

## Multi-Omics Approaches in the Study of Plants

Tiziana Maria Sirangelo\*

Department of Life Sciences, University of Modena and Reggio Emilia, Modena, Italy

\*Corresponding Author: Tiziana Maria Sirangelo, Department of Life Sciences, University of Modena and Reggio Emilia, Modena, Italy

**Abstract:** Technological advances in next-generation DNA sequencing have made using and collecting multi-omics data increasingly accessible. When compared to single omics studies, multi-omics approaches offer the opportunity to achieve a deeper understanding of relevant plants hot topics, like the ones involving growth, adaptation, development, and disease progression. In this review the importance of adopting an integrated multi-omics approach is discussed and some important plant phenomena are treated. More in detail, plant stress response, diseases and senescence are described, focusing on genomics, transcriptomics, proteomics, metabolomics-based research methods. Furthermore, software tools and bioinformatic resources devoted to support integrated multi-omics approaches are briefly examined and some of their limitations are highlighted. At the same time, open issues are outlined.

**Keywords:** 'omics, multi-omics approach, plants study, genomics, transcriptomics, metabolomics, proteomics

### 1. INTRODUCTION

Over the past decade, technological advances in next-generation DNA sequencing [1] have made using and collecting multi-omics data increasingly accessible, significantly reducing the cost of such activities. Genomics and transcriptomics have been more extensively used as well, and other omics technologies, such as proteomics and metabolomics, are now often incorporated into standard research methodologies. Overall, the adoption of multiple omics techniques is becoming a common practice in many research fields of life science. Consequently, many researchers are now developing comprehensive multi-omics experimental approaches and are trying to integrate the resulting different datasets to obtain a further understanding of the investigated levels [2].

When compared to traditional single omics studies, multi-omics ones offer the opportunity to achieve a further understanding of cells, organisms, and microbial communities, focusing on mechanisms involved in the growth, adaptation, development, and disease progression. Several studies are already available for humans [3] and plants [4].

Moreover, a proper integration of multi-omics data makes it possible to investigate biological pathways in a more comprehensive way. For instance, it is possible to understand how a given genotype influences the phenotype, as well as to individuate the molecular mediators that regulate the underlying mechanisms. It is also possible to identify key biological evidences in pathways that otherwise would not be as clear by considering single-omics approaches alone [5] (Figure 1).

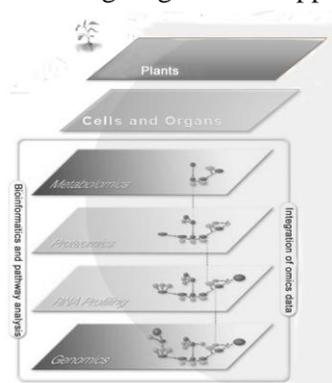


Figure1. Integrated multi-omics approach scheme

## 2. RELEVANCE OF MULTI-OMICS APPROACH IN PLANTS

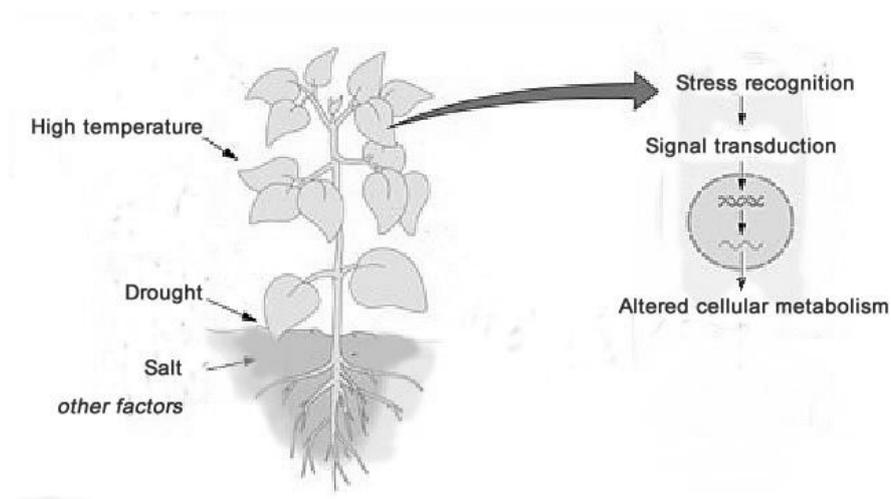
Nowadays, the available technologies make it possible to study the transcripts (mRNA, directly produced from the genes expressed at that time), the proteins and the metabolites present in a given cell, tissue and organ at any moment. These are strictly dependent on one another and are strongly affected by environmental conditions.

A wise integration of gene expression, protein profile and metabolites datasets makes it possible to understand the plant genetic potential and the possible role of each of these elements in response to unfavourable conditions.

In the next paragraph, we discuss the importance of this integrated approach in some cases, such as plant stress response, diseases, senescence.

### 2.1. Plants Stress Response

Abnormal climatic or soil conditions (shortage or excess of water, mineral nutrients, excess of salt) are referred to as abiotic stress factors (Figure 2). They can be responsible for changes in cellular metabolism that generally affect plant growth. Studying the molecular response of plants to abiotic stress is quite complicated, in fact, the response itself varies according to the structural complexity of the species involved, the type of stress and its intensity and duration [6].



**Figure2.** *The main factors responsible for abiotic plant stress*

The transcriptome study aims to monitor space-time variations of gene expression in plants under given conditions. Transcriptomics has been successfully used to identify gene interactions playing a key role in tolerance/susceptibility to stress, highlighting relevant findings by grouping genes with similar expression profiles [7].

The recent advances have not only allowed to understand the general biochemical state of plants, but also to identify proteins involved more specifically in the stress response [8].

Another important aspect to underline is that the metabolic and physiological responses are partly the consequence of subtle modulations of the activity of proteins which were already present and not synthesised in response to the stimuli. Such post-translational modifications are hard to study and only identifiable with proteomic studies [9].

This approach allowed to highlight some of the responses induced by the lack of water and high concentrations of salt on the root system, which plays a central role in plant abiotic stress, in different species. In detail, these studies revealed how levels of proteins involved in energy metabolism increase significantly already starting from the day following plant stress exposition [10].

In this context, it is clear that the integration of results obtained at transcriptomic and proteomic level allows a better understanding of the articulated responses consequent to exposure to stress conditions.

Metabolomics, similarly to other omic approaches, has a great importance in the study of plant stress response. It has been estimated that there could be two hundred thousand to one million metabolites in the plant kingdom alone, while at the single species level there could be a few thousands.

Metabolome analysis is relevant because it allows to verify the real effects caused by the variations observed at the transcriptomic and proteomic level. It also allows to obtain new data about molecules showing specific biological activities, including osmoregulation, oxidative stress response and/or compounds involved in obtaining information on the levels of plant hormones, such as ABA (abscisic acid), which play a central role in modulating stress responses [11]. For instance, in *Arabidopsis thaliana* leaves, investigations highlighted how low water and high salt levels cause the accumulation of many metabolites, including amino acids such as proline, gamma-aminobutyric acid and Krebs cycle acids [12].

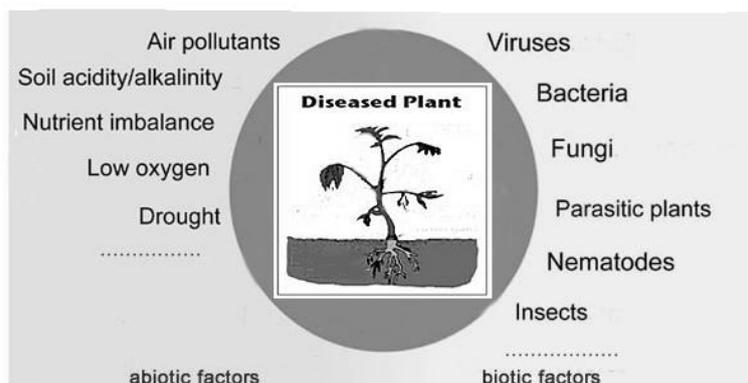
Given that, approaches based on the study of single genes, proteins or metabolites provide interesting but limited results.

A further understanding of the stress molecular response is possible by adopting an integrated multi-omics approach allowing not only to gather information at different molecular levels, but also to highlight the existing relationships among the processes that regulate cellular functions and from which plants adaptation mechanisms strictly depend.

An interesting research project [13] is about the development of sustainable viticulture models, today strongly limited by the scarce availability of rootstocks resistant to adverse cultivation conditions. In this work, the sequencing of some rootstocks, followed by transcriptomic, proteomic and metabolomic analyses have been carried out. The high level of integration of the results made it possible to obtain further information on the biological processes from which the vegetative and productive responses in different environmental conditions.

### 2.2. Plants Disease

Over the past decades, several studies aiming to analyze the interactions between plant immune response and pathogens were carried out. These pathogens include fungal, bacterial and viral agents, and may severely damage crop productivity both in terms of quantity and quality. Abiotic factors are a common cause of disease in plants as well (Figure 3), but in this work the focus is on the biotic ones.



**Figure3.** Some biotic and abiotic factors possibly causing diseases in plants

A very complicated cascade of defence responses is induced during plant-pathogen interactions: signals from these microorganisms are detected by plants immune systems through different mechanisms. Particularly, the molecules produced by pathogens (elicitors) are recognised, resulting into the activation of the basal immune system of plants, which prevents further colonization by incompatible pathogens and limits the spread of compatible pathogens [14] [15] [16]. Consequently, several defence reactions, such as production of reactive oxygen species, strengthening of plant cell wall, and synthesis of specific enzymes, are activated.

The identification of genes involved in such pathways in cultivated germplasm can be used to limit production losses caused by pathogens and the use of phytochemicals in agriculture.

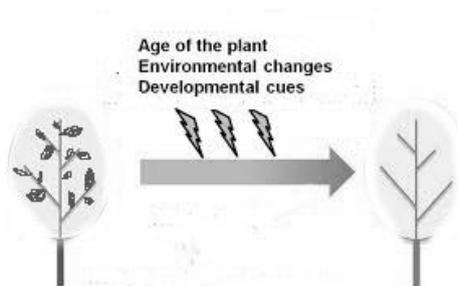
Therefore, one of the greatest priorities in studying plant-pathogen interactions is to identify plant resistance genes. Being the number of genes involved great, high-throughput methods are necessary to identify them and thus to counteract pathogens. The progresses in genomics and proteomics technologies were relevant, allowing the research in this area to reach significant results. Genomics methods identified the detailed organization of resistance gene clusters and the related mechanisms. Investigations also revealed the complex regulatory networks and the proteins involved in resistance [17] [18].

In disease research, it is important to consider that health of plants is strictly linked to their microbiome. This may be defined as the microbial communities of symbionts as well as their genetic material influencing the properties of the interactions within the microbiome itself and with its host. Therefore, a further understanding of the microbiome can reduce the incidence of plant diseases and, consequently, increase agricultural production. Several research studies have recently focused on this aim, and the development of several omics techniques has greatly contributed to this achievement [19].

Finally, in the last few years, the development of machine learning algorithms, which are a collection of analytic methods that automate the process for model generating and iteratively learn from data, provides effective tools to identify genes/proteins involved in plant pathologies [20]. Although machine learning techniques have been applied in different fields, only few studies have been conducted to predict plant pathogens on the basis of these datasets.

### 2.3. Plants Senescence

Studying senescence in plants is difficult because of the impact of several factors (Figure 4).



**Figure4.** Some factors possibly influencing plant senescence

In fact, developmental senescence is mainly determined by the plant genotype, but may be affected by external factors (both biotic and abiotic) as well as by agricultural management services. Depending on individual factors and their possible interactions, senescence can be accelerated or delayed. This affects physiological parameters such as those related to plant growth and development.

While the molecular mechanisms involved in leaf senescence are very well understood, senescence of other organs such as the flower, fruit, and root is less investigated.

Senescence involves major reprogramming of gene expression. The first transcriptome studies have been overcome by high-resolution temporal profiling researches [21] providing a very comprehensive characterization of the leaf transcriptome during *Arabidopsis* life, by considering several investigation levels, like age, RNA type, and organelles [22].

The key factors of cell physiological and biochemical reactions are the structurally and enzymatically active proteins. For this reason, the proteome analysis in senescence research is crucial and contributes to understand the related physiological mechanisms. The proteome studies of leaf senescence have provided new insights and can be considered complementary to the results of the transcriptome researches. Among them, we report studies with advanced proteome methods addressed to leaf senescence in *Arabidopsis* [23].

Senescence is the cause behind the degradation of macromolecules necessary to recycle nutrients sustaining plant metabolism [24]. Several studies report the influence of metabolites in leaf senescence progress [25].

In the light of the above, plant senescence research benefits from several omics areas for which new innovative technologies were available. Therefore, studies based on an integrated approach can be carried out, for instance characterizing phenotypes and addressing plant senescence at different omics levels (metabolomics, proteomics, transcriptomics) [26] [27].

### 3. Software Tools and Databases for Multi-Omics Integration

Several tools for multi-omics data integration are reported in a recent investigation [28]. Numerous resources, databases, software tools, and approaches, all available for free, assisting researchers in integrating multi-omics data, are detailed.

General multi-species resources about genes and proteins, such as GenBank and UniProt [29] [30] are also illustrated.

Similarly, context-specific curated databases and software tools, developed for different biological areas, including selected plant species, are reported. Such databases store data about genome, transcriptome, proteome and metabolome of various species. A relevant example for plants is Plant Metabolic Network or PMN [31].

Among software tools, about botany, are detailed:

- KaPPA-View, which integrates transcriptomics and metabolomics plant data to map pathways [32];
- MapMan, firstly developed to be used for *Arabidopsis* now includes more species. It integrates metagenomics, transcriptomics and metabolomics data, handle KEGG and KOG classification into clusters, and maps expression responses [33];
- VitisNet, about grapes, manages metagenomics, transcriptomics, proteomics and metabolomics data [34];
- MADMAX, a management and analysis database for multiple omics experiments. It integrates metagenomics transcriptomics and metabolomics data, statistical analysis and pathway mapping [35];
- MetaboAnalyst manipulates genomics, transcriptomics, proteomics and metabolomics data, runs data processing and statistical analysis, and develops pathway analysis [36].

Unfortunately several bioinformatics resources have different formats, many of them are not compatible with the most common standards. Moreover, different software packages may require their input data in non-standard formats and produce outputs which are incompatible with others.

Construction and visualization of network models is another important aim of multi-omics integration research investigation. Recently, multi-layer networks have been developed and allow the interpretation of specific interactions between different omics layers [37].

However, all the tools mentioned above still need improvements, like the ones concerning the availability of pathway databases providing links between genes, proteins, and metabolites. The few resources already existing [5] emphasize metabolic pathways, while other kinds of pathways are also important in plant science, like protein and metabolite signaling, gene activation and many others. Some of such pathways characterize existing commercial tools, but the lack of compatibility with many other bioinformatics software packages makes them difficult to integrate into multi-omics open-access pipelines [38].

Tools real ability to uncover multi-omics data relationships to analyse pathway cross-talk is still limited. Thus, new methodological approaches and software tools integrating multi-omics datasets for pathway analysis would be necessary. Machine learning approaches are also required for the model-based inference of multi-omics data for analysis of pathways [39].

#### 4. Conclusion and Discussion

While very large omics datasets are becoming more accessible, and multi-omics research approaches are being used more frequently, real multi-omics integration is an ongoing challenge. In fact, due to data differences from each omics layer, integrating multiple omics platforms is still a difficult task. Many specific analytical tools traditionally used for individual omics area are not well-suited to allow a proper integration among multiple omics disciplines. For instance, methods and storage techniques used for genomics investigations are often unsuitable for metabolomics, proteomics and transcriptomics. In the same way, qualitative methods commonly used in transcriptomics and proteomics are not suitable for the quantitative methods applied in genomics [38].

In the light of the above, high-quality multi-omics studies require: an appropriate experimental design; careful collection of multi-omics data and related meta-data; better tools for data integration; new resources and repositories for such datasets.

Even if problems about omics technologies could be controlled through the availability and the improvement of reference standards and well-defined laboratory protocols, today there are still significant variations in sample extraction and storage among different research laboratories.

REFERENCES

- [1] E.R. Mardis. Next-Generation DNA Sequencing Methods, *Annu. Rev. Genom. Hum. Genet.* vol. 9, pp. 387–402, (2008).
- [2] K. Yizhak, T. Benyamini, W. Liebermeister, E. Ruppin, T. Shlomi. Integrating quantitative proteomics and metabolomics with a genome-scale metabolic network model. *Bioinformatics*, vol. 26, pp. 255–260, (2010).
- [3] K. Cisek, M. Krochmal, J. Klein, H. Mischak. The application of multi-omics and systems biology to identify therapeutic targets in chronic kidney disease, *Nephrol. Dial. Transplant.* vol. 31, pp. 2003–2011, (2016).
- [4] Y. Hasin, M. Seldin and A. Lusis. Multi-omics approaches to disease. *Genome Biology*, vol.18, pp. 83, (2017).
- [5] A. Fabregat, S. Jupe, L. Matthews, K. Sidiropoulos, M. Gillespie, P. Garapati, R. Haw, B. Jassal, F. Korninger, B. May, et al. The Reactome Pathway Knowledgebase, *Nucleic Acids Res.*, vol. 46, pp.649–655, (2018).
- [6] E.A. Tattersall, J. Grimplet, L. Deluc, M.D. Wheatley, D. Vincent, et al. Transcript abundance profiles reveal larger and more complex responses of grapevine to chilling compared to osmotic and salinity stress, *Functional and Integrative Genomics*, n. 7, vol.4, pp. 317-333, (2007).
- [7] J. Kilian, D.Whitehead, J. Horak, D. Wanke, S. Weinl, O. Batistic, et al. The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses, *Plant Journal*, n. 50, vol. 2, pp. 347-363, (2007).
- [8] V.J. Nesatyy, M.J.F. Suter. Proteomics for the Analysis of Environmental Stress Response in Organisms, *Environmental Science & Technology*, n. 41, pp. 6891-6900, (2007).
- [9] C. Kannicht. Post-translational modification of proteins. Tools for functional proteomics (Second Edition), C. Kannicht, Humana Press, Totowa, (2008).
- [10] D. Ghosh, J. Xu. Abiotic stress responses in plant roots: a proteomics perspective, *Frontiers in plant Science*, n. 5, Article 6, (2014).
- [11] V. Arbona, M. Manzi, C. de Ollas, A. Gómez-Cadenas. Metabolomics as a tool to investigate abiotic stress tolerance in plants, *International Journal of Molecular Science*, n. 14, pp. 4885-4911, (2013)
- [12] K. Urano, K. Maruyama, Y. Ogata, Y. Morishita, M. Takeda, N. Sakurai, H. Suzuki , et al. Characterization of the ABA-regulated global responses to dehydration in Arabidopsis by metabolomics, “Plant Journal” n. 57, pp. 1065-1078, (2009).
- [13] M. Corso, B. Prinsi, A. Vannozzi, L. Espen and C. Bonghi. An Integrated Omic Approach for the Understanding of the Stress Physiology. Ager-Serres Project, Strategia innovativa per la selezione di nuovi portainnesti di vite, Publisher: Ed. L'Informatore Agrario, Editors: A.Scienza, O. Failla, L. Espen, pp.99-109, (2014).
- [14] R. Michelmore. Genomic approaches to plant disease resistance, *Current Opinion in Plant Biology*, vol. 3, pp. 125–131, (2000).
- [15] E.E. Schadt, J. Lamb, X. Yang, J. Zhu, S. Edwards, D. GuhaThakurta, et al. An integrative genomics approach to infer causal associations between gene expression and disease, *Nature genetics*, vol. 37, pp. 710-717, (2005).
- [16] P. Bittel and S. Robatzek S.: Microbe-Associated molecular patterns (MAMPs) probe plant immunity, *Curr Opin Plant Bio*, vol. 10, pp. 335-341, (2007).
- [17] F. Martinelli, R.Scalenghe, S. Davino, S. Panno, G. Scuderi, P. Ruisi, P. Villa, D. Stroppiana, M. Boschetti, L. R. Goulart, et al. Advanced methods of plant disease detection. A review, *Agronomy for Sustainable Development*, Springer Verlag/EDP Sciences/INRA, n. 35, vol. 1, pp.1-25, (2015).
- [18] J. Sperschneider, A.M. Catanzariti, K. DeBoer, B. Petre, D. Gardiner, K. Singh, et al. LOCALIZER: subcellular localization prediction of plant and effector proteins in the plant cell. *bioRxiv*. 092726, (2016).
- [19] R. L. Mondéjar, M. Kostovčik, S. Lladó, L. Carro, P. G. Fraile. Exploring the Plant Microbiome Through Multi-omics Approaches, *Probiotics in Agroecosystem*, Springer Nature Singapore, pp 233-268, (2017).
- [20] Y. Xin, G. Tingwei. Machine learning in plant disease research. *European Journal of BioMedical Research*, pp. 6-9, (2017).
- [21] E. Breeze, E. Harrison, S. McHattie, et al. High-resolution temporal profiling of transcripts during Arabidopsis leaf senescence reveals a distinct chronology of processes and regulation, *The Plant Cell*, vol. 23, pp. 873–894, (2011).
- [22] H.R. Woo, H.J. Koo, J. Kim, et al. Programming of plant leaf senescence with temporal and inter-organellar coordination of transcriptome in Arabidopsis. *Plant Physiology*, vol. 171, pp. 452–467, (2016).

- [23] S. Tolin, G. Arrigoni, A.R. Trentin, S. Veljovic-Jovanovic, M. Pivato, B. Zechman, A. Masi. Biochemical and quantitative proteomics investigations in Arabidopsis ggt1 mutant leaves reveal a role for the gamma-glutamyl cycle in plant's adaptation to environment. *Proteomics*, vol.13, pp. 2031–2045, (2013).
- [24] J.H. Schippers. Transcriptional networks in leaf senescence. *Current Opinion in Plant Biology*, vol. 27, pp. 77–83, (2015).
- [25] A. Wingler, T.L. Delatte, L.E. O'Hara, L.F. Primavesi, D. Jhurrea, M.J. Paul, H. Schluepmann. Trehalose 6-phosphate is required for the onset of leaf senescence associated with high carbon availability. *Plant Physiology*, vol. 158, pp. 1241–1251, (2012).
- [26] D. Houle, D.R. Govindaraju, S. Omholt. Phenomics: the next challenge. *Nature Reviews. Genetics*, vol. 11, pp. 855–866, (2010).
- [27] D.K. Großkinsky, S.J. Syaifullah and T. Roitsch. Integration of multi-omics techniques and physiological phenotyping within a holistic phenomics approach to study senescence in model and crop plants, *Journal of Experimental Botany*, vol. 69, n. 4, pp. 825–844, (2018).
- [28] B.M. Biswapriya, L. Carl, O. Michael, A.C. Laura. Integrated omics: Tools, advances and future approaches, *J. Mol. Endocrinol.*, vol. 62, pp. 21–45, (2019).
- [29] D.A. Benson, M. Cavanaugh, K. Clark, I. Karsch-Mizrachi, D.J. Lipman, J. Ostell, E.W. Sayers, GenBank. *Nucleic Acids Res.*, vol. 41, pp. 36–42, (2013).
- [30] UniProt Consortium. UniProt: The universal protein knowledgebase. *Nucleic Acids Res.* Vol. 46, pp. 2699, (2018).
- [31] P. Schläpfer, P. Zhang, C. Wang, T. Kim, M. Banf, L. Chae, K. Dreher, A.K. Chavali, R. Nilo-Poyanco, T. Bernard, et al. Genome-wide prediction of metabolic enzymes, pathways, and gene clusters in plants. *Plant Physiol.*, vol. 173, pp. 204, (2017).
- [32] T. Tokimatsu, N. Sakurai, H. Suzuki, H. Ohta, K. Nishitani, T. Koyama, T. Umezawa, et al. KaPPA-view: A web-based analysis tool for integration of transcript and metabolite data on plant metabolic pathway maps. *Plant Physiol.*, vol.138, pp. 1289–1300, (2005).
- [33] B. Usadel, A. Nagel, O. Thimm, H. Redestig, O.E. Blaesing, N. Palacios-Rojas, J. Selbig, J. Hannemann, M.C. Piques, D. Steinhauser, et al. Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of corresponding genes, and comparison with known responses. *Plant Physiol.*, vol.138, pp. 1195–1204, (2005).
- [34] J. Grimplet, G.R. Cramer, J.A. Dickerson, K. Mathiason, J. Van Hemert, A.Y. Fennell, VitisNet: Omics integration through grapevine molecular networks, *PLoS ONE*, vol. 4, e8365, (2009).
- [35] K. Lin, H. Kools, P.J. de Groot, A.K. Gavai, R.K. Basnet, F. Cheng, J. Wu, X. Wang, A. Lommen, G.J. Hooiveld, et al. MADMAX - Management and analysis database for multiple omics experiments, *J. Integr. Bioinform.*, vol. 8, pp.160, (2011).
- [36] J. Chong, O. Soufan, I. Caraus, J. Xia, C. Li, D.S. Wishart, G. Bourque, S. Li. MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis, *Nucleic Acids Res.* vol. 46, pp. 486–494, (2018).
- [37] M. Kivelä, A. Arenas, M. Barthelemy, J.P. Gleeson, Y. Moreno, Porter, M.A. Multilayer networks, *J. Complex Netw.*, vol. 2, pp. 203–271, (2014).
- [38] F.R. Pinu, D.J. Beale, A.M. Paten, K. Kouremenos, S. Swarup, H. J. Schirra and D. Wishart. Systems Biology and Multi-Omics Integration: Viewpoints from the Metabolomics Research Community, *Metabolites*, vol. 9, pp. 76, (2019).
- [39] K. Chaudhary, O.B. Poirion, L. Lu, L.X. Garmire. Deep Learning based multi-omics integration robustly predicts survival in liver cancer, *Clinical Cancer Research*, vol. 24, Issue 6, (2018).

**Citation:** Tiziana Maria Sirangelo, "Multi-Omics Approaches in the Study of Plants". *International Journal of Advanced Research in Botany (IJARB)*, vol. 5, no. 3, pp. 1-7, 2019. DOI: <http://dx.doi.org/10.20431/2455-4316.0503001>.

**Copyright:** © 2019 Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.