

Mass Spectrometry-Based Metabolomics for the Development of Bioinputs: A Review on Plant Growth-Promoting Yeasts

Miriã A. Souza,^{1a,b} Tallyta S. Teixeira,^{1a} Jorge C. Rodrigues Neto,^{1a} Caio O. G. Silva,^{1b}
Catharine A. Bomfim,^{1a} Eder A. Barbosa,^{1a} Félix G. de Siqueira^{1a} and
Patrícia V. Abdelnur^{1b*,a,b}

^aEmbrapa Agroenergia, Empresa Brasileira de Pesquisa Agropecuária, W3 Norte, PqEB,
70770-901 Brasília-DF, Brazil

^bInstituto de Química, Universidade Federal de Goiás, Campus Samambaia,
74690-900 Goiânia-GO, Brazil

^cDepartment of Biotechnology and Biomedicine, Technical University of Denmark,
Søltofts Plads 221, 2800 Kongens Lyngby, Denmark

Technologies offering economic, technical, and environmental advantages over conventional chemical products hold promise for replacing chemical pesticides and fertilizers. Due to the positive interactions between plant growth-promoting yeasts (PGPY) and plants, there has been significant focus on developing bioinputs derived from these microorganisms and their metabolites. However, to advance the development of more specific and effective bioinputs, while also ensuring the health of both producers and consumers, it is crucial to deepen our understanding of microorganism-plant interactions, plant responses to microbial presence, and the metabolic changes in plants following bioinput application. Mass spectrometry-based metabolomics offers valuable insights into the metabolite interactions between microorganisms and plants, proving essential for a comprehensive understanding of these processes. This review presents the current state of mass spectrometry-based metabolomics, highlighting targeted and untargeted approaches, sample preparation, instrumentation, and data analysis methods, as well as their application in investigating the benefits of PGPY. The findings aim to provide insights that can aid in the development of commercial bioinputs. Despite substantial research published in scientific journals, only a limited number of yeast-based bioinputs have been commercialized in recent years, underscoring both the challenges and opportunities within this field.

Keywords: plant growth promoter, microorganisms, biotechnology, sustainable agriculture, metabolomics

1. Introduction

Due to the negative impacts that agrochemicals have on human health and the environment, reducing the use of agrochemicals (synthetic pesticides, fertilizers and herbicides) has become a global necessity. Bioinputs have emerged as one of the main alternatives to the use of agrochemicals, replacing their use in agriculture in whole or in part.¹ Bioinputs can be defined as the product, process or technology of plant, animal or microbial origin that positively interferes with the growth, development

and response mechanism of plants.² They can be produced from enzymes, plant or microbial extracts, microorganisms, invertebrates and secondary metabolites, aimed at biological control, plant nutrition and the relief of biotic and abiotic stresses.³

The first records of bioinputs were in 1950, with an initial study of nitrogen-fixing bacteria by researcher Johanna Döbereiner.⁴ As a result, biological fertilizers and fungicides for biological control, inoculants, biostimulants and solubilizers have emerged, contributing to the current growth dynamics of the sector.⁵ In Brazil, the use of bioinputs grew by 15% in the 2023/2024 harvest, compared to the previous harvest, reaching sales of R\$ 5 billion. In the global market, the estimate is that the use of bioinputs will

*e-mail: patricia.abdelnur@embrapa.br

Editor handled this article: Giovanni Wilson Amarante (Executive)



grow between 13 and 14% by 2032, which corresponds to US\$ 45 billion, a value three times higher than the current one.⁶

Currently, Brazil has 1456 bioinputs registered by the Ministry of Agriculture and Livestock (MAPA), of which 69.4% come from microorganisms.⁵ Several microorganisms can be used in the production of bioinputs.⁷⁻⁹ The use of yeasts as plant growth promoting yeasts (PGPY) has grown, with the main genera used being *Candida* spp. (15%), *Rhodotorula* spp. (14%), *Cryptococcus* spp. (13%) and *Saccharomyces* sp. (9%).¹⁰

Yeasts are eukaryotic and unicellular fungi, grouped in the phylum *Ascomycetes* or *Basidiomycetes*. They reproduce through budding or binary fission, and they do not form a fruiting body.¹¹ Yeasts were firstly discovered as a fermentative agent in wine and beer and today they are widely used for the production of beverages, foods, organic acids, enzymes, proteins, lipids, pigments and others. Yeasts may form symbiotic or parasitic association with plants and animals or they can act as saprophytic agents, playing crucial ecological roles, such as bioremediation, microhabitat colonization, biodegradation of hydrocarbons, detoxification, and the production of metabolism by-products.^{10,12,13}

In the field of agriculture, yeasts have gained importance as plant growth promoting agents (PGP), called PGPY. They can provide phytohormones, enzymes

and amino acids, promoting plant growth and biomass production, even under stress conditions.¹⁴ PGPY were studied mainly in samples of roots, shoots, fruits, and seeds of cultivars such as wheat, corn, rice, tomato, beans, alfalfa, grapes and pomegranate, beetroot.¹⁰ The main PGPY reported in the literature¹⁰ are *Candida* spp., *Rhodotorula* spp., *Cryptococcus* spp., *Saccharomyces* sp., *Kluyveromyces* sp., *Meyerozyma* sp., *Torulaspora* sp. and *Pichia* sp.

PGPY promotes plant growth through direct and indirect mechanisms. The direct mechanisms involve the production of phytohormones and stress controllers such as indole-3-acetic acid (IAA), auxins, cytokinins, gibberellins (GAs), abscisic acid (ABA), ethylene (ET), brassinosteroides (BRs), jasmonic acid (JA), salicylic acid (SA) and strigolactones (SLs).¹⁵⁻¹⁷ In addition, the direct mechanisms can induce iron sequestration through siderophores, solubilization of salts by chelation of silicon ions, biological nitrogen fixation, oxidation of sulfur to sulfates that are absorbed by plants and bioremediation.¹⁷⁻¹⁹ On the other hand, indirect growth occurs by biological control of plant pathogens and diseases at a mechanistic level. In this case, the PGPY inhibits the growth of plant pathogens by secretion of cell wall degrading enzymes, competition for nutrients such as iron, inducing systemic resistance and producing volatile organic compounds (Figure 1).^{17,20}

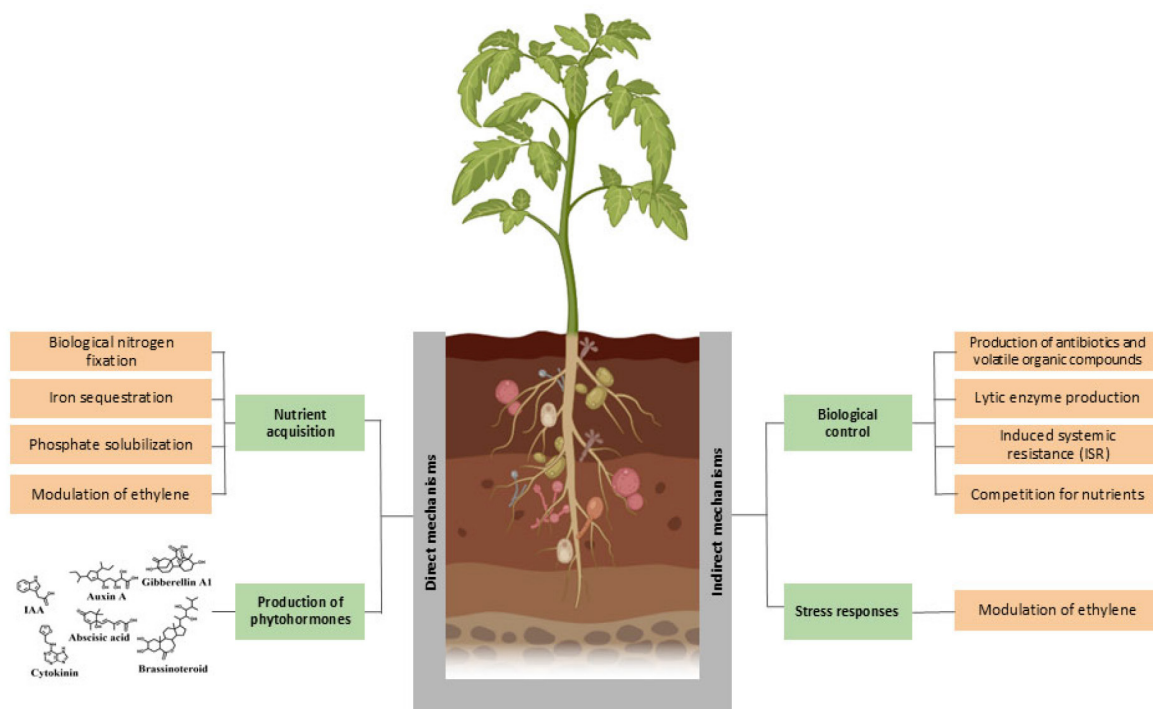


Figure 1. Growth mechanisms promoted by PGPY. Direct mechanisms involve the production of phytohormones and stress controllers, biological nitrogen fixation, phosphate solubilization, and iron sequestration via siderophores. Indirect growth mechanisms occur through biological control of pathogens through the production of cell wall degrading enzymes, and induction of systemic resistance.

Yeasts-based bioinputs are considered sustainable technologies, presenting economic, technical, and environmental advantages compared to conventional chemical products. They are good allies of crops, providing better growth, development, and response mechanisms in the metabolism of plants, and can replace, totally or partially, chemical pesticides and fertilizers. Yeast-based products can modify the plant metabolism increasing plant growth, the number of leaves, leaf area, chlorophyll concentrations and biomass production, not only under normal conditions, but also under conditions of stress and the deleterious effect of drought.^{14,18} In addition, yeast-based products can play an important role in soil health, fertility and nutrition, contributing to the sustainability of agricultural production and the implementation of a circular economy.

2. Beneficial Association of Yeasts and Plants

Yeasts are generally recognized as safe (GRAS) for use as plant growth promoting agents, and promote significant plant growth. One of their main characteristics is the (i) production of phytohormones, which is crucial in promoting plant growth and tolerance to biotic and abiotic stresses. Yeasts can facilitate the (ii) bioavailability of nutrients to plants by solubilizing phosphate necessary for (iii) plant growth and (iv) seed germination. Another important characteristic is the (v) biological control of pathogenic fungi, through competition with filamentous fungi, producing antifungal compounds or producing volatile organic compounds (VOCs).

2.1. Production of phytohormones

Yeasts can produce some phytohormones that promote plant growth and increase plant productivity. One group of growth-regulating phytohormones is the auxins, which have an indole ring.^{10,21,22} The presence of auxins increases the absorption of water and nutrients suitable for plants, improves root formation and plant development.²³⁻²⁵ In addition, it plays an important role in seed germination, cell differentiation, stress resistance and photosynthesis.²³ IAA is the main plant growth promoter in the auxin family, known to stimulate rapid and long-term responses in plants, regulating various physiological and developmental processes and metabolite biosynthesis.^{26,27}

Cytokinins are another group of phytohormones and promote plant growth through cell division and growth.²⁰ They also play an important role in seed germination, root elongation, nodule formation, flowering and vascular development.^{20,28,29} One of the most common cytokines is

zeatin, which can be synthesized by some yeasts such as *Sporobolomyces pararoseus*, *Metschnikowia pulcherrima* and *Aureobasidium pullulans*.²⁰

Other classes of phytohormones widely used are gibberellins, ethylene and abscisic acid.

Gibberellins help in general plant growth, acting in the promotion stages such as embryogenesis, seed germination, breaking dormancy, stem elongation, proper flowering, leaf/fruit senescence, leaf expansion, and fruit maturing.^{10,30,31} Recent studies³¹ demonstrate that yeasts such as *Yarrowia lipolytica* can synthesize gibberellic acids.

In addition to that, ET is essential for plant growth and development, fruit ripening and overcoming seed dormancy. However, when suffering any type of biotic and abiotic stress, ET is produced in high concentrations, suddenly inducing the death of the plant, it can inhibit plant growth and the development of roots after germination.^{30,32,33} Yeasts that have the enzyme 1-aminocyclopropane-1-carboxylic acid (ACCD) deaminase, degrade 1-aminocyclopropane-1-carboxylic acid (ACC), the immediate precursor of ET in higher plants.^{30,33} ACCD degrades ACC into α -ketobutyrate and ammonia (NH₃), inducing plant tolerance to stress and salt.³³

Finally, the ABA is involved in processes as seed dormancy, cell elongation, and flower induction.^{10,34} In addition, it can reduce the deleterious effects of water deficit and increase plant productivity. The yeast *Yarrowia lipolytica* synthesizes ABA, becoming a promoter host.³⁴

2.2. Bioavailability of nutrients

In addition to producing phytohormones, yeasts make nutrients more available. The production of NH₃ by yeast provides available nitrogen favoring plant growth.^{20,35} Some yeasts have the ability to solubilize and mineralize insoluble phosphate, making phosphorus available in soluble form for the plant. This solubilization is done through the mediation of organic acids, which chelate mineral ions or decrease the pH to bring P into solution.^{35,36} The acidification of microbial cells and their surrounding leads to the release of P-ions from the P-mineral by H⁺ substitution for calcium.³⁷ Strains of *Candida tropicalis* and *Lachancea thermotolerans* (Filippov) Kurtzman provided phosphorus through Ca₃(PO₄) solubilization.³⁸ *Rhodotorula* and *Yarrowia lipolytica* provide dissolved phosphorus through pH reduction and citric acid production, respectively.^{18,39} Similarly to phosphorus, yeasts convert the zinc that is available in soils in inorganic form to a soluble form available for assimilation by plants.¹⁰ In the case of iron, the siderophores produced by yeast help transport insoluble ferric iron from the environment

for uptake by plants, increase plant metal tolerance, and decrease metal availability and toxicity.³⁸

2.3. Plant and fruit growth

The application of yeasts in the rhizosphere or in the seeds is an efficient mechanism of microbial inoculation of the soil, being able to colonize the roots of the seedlings.³⁵ Nakayan *et al.*¹⁹ reported that *Meyerozyma guilliermondii* combined with a half dose of chemical fertilizer significantly improved the dry weight and nutrient uptake of maize and lettuce and the seed vigor index of maize and Chinese cabbage. Fernandez-San Millan *et al.*³⁵ highlighted 10 strains of yeast capable of increasing tobacco and lettuce growth, demonstrating that the effect of the yeast-plant interaction is not species-specific. In addition, the yeasts *Debaryomyces hansenii*, *Lachancea thermotolerans* and *Saccharomyces cerevisiae* promoted the development of maize seedlings, increasing the dry weight and chlorophyll content by 10%.³⁵

2.4. Seed germination

Yeasts can stimulate and accelerate seed germination because they can penetrate the seed endosperm (endophytic microflora) isolating amylolytic enzymes and physiologically active substances.⁴⁰ Furthermore, yeasts can quickly destroy the starch stock of seeds due to a release of extracellular amylase, which significantly increases the seed germination rate.⁴⁰ Fedotov *et al.*⁴⁰ reported that the use of yeasts *Rhodotorula*, *Cystofilobasidium*, *Sporobolomyces*, *Metschnikowia*, *Saccharomyces*, *Aureobasidium*, *Debaryomyces*, and *Cryptococcus* genera intensified the effect of seed stimulation and germination by more than 2-2.5 times, approximately from 11 to 31%. Yeast strains *Rhodotorula*, *Mrakia* and *Naganishia* increased seed germination percentage from day 3 to over 40% compared to 22% for uninoculated control seeds. In addition, these strains had an effect on seedling development, as the aerial part of the plant emerged earlier.⁴¹

2.5. Biological control

Plant diseases caused by pathogens are responsible for major crop losses worldwide, resulting in significant socioeconomic impact. Despite the reports, the effects of yeasts on crops and combating their pathogens are described in a more restricted way in relation to other microorganisms, such as bacteria and filamentous fungi. To apply yeasts as plant protection and biological control agents, it is necessary to understand the mechanisms

underlying the biocontrol activity of yeasts.

Yeasts use multiple mechanisms, which can occur simultaneously, increasing the antagonistic function, such as: (i) competition for nutrients and space, (ii) production of toxins, (iii) production of lytic enzymes, (iv) production of VOCs, (v) mycoparasitism and (vi) induction of resistance in plants (Table S1 in the Supplementary Information (SI) section).

Yeasts can act as biocontrol agents, biofungicide, bioinsecticide or bioherbicide and strengthen plant cell defense to prevent entry of pathogens. In addition, yeasts activate plant defense mechanisms, which produce compounds that reduce stress to environmental factors such as salinity, high temperature, drought, and metal toxicity.^{17,42} As yeasts grow faster than fungi, they reduce nutrient availability for pathogenic fungi and quickly colonize plant niches.⁴³ The production of VOCs such as alcohols, esters, aldehydes, ketones and lactones increases disease resistance and tolerance to abiotic stress. Toxins produced by yeast are capable of killing or inhibiting sensitive fungi, damaging the cell wall and cell membrane triggering apoptosis and inhibition of beta-glucan synthesis.¹⁰

Some examples are *Wickerhamomyces anomalus* yeasts, that significantly reduces the natural deterioration of the pear without causing any adverse change in the quality of the post-harvest pear fruit.⁴⁴ On the other hand, *Hanseniaspora uvarum* demonstrated efficiency in attracting and killing *Drosophila suzukii* present in grape plants.⁴⁵ The yeast *Torulaspora globosa* has antifungal action against the phytopathogenic fungus, *Colletotrichum sublineolum* and *Colletotrichum graminicola*, the causative agent of anthracnose in sorghum and maize, respectively.⁴⁶

There are opportunities to use bioinputs in the development for food security. However, there are just a few insecticides or herbicides yeast-based products available in the market. Information about plant/yeast interaction is still limited in terms of predicting and understanding the overall interaction relative to a sustainable environment. Yeast effects are influenced by species and genotype, inoculation methods, abiotic factors, soil type, fertilizer application, and the identification of VOCs, their functions and the roles played in the signaling system are crucial in the study of plant/yeast interactions.

A fact is the chemical mechanisms involved in PGPY/plant interactions, including those mediated by plant roots and the microbiome, have been little investigated. Understanding the characteristics of yeasts that promote plant growth and their interactions with plants is fundamental to the success of their application in the field. The use of mass spectrometry-based metabolomics has been expanding to overcome this gap.

3. Mass Spectrometry-Based Metabolomics Applied to PGPY

Metabolomics is an approach that studies the end products of cellular regulatory processes, commonly known as metabolites. Metabolites are small molecules (< 1500 Da) that provide a view of the state of a biological system at a specified moment, demonstrating how genotypic diversity affects phenotypic variation in plants and organisms.^{47,48} Metabolomics detect compounds in a sample and their concentrations, being a snapshot of the organism studied.⁴⁹ There are two strategies in metabolomics: (i) targeted and (ii) untargeted analysis. Targeted metabolomics is restricted to identifying and quantifying one or a group of pre-defined metabolites of a particular metabolic pathway.^{48,50} It is used to identify and quantify specific metabolites of interest.^{51,52} In addition, it is useful for the study of a genetic alteration, e.g., intracellular metabolites, metabolites secreted to the extracellular, nutrient usually absorbed by the cell and components not absorbed by the cell, but possibly transformed by extracellular enzymes.⁵³ Metabolic profiling constitutes an analytical approach employed for the identification and quantification of specific metabolites or metabolic classes, enabling the characterization of plant metabolic diversity and the discrimination of stress- or treatment-induced alterations.^{54,55} Untargeted metabolomics evaluates all measurable metabolites in a biological sample. It is an approach that focuses on annotating as many known and unknown compounds as possible in a sample.⁵⁶ The data are typically used for relative quantification between

sample groups, which can then be validated with targeted analyses.⁵⁶ The metabolic untargeted is subdivided into (i) fingerprint analysis (a large numbers of intracellular metabolites is observed); and (ii) footprinting analysis (known as metabolomic footprint, analyzes the profile of extracellular metabolites or the “exome”) (Figure 2).⁵³

Metabolic fingerprinting is an approach that provides scanning of a large number of intracellular metabolites, allowing comparison of the composition, interaction and changes of metabolite fingerprints.^{57,58} Metabolic footprinting provides information on the extracellular metabolites present in the sample.^{57,59} Biotransformation of substrate components is a part of the footprint.⁵³ In both cases, the extracellular or intracellular metabolite profile is constrained by experimental parameters, including extraction and chemical properties of the extraction solvent. In addition, the characterization of the metabolite profile can be influenced by the instrumentation used to identify and/or quantify the compounds of interest.

To elucidate bioactive metabolites, numerous factors must be considered for a reliable analysis. Factors such as biological and technical replicates must be taken into account in metabolomics analyses, because they can measure any source of biological fluctuation in data collected, limiting false acquisition disparities.⁶⁰ Biological replicates are measurements of biologically distinct samples, which represent random biological variation. In order to have the smallest variation of the biological system, a minimum of triplicate ($n = 3$) biological sampling is proposed with $n = 5$ preferred.^{61,62} Quality control (QC)

Metabolomics approaches

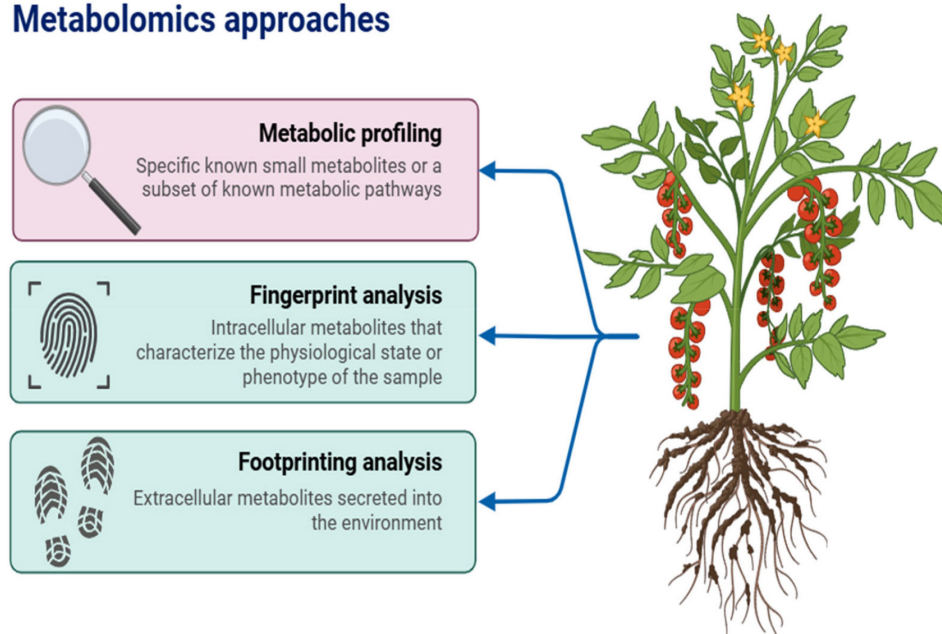


Figure 2. Metabolomics approaches: metabolic profiling, metabolic footprinting and metabolic fingerprinting.

samples are typically (i) pooled QC obtained by mixing small aliquots of each (or a few) biological samples to be studied, or (ii) commercially available QC samples with representative composition. With respect to the frequency of QC sample injections, 3-25 injections of samples between each QC injection were reported.⁶³ It is recommended that the reference sample is similar to the real samples of interest in complexity and composition.⁶⁴ Additionally, authentic commercially available standards can be added to the series of measurements to compare retention times for metabolite annotation.⁶⁴ A suitable experimental design for metabolomics investigations also includes analytical replicates, blanks, negative and positive controls to infer analytical and biological variations and assess data quality.⁶⁵

Technical replicates, on the other hand, are repeated measurements of samples processed using the same protocol to address equipment variations and human errors on the sample preparation.⁶⁰ In the laboratory, where variability is highly controlled, at least three replicates are required.

In metabolomics studies, a clear and robust experimental design is essential. Furthermore, at all stages of this study, it is crucial to follow predefined criteria throughout the experimental workflow, from material collection and storage, which precede sample preparation, to metabolite processing and annotation. The workflow will depend on the chosen metabolomics approach (targeted or untargeted).

Here, we will discuss the workflow for the untargeted metabolomics approach.

The main steps in mass spectrometry-based metabolomics untargeted approach include sample preparation, analytical techniques, data analysis and metabolite annotation. They are summarized in Figure 3, and they will be discussed below in detail.

3.1. Sample preparation

Sample collection and preparation is a fundamental step for the metabolomics analysis since the extract of the biological sample should reflect the original material.⁶⁷ For the sample preparation, some steps should be evaluated, such as the culture of the microorganism to obtain the metabolites, the sample collected, and the extraction of metabolites that will be analyzed.

Sample collection must ensure that the biological system sample is representative. When analyzing intracellular metabolism, the metabolism must be rapidly inhibited during sample collection to avoid metabolic turnover. This process is called quenching and is usually achieved by a rapid change in temperature.⁶⁸ For quenching, the most commonly used methods are cold methanol, liquid nitrogen, perchloric acid and acid/alkali.⁵³ Each one has advantages and disadvantages, as explained below.

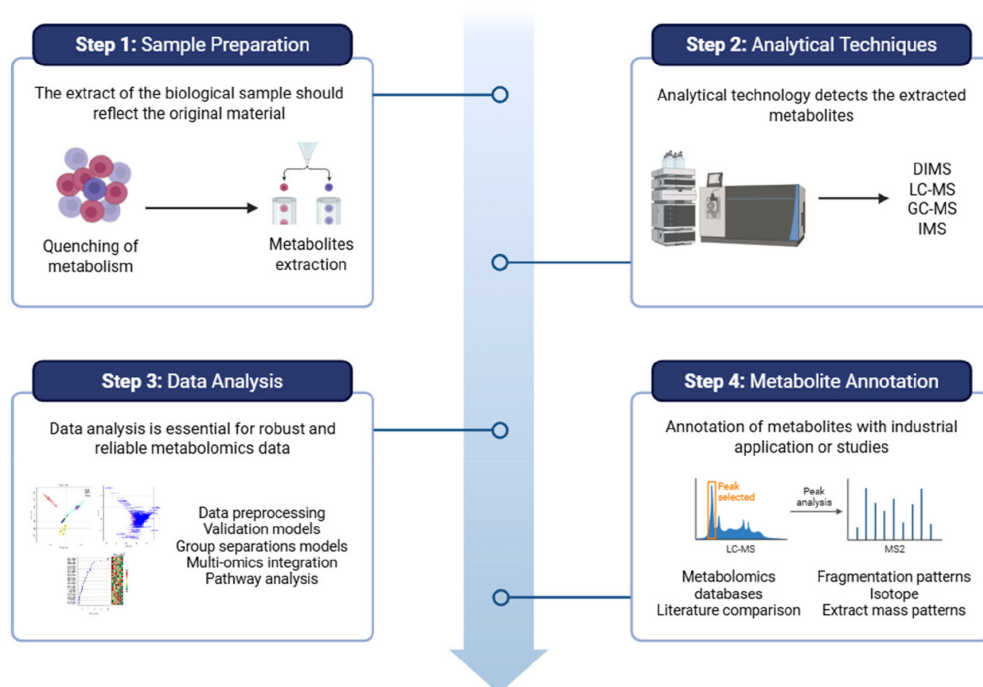


Figure 3. The workflow of metabolomics untargeted approach applied to bioinputs based on plant growth promoting yeasts. First, the sample preparation is required, and this step includes metabolites production and extraction with appropriate methods and solvents (step 1). Second, analytical techniques can be used based on research design and interests of it (step 2). Third, data are processed and analyzed to visualize the results (step 3). And finally, the metabolites are annotated and targeted new perspectives to industrial application or studies (step 4). For the targeted metabolomics approach, not covered in this workflow, other steps must be incorporated or waived. This figure was created with bioRender.⁶⁶

Quenching step using cold methanol allows the separation of intra and extracellular metabolites and is applicable to different microorganisms.⁵³ The disadvantages of this method are the difficulty in reproducing sample size, methanol water ration must be seriously controlled to avoid cell leakage and the need for technical adaptations for short time-scales.⁵³

Liquid nitrogen quenching reduces the sample temperature to -196°C and allows rapid, repeated sampling over short periods of time, however, it is not possible to separate intracellular and extracellular metabolites.⁵³ Inactivation of metabolism with perchloric acid is a procedure that allows inactivation of metabolism in fraction of seconds and analysis of nucleotides and amides, however, not all metabolites are stable at low pH and in the separation between extra- and intracellular compounds.⁵³ Also, the acid/alkali method is efficient for extraction of amines, amides present in the sample, but several metabolites are not stable in extreme pH and oxidative/reductive media.⁵³ Sample collection and tempering are sample pre-processing steps.

Most metabolites require an extraction step for further analysis. Therefore, an ideal extraction protocol must extract the largest possible number of metabolites and not be destructive or modify the molecules by chemical or physical means in simple or complex samples.⁶⁹ Due to the high diversity of physical and chemical characteristics (e.g., polarity, charged state, and acidity or basicity) of metabolites in biological samples, the solute-solvent interactions, dissociation of the solute, the matrix effect, molecular weight, and solvent solubility must all be considered when choosing the extraction solvent.^{53,70,71} In addition, the composition and polarity of the solvent, as well as the temperature and the period of extraction should be optimized because of the differences in chemical behavior, structure, and concentration of metabolites.⁶⁷ Liquid-liquid extraction is the most commonly used method.⁷² Metabolites are separated based on their differential solubilities in immiscible solvents, thus leading to partitioning. The choice of solvent depends on the polarity of the metabolites. Polar metabolites, such as amino acids, sugars, sugar phosphates, nucleotides, and polyamines, can be extracted with polar solvents.⁷² The most commonly used polar solvents are water (H_2O), methanol (CH_4O), ethanol ($\text{C}_2\text{H}_6\text{O}$), acetonitrile ($\text{C}_2\text{H}_3\text{N}$), isopropanol ($\text{C}_3\text{H}_8\text{O}$), and acetone ($\text{C}_3\text{H}_6\text{O}$).⁷² Nonpolar metabolites, such as lipids, fatty acids, ceramides, hormones, and cholesterol, can be extracted in chloroform (CHCl_3), methyl *tert*-butyl ether (MTBE, $\text{C}_5\text{H}_{12}\text{O}$), hexane (C_6H_{14}), and dichloromethane (CH_2Cl_2).⁷²

Biphasic liquid-liquid extraction is the most widely used

approach for partitioning polar and nonpolar metabolites in metabolomics workflows. It is obtained by combining a polar aqueous phase with a nonpolar organic solvent, such as methanol and chloroform ($\text{MeOH}/\text{CHCl}_3$), methanol, chloroform, and water ($\text{MeOH}/\text{CHCl}_3/\text{H}_2\text{O}$), acetone and water ($\text{CH}_3\text{COCH}_2/\text{H}_2\text{O}$), methanol, isopropanol, and water ($\text{MeOH}/\text{IPA}/\text{H}_2\text{O}$).⁷² Adjustments to the $\text{MeOH}:\text{CHCl}_3$ ratio (e.g., 1:1, 2:1, or 3:1) allow modulating the partitioning between lipids and polar metabolites; for highly polar compositions, pure methanol or strongly methanolic ratios (e.g., 9:1 $\text{MeOH}:\text{CHCl}_3$) are used.^{72,73}

Manipulating the pH of the aqueous phase represents a powerful strategy for enriching or displacing specific classes of metabolites through acid-base mismatch, although extraction is typically performed at neutral pH.⁷²

The choice of the best extraction method should be guided by the physicochemical diversity of the analytes and the metabolomic strategy (targeted or untargeted), as well as empirical validation of recovery and reproducibility.

3.2. Analytical techniques

Sophisticated analytical resources are required to detect all the extracted metabolites, which vary considerably in concentration as well as in their physical and chemical properties.⁵⁰ The choice of the analytical technology depends on the objectives of the study and is a compromise between the sensitivity, selectivity, and speed of the analysis.⁷⁰

Mass spectrometry (MS) is one of the fundamental techniques for analyzing a wide range of compounds, from small to high molecules, such as proteins.⁷⁴ MS offers high sensitivity, resolution, and robustness. It provides chemical information related to the structure of the chemical compound, such as mass, isotopic distribution pattern, and ionic fragmentation for structural elucidation based on the molecular mass/charge ratio of the ion (m/z).⁷⁰

MS methods are based on analysis of the biological samples by direct injection/infusion or after a chromatographic separation, such as GC-MS (gas chromatography-mass spectrometry), and LC-MS (liquid chromatography-mass spectrometry). In direct infusion mass spectrometry (DIMS), the sample moves directly into the source of the mass spectrometer without previous separation,⁷⁵ and samples are injected directly into the syringe pump MS.⁷⁶ DIMS allows acquisition of the MS in a few minutes, demonstrating analytical repeatability, as well as comparable classification and prediction capabilities.^{76,77} In addition, in the DIMS approach there is a greater possibility of ion suppression, differentiation of isobars and mass spectra complexity.

The combination of chromatography with MS is used for the precise quantification of chromatographic peaks, the separation of isobaric and/or isomeric compounds and the remarkable reduction of so-called matrix effects.⁷⁷ However, the analysis time increases significantly, when compared to analyses performed in DIMS. Biological samples are complex; therefore, the coupling of separation techniques is needed to reduce analytical interferences, such as ion suppression, overlapping peaks and matrix effects.⁷⁸ Ion suppression occurs when more easily ionizable species mask the presence of less ionizable ones.⁷⁸ Ion suppression has deleterious effects on both electrospray ionization (ESI) and atmospheric pressure chemical ionization (APCI), further impacting ESI.⁷⁹ The matrix effects, on the other hand, are caused by co-elution of matrix components on the ionization of the target analyte.⁸⁰ This effect changes the ionization of target analytes and the chromatographic response, leading to a reduction or increase in the sensitivity of the analysis.⁸⁰ The high concentration of interfering compounds reduces the ability of the analyte to reach the gas phase due to the high viscosity and surface tension of the droplets produced at the ESI interface.⁸⁰ In addition, interfering substances can neutralize analyte ions or affect the stability of the produced ions.⁸¹

The number of metabolites that can be annotated and the group of metabolites (polar, lipids, organic acid, etc.) vary according to the chromatographic technique used.⁶² The most commonly used experimental methods are GC-MS and LC-MS.^{78,82} GC-MS-based metabolomics is ideal for identifying and quantitating small-molecule metabolites (< 650 Da), including volatile organic compounds, lipids, derived primary metabolites and molecules small acids, alcohols, hydroxyl acids, amino acids, sugars, fatty acids, sterols, catecholamines, drugs, and toxins.^{83,84}

As most primary metabolites have higher boiling points (e.g., lactate, pyruvate, malic acid, glucose, palmitate, or similar metabolites), a GC-MS analysis requires a sample derivatization step to create volatile compounds. Derivatization improves volatility, thermal stability, detector sensitivity and response, and degradation of molecules at high temperatures.⁶² GC-MS can separate and quantify volatile metabolites, including organic residues, most amino acids, sugars, sugar alcohols, aromatic amines and fatty acids.⁸³ On the other hand, polar and non-polar compounds, non-volatile metabolites and large or thermolabile compounds are usually analyzed using LC-MS.^{62,83}

LC-MS methods have high resolving power, sensitivity, and specificity and provide a reproducible analysis of a wide range of metabolite classes.⁸⁵ The sample preparation is

simple and well established by the literature,⁸⁶ and it applies to complex mixtures, polar and non-polar compounds. No derivatization of the sample is required. On the other hand, the matrix effects and formation of multiple adducts can increase the difficulty of the interpretation. Moreover, information on fragmentation is required for a more comprehensive understanding of the investigated structures of the compound.⁸²

The development of ultra-performance liquid chromatography coupled to mass spectrometry (UPLC-MS) provided better chromatographic resolution, sensitivity, and shorter acquisition time.⁸⁷ The efficiency and speed of UPLC separation are due to smaller particles (below 2 μm) and higher pressure (12,000 to 15,000 psi).⁸⁸ Reverse-phase chromatography is used to separate nonpolar or mildly polar molecules.⁸⁹ Hydrophilic interaction chromatography (HILIC) is the choice for strong polar metabolites.⁹⁰ A concise review of separation techniques for MS-based metabolomics can be assessed at the works of Allen and McWhinney⁹¹ and Ovbude *et al.*⁹²

In MS, samples are ionized, separated according to the m/z of the compounds. In addition, mass spectrometers can also fragment them to elucidate their chemical structure more accurately (tandem mass spectrometry).⁹³ Tandem mass spectrometry (MS/MS) is composed of several processes: (i) ionization of sample molecules, (ii) mass selection of parent ions, (iii) fragmentation (iv) detection of daughter ions.⁹⁴

Most tandem mass spectrometry applications have been performed with triple quadrupole, ion trap quadrupole mass spectrometers (ITMS) and quadrupole time-of-flight (Q-TOF).

Once the sample is separated by chromatography, it passes through a quadrupole that separates and breaks down selected ions (precursor ions).⁹⁵ In addition, they use low collision energy to obtain structural molecular ion information.⁶⁰ After fragmentation of the precursor ion (MS^2), the mass spectra are produced.⁹⁵ In this step, it uses high collision energy to obtain fragment ions.⁶²

In triple quadrupole mass spectrometers, a continuous beam of ions passes through the instrument. The mass selection and fragmentation by collision induced dissociation (CID) of the precursor ions occur in different regions of the instrument (Q1 and Q2, respectively). The mass analysis of the product ions occurs by the second mass analyzer (Q3).⁹⁶ The quadrupoles Q1 and Q2 are independent of time and quadrupole Q3 is the only time-dependent function.⁹⁶ In ITMS, being tandem-in-time, pulses of precursor ions are generated and the precursor ions are selected. Through time-series or scan functions, CID and product ion mass analysis can be obtained.⁹⁴

Q-TOF-MS has a quadrupole coupled to a time-of-flight mass analyzer. After leaving the quadrupole, all ions acquired with the same kinetic energy are reaccelerated and enter the flight tube.⁹¹ The lighter ions, as they arrive at the detector first, they will have a shorter flight time than the heavier ions.⁹¹ The dispersion of kinetic energy and the spatial propagation of ions that have the same m/z can be corrected by reflectron device.⁹⁶ Two types of acquisition can be acquired in Q-TOF. The first mode (MS) provides an accurate mass scan of the fragmented precursor ion. The second mode (MS/MS) provides simultaneous fragmentation of precursor and product ions.⁹⁶

Additionally, there are some other interesting techniques which can be applied in metabolomics workflows, e.g., mass spectrometry imaging (MSI, also called IMS: imaging mass spectrometry). MSI is a powerful tool that investigates the spatial distribution of molecules, without labeling, using MALDI (matrix-assisted laser desorption/ionization), DESI (desorption electrospray ionization), or SIMS (secondary ion mass spectrometry) methods for ionization of molecules.^{97,98} MSI is able to visualize compounds that are secreted, which can be helpful in the elucidation of the action mode of one compound and their relationship with the biological mode. This technique provides complementary information about localization and distribution of molecules, which can be important to integrate OMICs tools.⁹⁹ Although this tool is interesting for better visualization of data that require spatial distribution information, some challenges of MSI need to be implemented for greater use. Among the challenges of MSI are sample preparation and identification of compounds.⁹⁸

3.3. Data analysis

Both untargeted and targeted metabolomics generate large datasets with distinct characteristics. The untargeted approach seeks to detect all measurable metabolites. The targeted approach is limited to a predefined set, with data volume depending on the number of metabolites analyzed and the degree of replication. The manual data analysis could easily lead to errors, so pattern recognition requires statistical treatment. After the samples are analyzed, the raw data needs to be processed and summarized in a data matrix.

There is a dichotomy between results simplification for the final user appreciation and deep data exploration to avoid missing critical information in data analysis. Chemometrics can organize and build visual patterns for an easier understanding of a large multivariate data analysis from complex data sets through clustering or group segregation.⁵²

Data handling and preprocessing are the pillars for robust and reliable metabolomics data before the proper analysis, providing accessible information about retention time, m/z values, and ion intensities extracted from the original data obtained from the instruments. Preprocessing is based on spectrum normalization and scaling, spectra clustering, precursor charge determination, spectrum denoising, and spectrum quality assessment. There are many open-source programs that offer data handling and preprocessing, which varies in terms of programming knowledge. XCMS, for example, can be used as an R software package demanding a moderate programming level or can be used in its website interface (XCMSOnline),¹⁰⁰ requiring no programming knowledge for data preprocessing.¹⁰¹

Principal component analysis (PCA) is the most used non-supervised chemometric method, where groups are segregated based only on their inherent differences. The PCA results in a two or three components graphic model with little lost information, and the projection of the variables responsible for the group segregation is evaluated as potential biomarkers.¹⁰² Other non-supervised methods such as hierarchical clustering, heatmap, and network analysis can be successfully applied to metabolomics studies.

In the supervised field, partial least square discriminant analysis (PLS-DA) is usually the best choice. A classification model is built based on previously labeled samples, optimizing the group segregation by the rotation of the component in search of the maximum separation.¹⁰³ PLS-DA uses a PLS regression and indicates the most important variables through a VIP (variable importance in projection) value that can be evaluated as compounds of interest. Since it is a classification model, it requires a model validation that can be achieved through a cross, internal or external validation. PLS components are built by finding a proper compromise between two purposes: describing the explanatory variables and predicting the response ones. A PLS-based classification should benefit from such a property in the direction of building typologies with an intrinsic prediction power.

Looking towards the “multi-omics age”, the metabolic pathway activity and analysis are potent tools used after the data processing and must be addressed. Data processing techniques, such as the mummichog algorithm, which is based on over-representation analysis (ORA), predicts the enriched pathway activity, comparing the significant peaks based on statistical importance and mapping metabolites in their respective biological pathway.¹⁰⁴

Same as in preprocessing, many software and website-based programs are used in multivariate statistical analysis,

each one with particular features and applications. Metaboanalyst 6¹⁰⁵ fits the description as an R-based friendly interface for metabolomics data analysis.¹⁰⁶ Data matrices can be imported, and prior to the multivariate analysis, variables are submitted to normalization and scaling. In a few steps, general statistics, biomarker analysis, two-factor/time-series analysis, and power analysis are achieved through PCA, PLS-DA, oPLS-DA (orthogonal partial least squares discriminant analysis), heatmaps, *k*-means clustering, random forest, and support vector machine, the most used methods for metabolomics analysis. The PLS-DA model is widely used because it is suitable as a variable selector and classifier when the underlying model is known or plausibly estimable. In the absence of this information, it is recommended to rely strictly on the cross-validation error and exercise extreme caution when interpreting or extrapolating results.^{82,107,108}

In summary, data preprocessing and processing methods must be wisely chosen for a correct data interpretation that can lead to the biological understanding of the study. Preprocessing must eliminate or at least diminish the technical and analytical errors derived from sample preparation and equipment drifts using data normalization and scaling. Moreover, the processing is usually based on the interpretation demand, whether group segregation, cluster visualization, or model classification.

3.4. Metabolite annotation

Compound elucidation is an essential step in metabolomics since it leads to a biological meaning through metabolic pathway visualization. The chemical and physical diversity of metabolites makes it difficult to identify them based on MS data. Information about the differentially expressed compounds can be extracted on loadings graphs (PCA), VIP scores (PLS-DA), and based on ROC (receiver operating characteristic) curves or heatmaps in the multivariate analysis. Identifying interest compounds is considered a bottleneck in untargeted metabolomics due to the great amount of known and elucidated metabolites (200,000+). Metabolic elucidation is usually performed through exact mass and isotope patterns (on high-resolution equipment), fragmentation patterns, metabolic databases, and literature comparison. Those techniques are not mutually excluded and can be used together for a more accurate metabolic elucidation.

At untargeted analyses, it is important to know the retention time and mass of metabolites of interest. Exact mass identification follows a logic where the molecular formula is generated and compared to metabolic databases based on the isotope pattern and the detected mass. The

high-resolution power of state-of-art equipment usually generates a list of molecular formulas and their respective metabolites. Therefore, it must be carefully analyzed to prevent erroneous interpretation.

Fragmentation patterns uses *m/z* and intensity values of fragment ions, retention time, neutral loss, and possible known rearrangements to elucidate metabolites, comparing the raw information with validated databases, such as Metlin,¹⁰⁹ MassBank,¹¹⁰ and Metfrag.¹¹¹ First, the *m/z* value of a molecular ion of interest is searched in a database. Then, the MS/MS spectra and retention times of the authentic compounds are compared with the molecules of interest in the sample and the identities of the molecules can be confirmed.⁶²

Metfrag¹¹¹ compares measured peaks with known fragments in candidate databases Kyoto Encyclopedia of Genes and Genomes (KEGG),¹¹² PubChem,¹¹³ ChemSpider¹¹⁴ to find metabolic matches. The results are ranked based on the fragment mass error and the number of explained peaks, with information about the fragment structures and a link to the database for a complete biological interpretation. Candidates are fragmented using a binding dissociation approach and these fragments are compared to the product ions in the measured mass spectrum to determine which candidates best explain the measured data.⁹⁹ The relationship between *m/z*, intensity and bond dissociation energy (BDE) of the combined peaks gives the candidate score, while a limited number of neutral loss rules (5 in total) are responsible for rearrangements.^{111,115}

Metfrag is a software that uses *in silico* fragmentation approach. This method is used for annotation of unknown molecules, where there is no reference spectrum for database matching.^{104,106} *In silico* fragmentation software compares the experimental spectra with a selection of fragments generated *in silico* and stored in molecular structure databases.^{116,117}

Feature-based molecular networking (FBMN) is another promising strategy for metabolite annotation and discovery in untargeted liquid chromatography-tandem mass spectrometry (LC-MS/MS) data. In FBMN analysis, LC-MS/MS spectra datasets are initially preprocessed with feature detection and alignment tools, such as MZMINE^{118,119} or XCMS.^{120,121} The obtained consensus spectra can be, for example, extracted and uploaded to the Global Natural Products Social Molecular Network (GNPS)¹²² web-platform for network assemblage and metabolite annotation by searching against the public GNPS crowdsourced MS/MS spectral libraries.¹²³ Results are usually visualized with Cytoscape¹²⁴ or MATLAB¹²⁵ software. Based on MS/MS fragmentation

pattern similarities, the detected features are organized in a molecular network, in which structurally related compounds are clustered together. Each node of the network represents a molecular feature, whereas the edges (connections) represent the degree of MS/MS fragment similarity between the features, calculated by “cosine scores”.

Molecular networking allows not only to identify previously known compounds in a sample by comparing with databases (i.e., dereplication) but also to discover novel structural analogs that are clustered with known compounds. FBMN has been increasingly applied for data visualization and annotation of untargeted metabolomics data. Cumulative data input into the GNPS database is expected to increase the elucidation of the novel microbial metabolite by this approach.¹²⁶ Since the molecular network uses MS² data and only qualitative information about the experiment is provided, some limitations such as low signal-to-noise ratio, insufficient analyte separation, and isomer separation may be crucial. In this sense, annotations of unknown molecules can be confirmed with an authentic chemical standard or by using a complementary technique, as nuclear magnetic resonance (NMR).¹²⁷

New bio/chemoinformatic tools that connect genomic and metabolomic data are also promising technologies that can expedite new microbial compounds. The Genomes-to-Natural Products (GNP) platform,¹²⁸ for example, uses a genome-guided approach to predict the chemical structure of hypothetical secondary metabolites. The predicted compounds are then fragmented *in silico* and matched with real LC-MS/MS chromatograms of crude microbial extracts, enabling their identification.¹²⁹ The inverse is also possible with tools such as the generalized retro-biosynthetic assembly prediction engine (GRAPE). The gene clusters responsible for synthesizing known polyketides and non-ribosomal peptides with yet unknown biosynthetic origin can be predicted based on the detailed chemical structure of these compounds.¹³⁰ However, these two platforms are limited to bacterial data, and their expansion to include fungal genomes is expected to drive the discovery of new fungal metabolites and their corresponding biosynthetic pathways.

4. Investigation of Metabolism of PGPY Using Mass Spectrometry-Based Metabolomics

The use of beneficial microorganisms as bioinputs is an effective tool for controlling plant diseases in agriculture, making crop management practices safer.^{131,132} Multiple biological controls are available, however, a better understanding of the interaction between plants,

environmental and pathogens is needed to define the best bioproduct to face each plant pathogen. In this way, metabolomic studies can be useful for comprehending metabolic changes of infected plants, helping to define a better strategy. Furthermore, metabolomic techniques can provide evidence about disease resistance metabolites through the study of the affected metabolic pathways.¹³³

Plant metabolic changes induced by infection of filamentous fungi can be seen through the reduction (suppression or downregulation), or even an increase (upregulation) of some concentration of the metabolite. Unveiling changes in plant metabolism in response to the plant-pathogen interactions by omics integration can be very helpful. An example is observed in experiments involving soybean types (resistance and susceptible) against *Phytophthora sojae* infection. In this case, transcriptomic (RNA-seq) and metabolomics (GC-MS) analysis suggested that many sugars (melezitose, levoglucosan, erythrose, trehalose, isomaltose, and 1-kestose), amino acids (tyramine, saccharopine, *N*-formyl-L-methionine, *N*- α -acetyl-L-ornithine, phenylacetaldehyde indole-3-acetamide, 4-hydroxybenzoic acid, *trans*-4-hydroxy-L-proline, treobeta hydroxy aspartate and *S*-carboxymethylcysteine), and secondary metabolites (mannitol, octanal, hypoxanthine and daidzein) play a defensive role against the pathogen.¹³³ Through identification of up and down regulated compounds it is still possible to identify the affected metabolic pathway, which in this case were related to primary metabolites (sugars, organics acids, and amino acids), and also secondary ones.¹³³

Primary and secondary metabolic pathways are usually affected by filamentous fungi infections. The metabolic response of maize genotypes (*Zea mays* L.) infected with *Peronosclerospora sorghi* showed that the pathogen substantially alters the metabolic profile of the plant, including significant changes in pathways of some sugars, amino acids and phenylpropanoids.¹³⁴

Another example can be seen in *Zingiber zerumbet* rhizomes after *Pythium myriotylum* infection. In this case, the pathogen altered the relative amount of the sesquiterpenoid zerumbone (ZER) and its precursor α -caryophyllene (humulene) in the plant.^{135,136} ZER, the monocyclic sesquiterpene compound, has been shown to have antiproliferative activity, anticancer, antinociceptive, antioxidant and anti-inflammatory, promoting apoptosis of tumor cells.¹³⁰

Furthermore, the omics studies can be focused on the study of enzymes and metabolites responsible for pathogenicity. For example, the hydroxycinnamic acid amide metabolites are known to be involved in abiotic and

biotic stress tolerance and can give resistance to disease in cocoa (*Theobroma cacao* L., chocolate tree).¹³⁷

The untargeted metabolomics was used to point out contrasting levels of susceptibility of *T. cacao* to *Phytophthora* spp., using LC-MS/MS.¹³⁷ Other proteomics and metabolomics (LC-MS) studies showed that *P. cinnamoni* pathogenicity access the plant stem after reprogramming of the plant proteome and accumulation of the stress-related hormones SA and JA.¹³⁸ This alteration in phytohormones as a pathogenicity mechanism was also annotation by targeted and untargeted metabolomics in *Phytophthora capsica*.¹³⁹

In addition to using metabolomics to assess changes in metabolites levels in plants, they can also be used for developing alternatives against phytopathogens. Therefore, following the metabolomics studies, strains can be applied as bioinputs in plants infected by filamentous pathogens. Whether for the understanding of altered metabolites, pathogenicity, or the pathogen control way during plant-pathogen interaction, metabolomic studies stand out due to the extensive information they can provide (Figure 4).

In vitro analysis and *in vivo* assays performed for

tomato, grape and apple fruits allowed the annotation of 35 metabolites secreted by strains of the yeast *Metschnikowia pulcherrima*, of which three compounds, namely acid 3-amino-5-methylhexanoic acid, biphenyl-2,3-diol and sinapaldehyde, were efficient in inhibiting *Botrytis cinerea* infection, the last one being the most active compound (reducing up to 90-100% of *B. cinerea* infection on tomatoes and apples).³⁵ Furthermore, the first two metabolites protected tomatoes against *Alternaria alternata* infection. Other metabolites secreted by the yeast *M. pulcherrima* have been identified with possible biotechnological applications, such as piperidine and protoemetine (alkaloids), *p*-coumaroyl quinic acid (phenylpropanoid), β -rhodomycin (antibiotic), hexadecanedioic acid (long-chain fatty acid) or taurocholic acid (bile acid).¹³⁹ Both identifications were made by UPLC-MS/MS.

Non-rhizobial endophytic yeasts, *Candida glabrata* and *Candida tropicalis*, when interacting with Rhizobium, were associated with the promotion of plant growth, producing IAA, ACC deaminase, siderophore, NH₃ and polyamine, nodulation and symbiosis in urdbean (*Vigna mungo* L.).¹⁴⁰ The yeasts were isolated from urdbean root nodules and

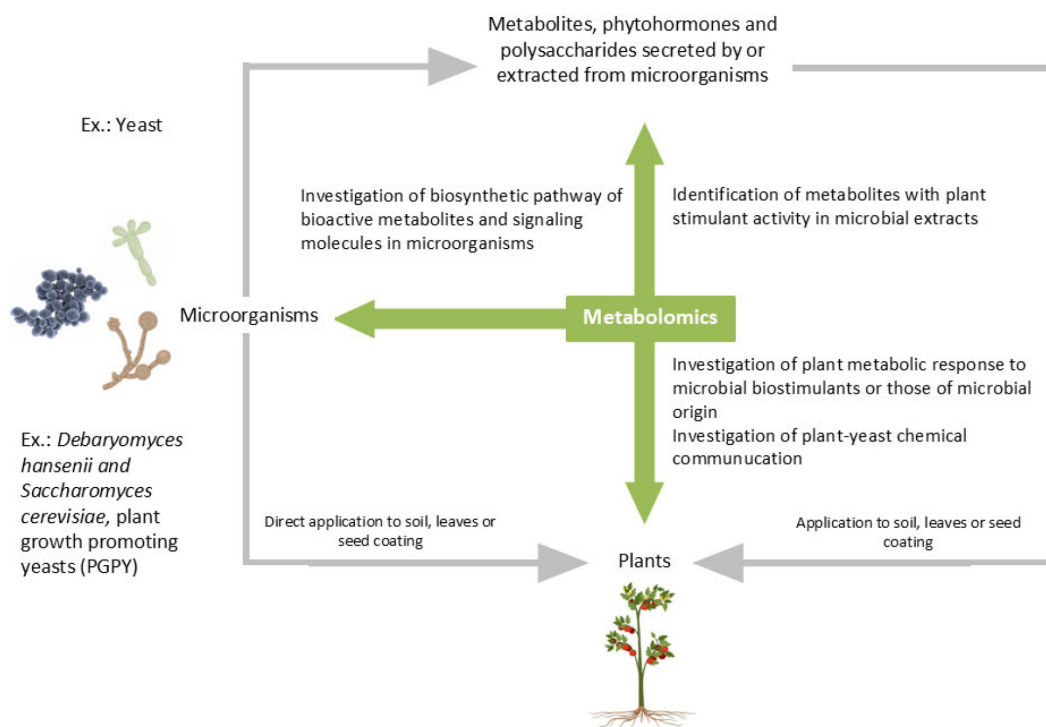


Figure 4. Possibilities of metabolomics approaches in research involving bioinputs based on plant growth promoting yeasts. Bioinputs based on yeast, such as *Debaryomyces hansenii* and *Saccharomyces cerevisiae*, can be directly employed as through the direct application of spores, mycelia, or cell suspensions to the soil, seeds, or leaves. Alternatively, bioactive metabolites, e.g., phytohormones and polysaccharides, can be extracted and then applied to the plant or soil in a process that does not rely on the direct interaction between live organisms and the plant. When microorganisms are employed as bioinputs, the metabolomics approach can be explored, for example, to investigate the plant-microbe chemical communication and the metabolic responses each organism exerts on its symbiont partner. In the scenario where metabolites (and not living microorganisms), metabolomics can be applied to unravel the main plant metabolic responses exerted by the microbial metabolites, as well as to identify bioactive metabolites in a complex microbial extract that are responsible for biocontrol activity or to investigate the biosynthetic pathways involved in the production of such bioactive metabolite.

their soluble and volatile metabolites involved during interactions were traced. The volatile profile was obtained by using a GC-MS platform customized with thermal desorber (TD). A total of 45 volatile organic compounds were annotated, including sulfonyls, organic trisulfides, fatty acyls, organic disulfides, alkanes, organooxygenic acids, phenols, benzene derivatives, diazinans, carboxylic acids and derivatives, dialkyl sulfides, alkanes, methoxybenzenes and benzophenones.¹⁴⁰

Besides metabolomics may be useful for the study of plant-pathogen interactions, it is important to understand how this study should be done and how each step should be conducted to obtain a result consistent with the research objective.¹⁴¹ The next section will deeply guide through all the steps in order to perform a mass spectrometry-based metabolomics analysis.

5. Challenges and Possible Advances of Plant Growth-Promoting Yeasts

The use of PGPY has been an area of great interest, as they are able to increase agricultural productivity directly and indirectly, contributing to soil bioremediation and the knowledge of these processes can lead to environmentally correct agricultural practices. The intrinsic characteristics of PGPY make them promising candidates in agricultural practices, promoting plant growth, providing protection against pathogens, reducing biotic and abiotic stress and favoring plant nutrition.

There is a huge discrepancy between the multitude of “biocontrol yeasts” described in scientific publications and the few yeast-based bioinputs that are commercialized. Commercial yeast-based biofertilizers include formulations containing bacteria and yeasts such as *C. tropicalis* and *S. cerevisiae*. Commercial biofungicides include products based on *Bacillus subtilis* strains, *M. pulcherrima*, *Monilinia fruticola*, *Candida oleophila*, *A. pullulans*, *Candida sake*, and *C. albidus*. For our knowledge, there is no yeast-based insecticide and herbicide commercially available.

The development of yeast-based bioinputs and bioinputs depends on several factors related to science, industry and legislation, and all of them must be together. Thus, exploring the bioinputs potential is a promising technique, although there is a need to investigate in detail the mechanisms of action by which PGPY achieves benefits in plants, the metabolic pathways and genes involved in the underlying mechanisms and plant/pathogen interactions. In addition, the search for new PGPY in varied ecosystems such as marine environments, soils and rhizosphere/phyllosphere of plants is essential.

Studies in induced resistance and chemical communication of roots with soil microorganisms can be done using advanced techniques, such as mass spectrometry-based metabolomics. With the rapid development of analytical techniques, significant advances in metabolomics research have been made to determine the interactions between phytopathogenic fungi and their hosts, fungal infection mechanisms and plant defense mechanisms. Such information can be useful for crop improvement, finding new resistant genes and targets for the development of fungicides. But there are still major challenges for metabolomics research in yeast.

A major challenge in metabolomics is the standardization of metabolomics workflows applied to bioinputs based on plant growth promoting yeasts (sample preparation, analytical techniques, data analysis, metabolite annotation, metabolite quenching and extraction). This standardization allows the construction of an integrative, efficient and reproducible workflow. The complexity of biological samples requires attention to analytical techniques regarding sensitivity, dynamic range, resolution and yields. Sample preparation is a crucial step. The success of this step guarantees a better metabolic impression of the sample. The analytical methodology must be carefully selected. Furthermore, the annotation of metabolites obtained represents a major bottleneck in metabolic analysis. Thousands of metabolites can be annotated in a mass spectrum and false metabolite annotations can occur due to isotopes and adducts. Furthermore, for a more comprehensive understanding of the structure of the compound, MS/MS information is essential. The information contained in this review will greatly support the standardization of metabolomics workflows.

Recently, the integration of multi-omics, such as genomic, transcriptomic, proteomic, and metabolomic data, allows us to identify new metabolites, metabolic pathways and investigate plant-pathogen interactions. The discovery of relevant biomarkers can be done through the combination of gene expression and regulation and protein synthesis and expression, contributing to biological research and agricultural applications, in addition to studies for future genetic modifications that favor the formation of VOCs.

In addition, it requires extensive research for industrial-scale development as a future perspective and its application in the field. It is necessary to improve the regulatory process around the production and commercialization of yeast-based products and their registration and testing must be done to make it more reliable. Studies must be done on concentration of yeast extracts, formulation

and applied dose. It is very important to identify that yeast-based bioinputs are safe to use and there is no risk associated with them, i.e., they are not human pathogens. In addition, quality and biosafety criteria for bioinputs, incentives and investments in research and innovation in plants highly responsive to bioinputs and the availability of new technologies covering new cultures, microorganisms, molecules and products must be established.

Supplementary Information

Supplementary information (table containing the main mechanisms that yeasts use to promote vegetative growth) is available free of charge at <http://jbcbs.sbq.org.br>, as PDF file.

Data Availability Statement

All data are available in the text.

Acknowledgments

The authors would like to thank the Brazilian Agricultural Research Corporation (EMBRAPA), Federal University of Goiás (UFG), Coordination for the Improvement of Higher Education Personnel (CAPES), and National Council for Scientific and Technological Development (CNPq) for support.

Author Contributions

MAS and PVA wrote the manuscript in consultation with TST, JCRN, and COGS. MAS, JCRN, and PVA have focused on the chemical analytical sessions, while TST, COGS, CAB, and FGS focused on biological applications. All authors read and approved the manuscript.



Miria A. Souza is graduated in Chemistry from the Goiano Federal Institute, she holds a Master's degree in Chemistry from the Federal University of Goiás. She is a PhD student in the Federal University of Goiás. Her research interests include plant and microorganism metabolomics applied to bioproducts using liquid chromatography coupled with mass spectrometry. She works as a Senior Metabolomics R&D Analyst at Ekoa Life Science, focusing on the bioinputs sector. His responsibilities include optimizing bioinput formulations, metabolite quantification, validating analytical methods (UPLC-PDA, UPLC-RID, and FTIR), quantifying reducing sugars and proteins, and

communicating results across the departments of the company.



Tallyta S. Teixeira is graduated in Bioprocess and Biotechnology Engineering from the Federal University of Tocantins, with a Master's degree in Biotechnology from the same university, and a PhD in Chemistry from UFG in partnership with EMBRAPA Agroenergia. She has research expertise encompassing biotechnology, Natural Products Chemistry, Analytical Chemistry, and Metabolomics. She currently works as a Senior R&D Analyst in Metabolomics at Santa Clara Group, with a focus on the bioinputs sector. Her responsibilities involve the development of products based on metabolites, stability testing, evaluation of analytical methods (DIMS, GC-FID, GC-MS/MS, UHPLC-MS/MS, UHPLC-DAD), and the communication of findings to multidisciplinary teams.



Jorge C. Rodrigues Neto is graduated in Chemistry from the Federal University of Goiás, he obtained both his Master's and doctoral degrees in Analytical Instrumentation at the same institution, in collaboration with Embrapa Agroenergia. His academic background has been focused on plant metabolomics and the application of multivariate statistical methods to complex datasets. He is currently an Assistant Researcher at Embrapa Agroenergy, where his work is dedicated to metabolomics of plants and microorganisms, applied to the development of innovative bioproducts in partnership with private companies. His research interests include the integration of analytical chemistry, chemometrics, and biotechnology for sustainable solutions.



Caio O. G. Silva is graduated in Biology from the University of Brasília, he holds MSc and PhD degrees in Molecular Biology from the same university. He has worked as research associate at Embrapa Agroenergy and the Norwegian University of Natural Sciences. After four years as an independent postdoc fellow at Technical University of Denmark, he is now Research Specialist in Lignin Biotechnology at the same institution. His main research areas are plant cell wall enzymology, lignin bioconversion, and fungal metabolomics.



Catharine A. Bomfim has bachelor's in Biotechnology from the Federal University of Bahia. She holds a master and PhD in Microbial Biology from the University of Brasília. She is R&D coordinator at Ekoa Life Sciences, currently developing activities in applied microbiology and biological inputs.



Eder A. Barbosa is graduated in Biology from the University of Brasília, he holds a master's degree and a doctorate in Molecular Biology from the same university. His expertise includes molecular biology, protein and oligosaccharide biochemistry and mass spectrometry (including MALDI-imaging) applied to metabolomics and biomolecule discovery. He has conducted postdoctoral research at Embrapa and the University of Brasília, focusing on plant-pathogen interactions and glycosylation mechanisms. Currently, he works as an innovation fellow at Embrapa Agroenergia (Brasília-DF) assisting in the development of sustainable bioproducts.



Félix G. de Siqueira is graduated in Biology from the Federal University of Lavras, he holds a Master's degree in Agricultural Microbiology in the same university. His PhD was in Molecular Biology in the University of Brasília, and a part of work was done in Purdue University. He is a researcher at Embrapa Agroenergia at the Laboratory of Biochemical Processes. He has experience in industrial microbiology, enzymology and bioprospecting in filamentous fungi with biotechnological potential. He also works with biological pre-treatment using macrofungi and other filamentous fungi, cultivating them in residual plant biomass to obtain bio-inputs and biofuels in the biorefinery and circular economy model.



Patrícia V. Abdelnur is graduated in Chemistry from the Federal University of São Carlos, she holds a master's degree in Organic Chemistry from the same university and a doctorate from the State University of Campinas. She completed a sandwich PhD at the Mass Spectrometry Research Center at Vanderbilt University in the USA. She is an associate professor in the postgraduate program in Chemistry at the Federal University of Goiás and a Researcher at Embrapa Agroenergia. Her research area is

Metabolomics of Plants and Microorganisms applied to Bioenergy, Bioproducts and Biotechnology. She is currently a Guest Scientist in the Helmholtz-Zentrum für Umweltforschung (UFZ) in Germany.

References

1. Kumari, M.; Swarupa, P.; Kesari, K. K.; Kumar, A.; *Life* **2022**, 13, 12. [Crossref]
2. Ministério da Agricultura e Pecuária; *Bioinputs - Increasing Brazilian Biodiversity*. [Link] accessed in October 2025
3. Kumar, V. V. In *Role of Rhizospheric Microbes in Soil*, vol. 1; Meena, V. S., ed.; Springer: Singapore, 2018, p. 377-398. [Crossref]
4. Baldani, J. I.; Baldani, V. L. D.; *An. Acad. Bras. Cienc.* **2005**, 77, 549. [Crossref]
5. Jacquet, F.; Jeuffroy, M.-H.; Jouan, J.; Le Cadre, E.; Litrico, I.; Malausa, T.; Reboud, X.; Huyghe, C.; *Agron. Sustainable Dev.* **2022**, 42, 8. [Crossref]
6. CropLife Brazil Research; *Bioinputs in Brazil, Industry Market Data*. [Link] accessed in October 2025
7. Sudheer, S.; Johny, L.; Srivastava, S.; Adholeya, A.; *Symbiosis* **2023**, 89, 259. [Crossref]
8. Mazoyon, C.; Firmin, S.; Bensaddek, L.; Pecourt, A.; Chabot, A.; Faucon, M.-P.; Sarazin, V.; Dubois, F.; Duclercq, J.; *Microorganisms* **2023**, 11, 1847. [Crossref]
9. Palmieri, D.; Ianiri, G.; Del Grosso, C.; Barone, G.; De Curtis, F.; Castoria, R.; Lima, G.; *Horticulturae* **2022**, 8, 577. [Crossref]
10. Nimsi, K. A.; Manjusha, K.; Kathiresan, K.; Arya, H.; *J. Appl. Microbiol.* **2023**, 134, 1xac088. [Crossref]
11. Knop, M.; *C. R. Biol.* **2011**, 334, 599. [Crossref]
12. Jofre, F. M.; Hernández-Pérez, A. F.; dos Santos, J. C.; Felipe, M. G. A.; *Ind. Crops Prod.* **2021**, 170, 113812. [Crossref]
13. Mohiuddin, O.; Harvey, A. P.; Ledesma, M. T. O.; Velasquez-Orta, S.; *Electron. J. Biotechnol.* **2024**, 69, 30. [Crossref]
14. Hammad, S. A. R.; Ali, O. A. M.; *Ann. Agric. Sci.* **2014**, 59, 133. [Crossref]
15. Nassar, A. H.; El-Tarabily, K. A.; Sivasithamparam, K.; *Biol. Fertil. Soils* **2005**, 42, 97. [Crossref]
16. Amprayn, K.; Rose, M. T.; Kecskés, M.; Pereg, L.; Nguyen, H. T.; Kennedy, I. R.; *Appl. Soil Ecol.* **2012**, 61, 295. [Crossref]
17. Fu, S.-F.; Sun, P.-F.; Lu, H.-Y.; Wei, J.-Y.; Xiao, H.-S.; Fang, W.-T.; Cheng, B.-Y.; Chou, J.-Y.; *Fungal Biol.* **2016**, 120, 433. [Crossref]
18. Mundra, S.; Arora, R.; Stobdan, T.; *World J. Microbiol. Biotechnol.* **2011**, 27, 2387. [Crossref]
19. Nakayan, P.; Hameed, A.; Singh, S.; Young, L.-S.; Hung, M.-H.; Young, C.-C.; *Plant Soil* **2013**, 373, 301. [Crossref]
20. Hernández-Fernández, M.; Cordero-Bueso, G.; Ruiz-Muñoz, M.; Cantoral, J. M.; *Plants* **2021**, 10, 822. [Crossref]

21. Yanagishima, N.; Shimoda, C.; *Plant Cell Physiol.* **1967**, *8*, 109. [Crossref]
22. Marques, A. R.; Resende, A. A.; Gomes, F. C. O.; Santos, A. R. O.; Rosa, C. A.; Duarte, A. A.; de Lemos-Filho, J. P.; dos Santos, V. L.; *Braz. J. Microbiol.* **2021**, *52*, 1417. [Crossref]
23. Rehman, F.; Kalsoom, M.; Adnan, M.; Toor, M. D.; Zulficar, A.; *SunText Rev. Biotechnol.* **2020**, *1*, 110. [Crossref]
24. Oleńska, E.; Małek, W.; Wójcik, M.; Swiecicka, I.; Thijs, S.; Vangronsveld, J.; *Sci. Total Environ.* **2020**, *743*, 140682. [Crossref]
25. Saini, S.; Kaur, N.; Pati, P. K.; *Ecotoxicol. Environ. Saf.* **2021**, *223*, 112578. [Crossref]
26. Labeeuw, L.; Khey, J.; Bramucci, A. R.; Atwal, H.; de la Mata, A. P.; Harynuk, J.; Case, R. J.; *Front. Microbiol.* **2016**, *7*, 828. [Crossref]
27. Casanova-Sáez, R.; Mateo-Bonmatí, E.; Ljung, K.; *Cold Spring Harbor Perspect. Biol.* **2021**, *13*, a039867. [Crossref]
28. Werner, T.; Schmölling, T.; *Curr. Opin. Plant Biol.* **2009**, *12*, 527. [Crossref]
29. Sosnowski, J.; Truba, M.; Vasileva, V.; *Agriculture* **2023**, *13*, 724. [Crossref]
30. Arraes, F. B. M.; Beneventi, M. A.; de Sa, M. E. L.; Paixão, J. F. R.; Albuquerque, E. V. S.; Marin, S. R. R.; Purgatto, E.; Nepomuceno, A. L.; Grossi-de-Sa, M. F.; *BMC Plant Biol.* **2015**, *15*, 213. [Crossref]
31. Kildegaard, K. R.; Arnesen, J. A.; Adiego-Pérez, B.; Rago, D.; Kristensen, M.; Klitgaard, A. K.; Hansen, E. H.; Hansen, J.; Borodina, I.; *Metab. Eng.* **2021**, *66*, 1. [Crossref]
32. Orozco-Mosqueda, M. C.; Glick, B. R.; Santoyo, G.; *Microbiol. Res.* **2020**, *235*, 126439. [Crossref]
33. Jha, C. K.; Sharma, P.; Shukla, A.; Parmar, P.; Patel, R.; Goswami, D.; Saraf, M.; *Physiol. Mol. Plant Pathol.* **2021**, *115*, 101664. [Crossref]
34. Arnesen, J. A.; Jacobsen, I. H.; Dyekjær, J. D.; Rago, D.; Kristensen, M.; Klitgaard, A. K.; Randelovic, M.; Martinez, J. L.; Borodina, I.; *FEMS Yeast Res.* **2022**, *22*, foac015. [Crossref]
35. Fernandez-San Millan, A.; Farran, I.; Larraya, L.; Ancin, M.; Arregui, L. M.; Veramendi, J.; *Microbiol. Res.* **2020**, *237*, 126480. [Crossref]
36. Khan, Md. S.; Zaidi, A.; Ahmad, E. In *Phosphate Solubilizing Microorganisms*; Khan, M. S.; Zaidi, A.; Musarrat, J., eds.; Springer International Publishing: Cham, 2014, p. 31. [Crossref]
37. Trivedi, P.; Sa, T.; *Curr. Microbiol.* **2008**, *56*, 140. [Crossref]
38. Kowalska, J.; Krzyńska, J.; Tyburski, J.; *Agriculture* **2022**, *12*, 1404. [Crossref]
39. Mukherjee, S.; Sen, S. K.; *J. Sci. Food Agric.* **2015**, *95*, 1491. [Crossref]
40. Fedotov, G. N.; Shoba, S. A.; Fedotova, M. F.; Stepanov, A. L.; Streletsky, R. A.; *Eurasian Soil Sci.* **2017**, *50*, 573. [Crossref]
41. Tapia-Vázquez, I.; Sánchez-Cruz, R.; Arroyo-Domínguez, M.; Lira-Ruan, V.; Sánchez-Reyes, A.; Sánchez-Carbente, M. R.; Padilla-Chacón, D.; Batista-García, R. A.; Folch-Mallol, J. L.; *Microbiol. Res.* **2020**, *232*, 126394. [Crossref]
42. Ruspi, C.; Pierantoni, D. C.; Conti, A.; Favaro, L.; Antinori, M. E.; Puglisi, E.; Corte, L.; Cardinali, G.; *Curr. Plant Biol.* **2024**, *39*, 100357. [Crossref]
43. Ferraz, P.; Cássio, F.; Lucas, C.; *Front. Microbiol.* **2019**, *10*, 1766. [Crossref]
44. Zhang, Q.; Zhao, L.; Li, B.; Gu, X.; Zhang, X.; Boateng, N. S.; Zhang, H.; *Biol. Control* **2020**, *148*, 104305. [Crossref]
45. Bianchi, F.; Spitaler, U.; Castellan, I.; Cossu, C. S.; Brigadoi, T.; Duménil, C.; Angeli, S.; Robatscher, P.; Vogel, R. F.; Schmidt, S.; Eisenstecken, D.; *Insects* **2020**, *11*, 810. [Crossref]
46. Magri, M. M.; Tauk-Tornisielo, S. M.; Ceccato-Antonini, S. R.; *Braz. Arch. Biol. Technol.* **2011**, *54*, 1. [Crossref]
47. German, J. B.; Hammock, B. D.; Watkins, S. M.; *Metabolomics* **2005**, *1*, 3. [Crossref]
48. Danzi, F.; Pacchiana, R.; Mafficini, A.; Scupoli, M. T.; Scarpa, A.; Donadelli, M.; Fiore, A.; *Signal Transduction Targeted Ther.* **2023**, *8*, 137. [Crossref]
49. Rodrigues-Neto, J. C. R.; Vieira, L. R.; Ribeiro, J. A. A.; de Sousa, C. A. F.; Souza Júnior, M. T.; Abdelnur, P. V.; *Sci. Rep.* **2021**, *11*, 18271. [Crossref]
50. Shriver, L. P.; *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 1. [Crossref]
51. Campos, C. G.; Veras, H. C. T.; Ribeiro, J. A. A.; Costa, P. P. K. G.; Araújo, K. P.; Rodrigues, C. M.; de Almeida, J. R. M.; Abdelnur, P. V.; *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 2646. [Crossref]
52. Rodrigues-Neto, J. C.; Correia, M. V.; Souto, A. L.; Ribeiro, J. A. A.; Vieira, L. R.; Souza Jr., M. T.; Rodrigues, C. M.; Abdelnur, P. V.; *Metabolomics* **2018**, *14*, 142. [Crossref]
53. Villas-Bôas, S. G.; Mas, S.; Åkesson, M.; Smedsgaard, J.; Nielsen, J.; *Mass Spectrom. Rev.* **2005**, *24*, 613. [Crossref]
54. Tugizimana, F.; Piater, L.; Dubery, I.; *South Afr. J. Sci.* **2013**, *109*, 11. [Crossref]
55. Kotta, L. R.; Vijayalakshmi, A.; *Ann. Phytomed.* **2024**, *13*, 469. [Crossref]
56. Amer, B.; Deshpande, R. R.; Bird, S. S.; *Metabolites* **2023**, *13*, 648. [Crossref]
57. Pinu, F. R.; Villas-Boas, S. G.; *Metabolites* **2017**, *7*, 43. [Crossref]
58. Liu, F.; Wu, T.; Tian, A.; He, C.; Bi, X.; Lu, Y.; Yang, K.; Xia, W.; Ye, J.; *Anal. Chim. Acta* **2023**, *1279*, 341809. [Crossref]
59. Urzì, C.; Hertig, D.; Meyer, C.; Maddah, S.; Nuoffer, J.-M.; Vermathen, P.; *Int. J. Mol. Sci.* **2022**, *23*, 6555. [Crossref]
60. Blainey, P.; Krzywinski, M.; Altman, N.; *Nat. Methods* **2014**, *11*, 879. [Crossref]
61. Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.-M.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi,

- R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reilly, M. D.; Thaden, J. J.; Viant, M. R.; *Metabolomics* **2007**, *3*, 211. [Crossref]
62. Zeki, Ö. C.; Eylem, C. C.; Reçber, T.; Kir, S.; Nemutlu, E.; *J. Pharm. Biomed. Anal.* **2020**, *190*, 113509. [Crossref]
63. Cajka, T.; Fiehn, O.; *Anal. Chem.* **2016**, *88*, 524. [Crossref]
64. Berg, M.; Vanaerschot, M.; Jankevics, A.; Cuypers, B.; Breitling, R.; Dujardin, J.-C.; *Comput. Struct. Biotechnol. J.* **2013**, *4*, e201301002. [Crossref]
65. Zhou, B.; Xiao, J. F.; Tuli, L.; Ransom, H. W.; *Mol. Biosyst.* **2012**, *8*, 470. [Crossref]
66. bioRender. [Link] accessed in October 2025
67. Abdelnur, P. V.; Caldana, C.; Martins, M. C. M.; *Chem. Biol. Technol. Agric.* **2014**, *1*, 22. [Crossref]
68. Dunn, W. B.; Winder, C. L. In *Methods in Enzymology*, vol. 500; Jameson, D.; Verma, M.; Westerhoff, H. V., eds.; Elsevier, 2011, p. 277-297. [Crossref]
69. Wang, G.; Wang, X.; Wang, T.; van Gulik, W.; Noorman, H. J.; Zhuang, Y.; Chu, J.; Zhang, S.; *Biotechnol. J.* **2019**, *14*, 1900009. [Crossref]
70. Lei, Z.; Huhman, D. V.; Sumner, L. W.; *J. Biol. Chem.* **2011**, *286*, 25435. [Crossref]
71. Lv, W.; Shi, X.; Wang, S.; Xu, G.; *TrAC, Trends Anal. Chem.* **2019**, *120*, 115302. [Crossref]
72. Hajnajafi, K.; Iqbal, M. A.; *Proteome Sci.* **2025**, *23*, 5. [Crossref]
73. Lorenz, M. A.; Burant, C. F.; Kennedy, R. T.; *Anal. Chem.* **2011**, *83*, 3406. [Crossref]
74. Gathungu, R. M.; Kautz, R.; Kristal, B. S.; Bird, S. S.; Vouras, P.; *Mass Spectrom. Rev.* **2020**, *39*, 35. [Crossref]
75. Garrett, R.; Vaz, B. G.; Hovell, A. M. C.; Eberlin, M. N.; Rezende, C. M.; *J. Agric. Food Chem.* **2012**, *60*, 4253. [Crossref]
76. González-Domínguez, R.; González-Domínguez, Á.; Sayago, A.; Fernández-Recamales, Á.; *Metabolites* **2020**, *10*, 229. [Crossref]
77. Habchi, B.; Alves, S.; Paris, A.; Rutledge, D. N.; Rathahao-Paris, E.; *TrAC, Trends Anal. Chem.* **2016**, *85*, 128. [Crossref]
78. Garcia, D. E.; Baidoo, E. E.; Benke, P. I.; Pingitore, F.; Tang, Y. J.; Villa, S.; Keasling, J. D.; *Curr. Opin. Microbiol.* **2008**, *11*, 233. [Crossref]
79. Furey, A.; Moriarty, M.; Bane, V.; Kinsella, B.; Lehane, M.; *Talanta* **2013**, *115*, 104. [Crossref]
80. Panuwet, P.; Hunter Jr., R. E.; D'Souza, P. E.; Chen, X.; Radford, S. A.; Cohen, J. R.; Marder, M. E.; Kartavenka, K.; Ryan, P. B.; Barr, D. B.; *Crit. Rev. Anal. Chem.* **2016**, *46*, 93. [Crossref]
81. King, R.; Bonfiglio, R.; Fernandez-Metzler, C.; Miller-Stein, C.; Olah, T.; *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 942. [Crossref]
82. Gromski, P. S.; Muhamadali, H.; Ellis, D. I.; Xu, Y.; Correa, E.; Turner, M. L.; Goodacre, R.; *Anal. Chim. Acta* **2015**, *879*, 10. [Crossref]
83. Zhang, A.; Sun, H.; Wang, P.; Han, Y.; Wang, X.; *Analyst* **2012**, *137*, 293. [Crossref]
84. Fiehn, O.; *Curr. Protoc. Mol. Biol.* **2016**, *114*, 21.33.1. [Crossref]
85. Mulvaney, R. L. In *Nitrogen Isotope Techniques*; Knowles, R.; Blackburn, T. H., eds.; Academic Press, 1993, ch. 2. [Crossref]
86. Giavalisco, P.; Li, Y.; Matthes, A.; Eckhardt, A.; Hubberten, H.-M.; Hesse, H.; Segu, S.; Hummel, J.; Köhl, K.; Willmitzer, L.; *Plant J.* **2011**, *68*, 364. [Crossref]
87. Alanazi, S.; *Anal. Sci. Adv.* **2025**, *6*, e70024. [Crossref]
88. Raftery, D.; *Mass Spectrometry in Metabolomics: Methods and Protocols*; Springer: New York, 2014. [Crossref]
89. Roberts, L. D.; Souza, A. L.; Gerszten, R. E.; Clish, C. B.; *Curr. Protoc. Mol. Biol.* **2012**, *98*, 30.2.1. [Crossref]
90. Bajad, S. U.; Lu, W.; Kimball, E. H.; Yuan, J.; Peterson, C.; Rabinowitz, J. D.; *J. Chromatogr. A* **2006**, *1125*, 76. [Crossref]
91. Allen, D.; McWhinney, B.; *Clin. Biochem. Rev.* **2019**, *40*, 135. [Link] accessed in October 2025
92. Ovbude, S. T.; Sharmeen, S.; Kyei, I.; Olupathage, H.; Jones, J.; Bell, R. J.; Powers, R.; Hage, D. S.; *J. Chromatogr. B* **2024**, *1239*, 124124. [Crossref]
93. van Agthoven, M. A.; Lam, Y. P. Y.; O'Connor, P. B.; Rolando, C.; Delsuc, M.-A.; *Eur. Biophys. J.* **2019**, *48*, 213. [Crossref]
94. Johnson, J. V.; Yost, R. A.; Kelley, P. E.; Bradford, D. C.; *Anal. Chem.* **1990**, *62*, 2162. [Crossref]
95. Mittal, R. D.; *Indian J. Clin. Biochem.* **2015**, *30*, 121. [Crossref]
96. Chernushevich, I. V.; Loboda, A. V.; Thomson, B. A.; *J. Mass Spectrom.* **2001**, *36*, 849. [Crossref]
97. Esquenazi, E.; Yang, Y.-L.; Watrous, J.; Gerwick, W. H.; Dorrestein, P. C.; *Nat. Prod. Rep.* **2009**, *26*, 1521. [Crossref]
98. Buchberger, A. R.; DeLaney, K.; Johnson, J.; Li, L.; *Anal. Chem.* **2018**, *90*, 240. [Crossref]
99. Dunham, S. J. B.; Ellis, J. F.; Li, B.; Sweedler, J. V.; *Acc. Chem. Res.* **2017**, *50*, 96. [Crossref]
100. XCMSOnline - Metabolomic and Lipidomic Platform. [Link] accessed in October 2025
101. Gowda, H.; Ivanisevic, J.; Johnson, C. H.; Kurczy, M. E.; Benton, H. P.; Rinehart, D.; Nguyen, T.; Ray, J.; Kuehl, J.; Arevalo, B.; Westenskow, P. D.; Wang, J.; Arkin, A. P.; Deutschbauer, A. M.; Patti, G. J.; Siuzdak, G.; *Anal. Chem.* **2014**, *86*, 6931. [Crossref]
102. Cordella, C. B. Y. In *Analytical Chemistry*; Krull, I. S., ed.; InTech, 2012. [Crossref]
103. Barker, M.; Rayens, W.; *J. Chemom.* **2003**, *17*, 166. [Crossref]
104. Li, S.; Park, Y.; Duraisingham, S.; Strobel, F. H.; Khan, N.; Soltow, Q. A.; Jones, D. P.; Pulendran, B.; *PLoS Comput. Biol.* **2013**, *9*, e1003123. [Crossref]
105. MetaboAnalyst 6.0. [Link] accessed in October 2025
106. Pang, Z.; Chong, J.; Zhou, G.; Morais, D. A. L.; Chang, L.; Barrette, M.; Gauthier, C.; Jacques, P.-É.; Li, S.; Xia, J.; *Nucleic Acids Res.* **2021**, *49*, W388. [Crossref]
107. Ruiz-Perez, D.; Guan, H.; Madhivanan, P.; Mathee, K.; Narasimhan, G.; *BMC Bioinf.* **2020**, *21*, 2. [Crossref]

108. Bro, R.; *Chemom. Intell. Lab. Syst.* **2025**, *264*, 105461. [Crossref]
109. Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G.; *Ther. Drug Monit.* **2005**, *27*, 747. [Crossref]
110. Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; Oda, Y.; Kakazu, Y.; Kusano, M.; Tohge, T.; Matsuda, F.; Sawada, Y.; Hirai, M. Y.; Nakanishi, H.; Ikeda, K.; Akimoto, N.; Maoka, T.; Takahashi, H.; Ara, T.; Sakurai, N.; Suzuki, H.; Shibata, D.; Neumann, S.; Iida, T.; Tanaka, K.; Funatsu, K.; Matsuura, F.; Soga, T.; Taguchi, R.; Saito, K.; Nishioka, T.; *J. Mass Spectrom.* **2010**, *45*, 703. [Crossref]
111. Ruttkies, C.; Schymanski, E. L.; Wolf, S.; Hollender, J.; Neumann, S.; *J. Cheminf.* **2016**, *8*, 3. [Crossref]
112. KEGG: Kyoto Encyclopedia of Genes and Genomes. [Link] accessed in October 2025
113. PubChem. [Link] accessed in October 2025
114. ChemSpider. [Link] accessed in October 2025
115. Wolf, S.; Schmidt, S.; Müller-Hannemann, M.; Neumann, S.; *BMC Bioinf.* **2010**, *11*, 148. [Crossref]
116. Böcker, S.; *Curr. Opin. Chem. Biol.* **2017**, *36*, 1. [Crossref]
117. Blaženović, I.; Kind, T.; Ji, J.; Fiehn, O.; *Metabolites* **2018**, *8*, 31. [Crossref]
118. Pluskal, T.; Korf, A.; Smirnov, A.; Schmid, R.; Fallon, T. R.; Du, X.; Weng, J.-K. In *Processing Metabolomics and Proteomics Data with Open Software: A Practical Guide*; Winkler, R., ed.; The Royal Society of Chemistry: London, 2020, p. 232-254. [Crossref]
119. MZMINE. [Link] accessed in October 2025
120. Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G.; *Anal. Chem.* **2006**, *78*, 779. [Crossref]
121. Domingo-Almenara, X.; Siuzdak, G. In *Computational Methods and Data Analysis for Metabolomics*; Li, S., ed.; Springer US: New York, 2020, p. 11-24. [Crossref]
122. GNPS: Global Natural Products Social Molecular Networking. [Link] accessed in October 2025
123. Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kaponov, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W.-T.; Crisemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderón, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C.-C.; Floros, D. J.; Gavilan, R. G.; Kleigrewe, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C.-C.; Yang, Y.-L.; Humpf, H.-U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; Boya, P. C. A.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodríguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P.-M.; Phapale, P.; Nothias, L.-F.; Alexandrov, T.; Litaudon, M.; Wolfender, J.-L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D.-T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. Ø.; Pogliano, K.; Linington, R. G.; Gutiérrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N.; *Nat. Biotechnol.* **2016**, *34*, 828. [Crossref]
124. Cytoscape. [Link] accessed in October 2025
125. MATLAB. [Link] accessed in October 2025
126. Nothias, L.-F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; Sarvepalli, A.; Protsyuk, I.; Ernst, M.; Tsugawa, H.; Fleischauer, M.; Aicheler, F.; Aksenov, A. A.; Alka, O.; Allard, P.-M.; Barsch, A.; Cachet, X.; Caraballo-Rodríguez, A. M.; da Silva, R. R.; Dang, T.; Garg, N.; Gauglitz, J. M.; Gurevich, A.; Isaac, G.; Jarmusch, A. K.; Kameník, Z.; Kang, K. B.; Kessler, N.; Koester, I.; Korf, A.; Le Gouellec, A.; Ludwig, M.; Martin, H. C.; McCall, L.-I.; McSayles, J.; Meyer, S. W.; Mohimani, H.; Morsy, M.; Moyne, O.; Neumann, S.; Neuweger, H.; Nguyen, N. H.; Nothias-Esposito, M.; Paolini, J.; Phelan, V. V.; Pluskal, T.; Quinn, R. A.; Rogers, S.; Shrestha, B.; Tripathi, A.; Van Der Hooft, J. J. J.; Vargas, F.; Weldon, K. C.; Witting, M.; Yang, H.; Zhang, Z.; Zubeil, F.; Kohlbacher, O.; Böcker, S.; Alexandrov, T.; Bandeira, N.; Wang, M.; Dorrestein, P. C.; *Nat. Methods* **2020**, *17*, 905. [Crossref]
127. Aron, A. T.; Gentry, E. C.; McPhail, K. L.; Nothias, L.-F.; Nothias-Esposito, M.; Bouslimani, A.; Petras, D.; Gauglitz, J. M.; Sikora, N.; Vargas, F.; Van Der Hooft, J. J. J.; Ernst, M.; Kang, K. B.; Aceves, C. M.; Caraballo-Rodríguez, A. M.; Koester, I.; Weldon, K. C.; Bertrand, S.; Roullier, C.; Sun, K.; Tehan, R. M.; Boya, P. C. A.; Christian, M. H.; Gutiérrez, M.; Ulloa, A. M.; Tejeda Mora, J. A.; Mojica-Flores, R.; Lakey-Beitia, J.; Vázquez-Chaves, V.; Zhang, Y.; Calderón, A. I.; Tayler, N.; Keyzers, R. A.; Tugizimana, F.; Ndlovu, N.; Aksenov, A. A.; Jarmusch, A. K.; Schmid, R.; Truman, A. W.; Bandeira, N.; Wang, M.; Dorrestein, P. C.; *Nat. Protoc.* **2020**, *15*, 1954. [Crossref]
128. Johnston, C. W.; Skinnider, M. A.; Wyatt, M. A.; Li, X.; Ranieri, M. R. M.; Yang, L.; Zechel, D. L.; Ma, B.; Magarvey, N. A.; *Nat. Commun.* **2015**, *6*, 8421. [Crossref]
129. Blin, K.; Wolf, T.; Chevrette, M. G.; Lu, X.; Schwalen, C. J.; Kautsar, S. A.; Suarez Duran, H. G.; de los Santos, E. L. C.;

- Kim, H. U.; Nave, M.; Dickschat, J. S.; Mitchell, D. A.; Shelest, E.; Breitling, R.; Takano, E.; Lee, S. Y.; Weber, T.; Medema, M. H.; *Nucleic Acids Res.* **2017**, *45*, W36. [Crossref]
130. Dejong, C. A.; Chen, G. M.; Li, H.; Johnston, C. W.; Edwards, M. R.; Rees, P. N.; Skinnider, M. A.; Webster, A. L. H.; Magarvey, N. A.; *Nat. Chem. Biol.* **2016**, *12*, 1007. [Crossref]
131. Gardener, B. B. M.; Fravel, D. R.; *Plant Health Prog.* **2002**, *3*. [Crossref]
132. Thambugala, K. M.; Daranagama, D. A.; Phillips, A. J. L.; Kannangara, S. D.; Promputtha, I.; *Front. Cell. Infect. Microbiol.* **2020**, *10*, 604923. [Crossref]
133. Zhu, L.; Zhou, Y.; Li, X.; Zhao, J.; Guo, N.; Xing, H.; *Front. Plant Sci.* **2018**, *9*, 1530. [Crossref]
134. Sireesha, Y.; Sekhar, M.; Velazhahan, R.; *J. Pure Appl. Microbiol.* **2015**, *9*, 1355. [Link] accessed in October 2025
135. Keerthi, D.; Geethu, C.; Nair, R. A.; Pillai, P.; *Appl. Biochem. Biotechnol.* **2014**, *172*, 2593. [Crossref]
136. Jegannathan, S. D.; Arul, S.; Dayalan, H.; *Nutr. Cancer* **2016**, *68*, 865. [Crossref]
137. Knollenberg, B. J.; Li, G.-X.; Lambert, J. D.; Maximova, S. N.; Guiltinan, M. J.; *Front. Plant Sci.* **2020**, *11*, 617520. [Crossref]
138. Saiz-Fernández, I.; Milenković, I.; Berka, M.; Černý, M.; Tomšovský, M.; Brzobohatý, B.; Kerchev, P.; *Int. J. Mol. Sci.* **2020**, *21*, 8525. [Crossref]
139. Mhlongo, M. I.; Piater, L. A.; Steenkamp, P. A.; Labuschagne, N.; Dubery, I. A.; *Metabolites* **2020**, *10*, 466. [Crossref]
140. Thanuja, K.; Annadurai, B.; Thankappan, S.; Uthandi, S.; *Arch. Microbiol.* **2020**, *202*, 2739. [Crossref]
141. Castro-Moretti, F. R.; Gentzel, I. N.; Mackey, D.; Alonso, A. P.; *Metabolites* **2020**, *10*, 52. [Crossref]

Submitted: August 21, 2025

Final version online: October 22, 2025