res.* **244**, 161–169, [https://doi.org/10.1016/0008-6215\(93\)80012-4](https://doi.org/10.1016/0008-6215(93)80012-4) (1993).
57. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30, <https://doi.org/10.1093/nar/28.1.27> (2000).
58. Simons, R. *Prenylated isoflavonoids from soya and licorice Analysis, induction and in vitro estrogenicity.* (Wageningen University, 2011).

Acknowledgements

The authors would like to thank the Brazilian Agricultural Research Corporation (EMBRAPA), the Federal Foundation for the Brazilian Research and Development (FINEP), the Coordination for the Improvement of Higher Education Personnel (CAPES), the National Council for Scientific and Technological Development (CNPQ) and Complex of Research Support Centers of State University of Maringá (COMCAP-UEM). This project was funded by the grants 401590/2014-3-CNPq, 168384/2017-5-CNPq, 155675/2016-8 and 0673/13-CT-Infra-FINEP, 09/2016 – Programa de Pesquisa e Aplicada - Fundação Araucária de Apoio ao Desenvolvimento Científico e Tecnológico do Estado do Paraná This study was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil* (CAPES) – Finance Code 001.

Author contributions

E.S., C.C., E.N., M.M. and E.P. conceived the project. E.N. and J.G. performed the inoculation and greenhouse experiments. ES performed extractions and metabolomic analysis under the supervision of C.P., J.G., E.N. and E.P. E.S., R.P., C.P., E.N. and E.P. wrote the paper with input from all co-authors. All authors reviewed and approved the manuscript.

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-019-56782-4>.

Correspondence and requests for materials should be addressed to E.J.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020