Biological Times

Volume No. 4 | Issue No. 12 | Pages 39-41

Metabolite Profiling and Analytical Techniques in Plant Genetics and Crop Improvement

Ayaz Haider Khan*[1], Fatima Toufiq[2], Tahira Karamat[2], Abdul Kareem[1] and Aftab Akhtar[2]

1. Department of Plant Breeding and Genetics, University of Agriculture Faisalabad, Pakistan
2. Department of Chemistry, University of Agriculture Faisalabad, Pakistan

*Corresponding Author: **mrayazhaiderkhan@gmail.com**

ABSTRACT

Plant breeding should be able to enhance yield, nutritional and stress resistance. Comprehensive profiling of small molecules is called metabolomics which joins genotype, environment and phenotype and therefore becomes an irreplaceable toolkit in the study of plant genetics and crop enhancement. Hundreds of thousands of primary and secondary metabolites are found in plant tissues and it has been estimated that plant metabolome has over 200 000 compounds, though only a small fraction of these have been deposited in publicly accessible databases. It is now feasible to identify thousands of metabolite features using the high throughput platforms including gas chromatology mass spectrometry (GC MS), liquid chromatology mass spectrometry (LC MS), nuclear magnetic resonance (NMR) spectroscopy and capillary electrophoresis mass spectrometry (CE MS) although the metabolite identification continues to act as a bottleneck. Untargeted LC MS data sets normally annotate 2 to 15 percent of detected peaks.

Keywords: Anti-nutritional factors (ANFs); CRISPR-Cas9; Genome editing; Gossypol; Phytic acid; Plant breeding; Metabolite profiling; Plant metabolomics

To cite this article: Khan AH, F Toufiq, T Karamat, A Kareem, A Akhtar. Metabolite Profiling and Analytical Techniques in Plant Genetics and Crop Improvement. Biological Times. 2025. December 4(12): 39-41.

Introduction

The metabolome is the end products of gene expression and is directly correlated with the physiological phenotypes. Plants produce an enormous range of primary metabolites such as sugars, amino acids, organic acids and lipids and secondary metabolites such as phenolics, terpenoids, alkaloids and others that result in growth, stress responses and interactions with other microbes [1, 2]. Recent studies indicate that plants can synthesize more than 200 000 metabolites; yet in the Universe of public databases like HMDB, KNApSAcK or METLIN we can find a couple thousand metabolites in plants. The implication of this is that metabolomics is becoming a critical instrument that can be used by plant scientists to study metabolic diversity, as well as to relate phenotype to genotype, and in this regard, to improve crop species [1, 3].

Metabolomics involves untargeted profiling, in which metabolites are scanned to as many compounds as possible and targeted profiling, which targets particular compounds or pathways. Targeted methods offer great sensitivity and precise quantification of established metabolites whereas untargeted methods identify new or unexpected chemicals although with challenges to identification and quantification [4]. Since metabolome is an indicator of interactions between genes and the environment, metabolomic profiles can demonstrate genotype, developmental stages and environmental conditions variation and give indicators of selection in breeding programs.

Metabolomic Analytical Techniques and Platforms.

Gas Chromatography Mass Spectrometry (GCMS).

Detection is done by electron impact mass spectrometry and separation is done by gas chromatography of the derivatized compounds[7]. GCMS has a high chromatographic resolution, which is suitable at separating isomeric metabolites, and has strong and reproducible spectral libraries. Other than the orthodox single dimensional GCMS, two dimensional comprehensive GC improves the sample throughput and separates the co eluting peaks to increase the detection of the complex plant matrices[1]. The use of GCMS in the stress-induced primary metabolite profiling has become common. As an illustration, drought stressed rice and wheat contain high amounts of amino acid like leucine, isoleucine, valine and proline [8]. Profiling of eight wheat cultivars using GCMS also demonstrated the higher concentration of amino acids during drought. In rice, it was found that GCMS detected production of more hydroxycinnamic acid, ferulic acid and other phenolic acids during times of drought. These cases demonstrate that GCMS may be used to detect the metabolites related to the tolerance of abiotic stress that could be used as breeding biomarkers[9].

Liquid Chromatography Mass Spectrommetry (LCMS).

High resolution LCMS instruments are a combination of liquid chromatography and electrospray ionisation or atmospheric pressure chemical ionisation, and enable the detection of both primary and secondary metabolites [10]. LCMS can be used in targeted mode with triple quadrupole MS to be used to obtain precise quantification of results or in untargeted mode with high resolution instruments to obtain full scan and MSMS spectral results. Since a large fraction of the metabolites produced by plants can be identified only by their mass, the datasets produced by LCMS usually identify thousands of features, but only 2 to 15 percent of them can be annotated via reference libraries [11]. The low rate of annotation underscores the need to have a better spectral database, ionisation technique and computational aids.

Nuclear Magnetic Resonance (NMR) Spectroscopy.

With NMR spectroscopy, structural data are obtained, as well as absolute quantification is given without derivatization. Though less sensitive than MS based techniques and it is necessary to use large sample masses, NMR is good at filling physical properties of metabolites, and ligand protein and protein binding sites [12]. As a result NMR is commonly employed to verify structures of compounds discovered using MS or to study large quantities of metabolites in plant tissues (sugars and organic acids). NMR NMR has also proven to be an important technique in the study of the interactions of plant pathogens and the nutrient composition of plants through advanced NMR. An example of this is combined spectroscopic analyses of polyphenols, lignin and flavonoids which were related to resistance to southern corn leaf blight[13].

Mass Spectrometry Capillary Electrophoresis (CEMS) and Other Methods.

CEMS has great resolution in the separation of charged, neutral, polar and hydrophobic metabolites and is also sensitive in the detection of small ionic metabolites [14]. It is used together with LCMS to isolate highly polar compounds, including organic acids, amino acids and nucleotides, which have low solubility in LC. Spatially resolved metabolite maps are obtained using mass spectrometry imaging, and matrix assisted laser desorption ionisation, as a result of which it is possible to investigate specific metabolic differences in tissues, and also to investigate plant microbe interactions [15]. Field asymmetric waveform ion mobility spectrometry with MS has the capability to selectively detect volatile metabolites and enhance ion performance. These new techniques as single cell metabolomics employ new ionisation approaches to profile the metabolites at cellular scale, which offers insights into cell type specific metabolism and possible heterogeneity across tissues [16].

In Plant Metabolomics Data Acquisition and Processing.

Metabolite profiling starts at the phase of collecting the samples with utmost care, quenching of metabolic activity and extraction. GCMS involves chemical derivatization to volatilise polar compounds and LCMS involves the use of solvent mixtures to extract a diverse range of metabolites. In targeted metabolomics, internal standards are typically used to eliminate variability as well as quantify metabolites correctly[17].

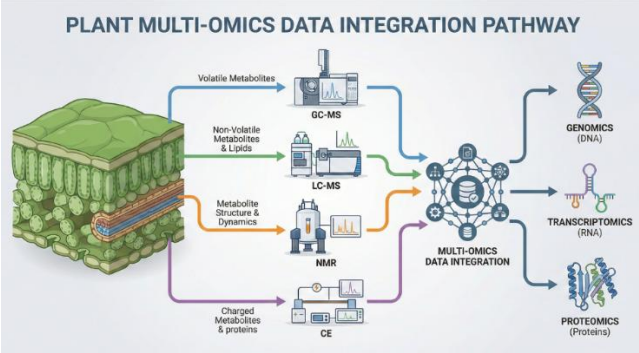

Figure 1: The integrated workflow of plant metabolomics.

Multi Omics Data Integration.

The metabolomics data are becoming more and more used with genomics, transcriptomics and proteomics in order to relate the metabolite variation with the genetic variation. Quantitative trait loci mapping or genome wide association studies may be used together with analytical platforms to determine candidate genes of the metabolite traits [1]. Multi omics models bridge the connection between genes and transcripts, proteins and metabolites. Omics Combinations of both metabolomics and transcriptomics have revealed genes and metabolic pathways involved in resistance to disease and tolerance to stress, and have shown the utility of combined omics in breaking down complex traits [19].

Plant Genetics and Crop Improvement Integration.

Abiotic and Biotic Stress metabolomics Studies.

One of the primary breeding patterns is stress tolerance. Metabolomics has been extensively used in the description of plant response to drought, salinity, heat and nutrient deficiencies [20]. Drought tolerant lines of wheat contain high amino acids and sugars, and salt tolerant rice contains proline, sucrose as well as organic acids. Maize heat stress metabolic profiling revealed increased carbohydrates and sugar alcohols. Biochemical markers that can be applied to test germplasm in their stress tolerance include such metabolite signatures [21].

Metabolomics has been used to study biotic stress to determine defence metabolites and pathways. The use of analytical techniques has demonstrated that pathogen-infested plants generate various metabolites which are polyphenols, lignin and flavonoids which are linked to disease resistance.

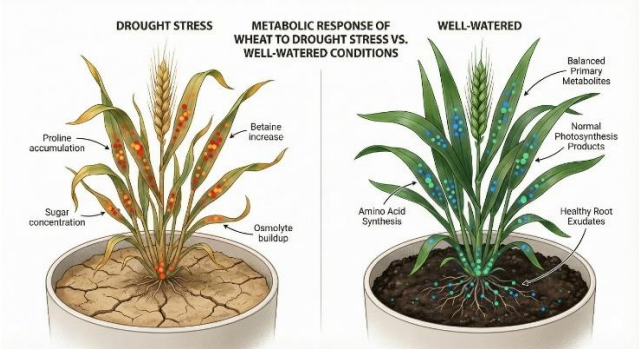

Figure 2: Conceptual visualization of distinct metabolic profiles in stressed versus non-stressed plants.

Biomarker Discovery and Metabolite Assisted Breeding.

Metabolite profiling is a fast method of phenotyping in breeding programs. Combining metabolomics with association mapping has been used to provide genomic links between kernel composition, starch content and provitamin A content in crop. Metabolites have been utilized in predicting heterosis and also in taking advantage of metabolic diversity within breeding populations [22].

The use of metabolomics also supplements the use of marker assisted selection and genomic selection. Metabolite information can also be used to enhance the accuracy of genomic prediction models when combined with genomic prediction models since the metabolites are in a closer relationship with the phenotype than the DNA markers. Metabolite biomarkers can be used in early generation selection where the breeder selects high-quality nutritionally or stress responsive genotypes by screening large populations [23].

New Technologies and Data Science Methods.

Single Cell Metabolomics and Techniques with Spatial Resolution.

Conventional metabolomics studies bulk tissue extracts and therefore make cell types to average metabolite signals. Single cell metabolomics is capable of identifying metabolites in single cells unlabelled, and also visualizing local distributions of metabolites [24].

Metabolic Flux Analysis

Isotopic labeling combined with analytical platforms is employed in the metabolic flux analysis to quantify the rate of metabolic reactions and build metabolic networks. Metabolic flux analysis in crop improvement can be useful to determine bottlenecks in crop improvement and to lead efforts to improve the efficiency of resource utilization.

Problems and Future developments.

In spite of such massive progress, metabolomics in plant breeding has multiple challenges that hamper its regular use. The requirement of derivatizing of GCMS and low sensitivity of NMR restricts their use to particular classes of compounds [27]. Untargeted LCMS is limited to low levels of annotation and inter-laboratory variability. High throughput metabolomics is also expensive and expensive in terms of skill and expertise, which prevents large-scale adoption [28].

There is an additional challenge of data integration. Multi omics techniques have produced elaborate datasets that demand hi-tech bioinformatics and statistical interventions. The open source software and standardisation of collaborative efforts to workflow development will be adopted more broadly.

Conclusion

High throughput methods provide a complete profiling of the metabolites, whereas advanced data processing and identification free methods overcome the problem of the huge metabolic diversity. Combining metabolomics with genomics, transcriptomics and proteomics, the interrelationship between metabolic phenotypes and genetic loci and identification of candidate genes regulating key traits are achieved. The metabolite profiling offers the breeders with quantitative biomarkers of stress tolerance, nutritional quality and heterosis and speed up the selection and boost the specificity of crop development. New technologies and data based methods provide unparalleled chances to decompose cellular metabolism and combine multidimensional data. Regardless of the fact that the challenges still exist, the interdisciplinary approaches will make metabolomics a must-have tool in the twenty first century of plant breeding.

References

[1] Salam, U., et al., Plant Metabolomics: An Overview of the Role of Primary and Secondary Metabolites against Different Environmental Stress Factors. Life, 2023. 13(3): p. 706.

[2] Khan, N., Exploring Plant Resilience Through Secondary Metabolite Profiling: Advances in Stress Response and Crop Improvement. Plant, Cell & Environment, 2025. 48(7): p. 4823-4837.

[3] Arbona, V., et al., Metabolomics as a Tool to Investigate Abiotic Stress Tolerance in Plants. International Journal of Molecular Sciences, 2013. 14(3): p. 4885-4911.

[4] Lelli, V., A. Belardo, and A. Maria Timperio, From Targeted Quantification to Untargeted Metabolomics. Metabolomics - Methodology and Applications in Medical Sciences and Life Sciences, 2021.

[5] Akiyama, M., Multi-omics study for interpretation of genome-wide association study. Journal of Human Genetics, 2020. 66(1): p. 3-10.

[6] Yousri, N.A., O.M.E. Albagha, and S.C. Hunt, Integrated epigenome, whole genome sequence and metabolome analyses identify novel multi-omics pathways in type 2 diabetes: a Middle Eastern study. BMC Medicine, 2023. 21(1).

[7] C. Moldoveanu, S. and V. David, Derivatization Methods in GC and GC/MS. Gas Chromatography - Derivatization, Sample Preparation, Application, 2019.

[8] Kumar, M., et al., Metabolomics and Molecular Approaches Reveal Drought Stress Tolerance in Plants. International Journal of Molecular Sciences, 2021. 22(17): p. 9108.

[9] Swarcewicz, B., et al., Effect of drought stress on metabolite contents in barley recombinant inbred line population revealed by untargeted GC–MS profiling. Acta Physiologiae Plantarum, 2017. 39(8).

[10] Sarkar, J., R. Singh, and S. Chandel, Understanding LC/MS-Based Metabolomics: A Detailed Reference for Natural Product Analysis. PROTEOMICS – Clinical Applications, 2024. 19(1).

[11] Bianchi, F., et al., MS-Based Analytical Techniques: Advances in Spray-Based Methods and EI-LC-MS Applications. Journal of Analytical Methods in Chemistry, 2018. 2018: p. 1-24.

[12] Fei, Q., et al., Combining NMR and MS with Chemical Derivatization for Absolute Quantification with Reduced Matrix Effects. Analytical Chemistry, 2019. 91(6): p. 4055-4062.

[13] Cai, T., et al., HPLC-MRM relative quantification analysis of fatty acids based on a novel derivatization strategy. The Analyst, 2014. 139(23): p. 6154-6159.

[14] Song, X.-C., et al., Application of Ion Mobility Spectrometry and the Derived Collision Cross Section in the Analysis of Environmental Organic Micropollutants. Environmental Science & Technology, 2023. 57(51): p. 21485-21502.

[15] Musatadi, M., et al., Electromembrane Extraction in Suspect Screening of Polar Organic Xenobiotics and their Metabolites in Human Urine: A New Approach to Enhance Compound Annotation? Analytical Chemistry, 2025. 97(18): p. 9683-9690.

[16] Aly, N.A., et al., Utilizing ion mobility spectrometry-mass spectrometry for the characterization and detection of persistent organic pollutants and their metabolites. Analytical and Bioanalytical Chemistry, 2021. 414(3): p. 1245-1258.

[17] Masson, P., et al., Optimization and Evaluation of Metabolite Extraction Protocols for Untargeted Metabolic Profiling of Liver Samples by UPLC-MS. Analytical Chemistry, 2010. 82(18): p. 7779-7786.

[18] Beniddir, M.A., et al., Advances in decomposing complex metabolite mixtures using substructure- and network-based computational metabolomics approaches. Natural Product Reports, 2021. 38(11): p. 1967-1993.

[19] Rabinowitz, J., et al., Metabolite discovery through global annotation of untargeted metabolomics data. 2021.

[20] Khan, N., et al., Advances in detection of stress tolerance in plants through metabolomics approaches. 2017.

[21] Raza, A., Metabolomics: a systems biology approach for enhancing heat stress tolerance in plants. Plant Cell Reports, 2020. 41(3): p. 741-763.

[22] Alseekh, S., et al., Crop metabolomics: from diagnostics to assisted breeding. Metabolomics, 2018. 14(11).

[23] Fontanesi, L., Metabolomics and livestock genomics: Insights into a phenotyping frontier and its applications in animal breeding. Animal Frontiers, 2016. 6(1): p. 73-79.

[24] Ali, A., et al., Single-cell metabolomics by mass spectrometry: Advances, challenges, and future applications. TrAC Trends in Analytical Chemistry, 2019. 120: p. 115436.
[25] Galal, A., M. Talal, and A. Moustafa, Applications of machine learning in metabolomics: Disease modeling and classification. Frontiers in Genetics, 2022. 13.
[26] Liebal, U.W., et al., Machine Learning Applications for Mass Spectrometry-Based Metabolomics. Metabolites, 2020. 10(6): p. 243.
[27] Beale, D.J., et al., Review of recent developments in GC–MS approaches to metabolomics-based research. Metabolomics, 2018. 14(11).
[28] Salem, M.A., et al., Metabolomics in the Context of Plant Natural Products Research: From Sample Preparation to Metabolite Analysis. Metabolites, 2020. 10(1): p. 37.
[29] Nian, M., X. Chen, and M. Fang, Addressing Pitfalls of Metabolomics for Toxicology: A Call for Standardization, Reproducibility and Data Sharing. Chemical Research in Toxicology, 2025. 38(7): p. 1150-1156.