Tutor/s

Dr. José Fermín López Sánchez
Departament de Química Analítica

Dr. Joaquim Jaumot Soler
CSIC-IDAEA

Grau de Química

Treball Final de Grau

Metabolomic study of physical stressors on rice crops

Estudi metabolòmic dels efectes d'agents estressants sobre cultius d'arròs

Roger Bujaldón Carbó

June 2015

Universitat de Barcelona

Campus d'Excel·lència Internacional
B:KC Barcelona Knowledge Campus

In order to succeed, your desire for success should be greater than your fear of failure.

Bill Cosby

En primer lloc, agrair als meus dos tutors: al Dr Joaquim Jaumot, per la seva dedicació i per tot el que he après, i al Dr. José Fermín López per haver-me guiat quan ho he necessitat durant el transcurs del treball.

També agrair als companys del laboratori per haver-me fet sentir un més del grup durant el temps que hi he estat, i en especial a la Meritxell, per haver-me ensenyat l'art de cuidar l'arròs i tantes altres coses, i també per la paciència i dedicació que ha tingut.

Als amics i companys, que han contribuït amb tants bons moments al llarg de la carrera i sempre hi han sigut per alleujar els més difícils.

Finalment, a la meva família pel suport incondicional que sempre he tingut i per què sempre m'han animat a continuar endavant. *17*

REPORT

CONTENTS

1. SUMMARY

Metabolomics is defined as the study of biochemical processes that involve small molecules, namely metabolites. Metabolomics has an important role on studies aimed at the understanding of living organisms at molecular level. Since metabolites are involved in most of biological mechanisms, their quantification and identification serve as an approximation of cellular state. To achieve so, it analyzes the metabolites of organisms submitted to any perturbation and compares the profiles to non-perturbed organisms. Therefore, it can be found how this perturbation affects the organism's metabolome.

In this work a lipidomic study, which only involves the hydrophobic metabolites, will be performed on rice. It aims to evaluate the effect of water scarcity as a physical stressor and influence of the harvesting hour. The study includes lipid extraction, analysis using LC-MS and data processing by chemometric means. Additionally, an optimization of the extracting method is included, studying three different factors.

Keywords: Metabolomics, rice, physical stressors, chemometrics, lipidomics, lipid extraction, LC-MS, drought

2. RESUM

La metabolòmica es defineix com l'estudi dels processos bioquímics que involucren molècules petites, és a dir, metabòlits. La metabolòmica té un paper important en estudis destinats al coneixement d'organismes vius a nivell molecular. Com que els metabòlits estan involucrats en la majoria de mecanismes biològics, la seva quantificació i identificació serveixen com a lectura de l'estat cel·lular. Per aconseguir-ho, s'analitzen els metabòlits d'organismes sotmesos a una pertorbació per tal de comparar-ne el perfil amb el d'organismes als quals no se'ls ha aplicat. Per tant, es pot descobrir com afecta la pertorbació al metaboloma de l'organisme.

En aquest treball, es durà a terme un estudi lipidòmic, el qual està centrat en els metabòlits hidrofòbics, en plantes d'arròs. Es pretén veure com l'afecta la manca d'aigua com a agent estressant i la influència de l'hora de la collita. L'estudi inclou l'extracció de lípids, l'anàlisi mitjançant LC-MS i processament de dades amb mètodes quimiomètrics. A més, l'estudi inclou l'optimització del mètode d'extracció, que estudia tres factors diferents.

Paraules clau: Metabolòmica, arròs, agent estressant, quimiometria, lipidòmica, extracció de lípids, LC-MS, sequera

3. INTRODUCTION

The understanding of living organisms at molecular level is a major challenge in biological studies, which are basically aimed at the comprehension of all biochemical mechanisms involved. The full comprehension of all processes, though, not only implies the functioning related to the organism itself, but its response to any perturbation. Actually, the responses observed to external changes serve as an important base to perform this kind of studies.

3.1. STAGES OF MOLECULAR LEVEL STUDIES

The biological studies at molecular level are classified in different stages, concerning the degree of complexity of the targeted molecular units. They include genes, the different forms of RNA, proteins and metabolites [1].

3.1.1. The "omics" cascade

The study related on each stage or group of molecular units is commonly known as "omics". Thus, the study of the different levels are genomics, transcriptomics, proteomics and metabolomics, respectively, and their sequence is called the "omics" cascade [1], as represented on figure 1.

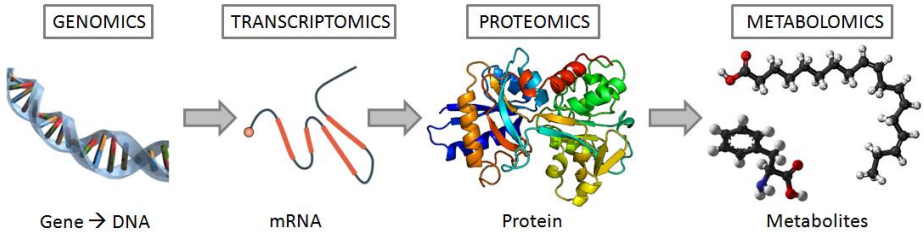

Figure 1. The "omics" cascade sequence.

A complete dataset of all "omics", and specially the integration between them, could provide a comprehensible description of biological responses of the studied organism to any change or perturbation, whether caused by disease, genetic or environmental origin. This could represent

not only the biological understanding of the organism, but also the possibility to predict its state by analyzing just the involved components [1,2].

In spite of the importance of the information provided by all "omics", their interest regarding the effects attached to perturbations is not equal. Whereas genomics aims the analysis of the complete genome to understand the function of single genes, the most functional information is obtained from the other "omics": transcriptomics, which studies the genes expression, proteomics and metabolomics [2].

Nevertheless, it is considered that metabolomics is the endpoint of the "omics" cascade, because the changes it suffers are the ultimate answer of an organism to any influence. Hence, the metabolome is the closest to phenotype, and the most accurate readout of cellular state in an organism [1,2].

3.1.2. Metabolomics and lipidomics

Metabolomics can be defined as the study aimed at the identification and quantification of all metabolites in a biological system or organism, whereas the whole set of metabolites present in an organism is called metabolome [3]. Metabolites are molecules that somehow participate into the chemical reactions involved in the metabolic processes of a cell, which are mainly required for its maintenance, growth and normal function. Metabolites are subject to continued chemical transformations, so they provide valuable information related with the cell state and phenotype. The term is generally applied to small molecules [1,3].

The identification and quantification of the whole metabolome appears to be a difficult goal, though. And it certainly is, mainly due to the enormous amount of metabolites present in an organism, and their different natures. In fact, the amount of metabolites concerning an organism can be very large. It increases even more when considering secondary metabolites, which are not directly involved in the growth, development or reproduction but can be of great importance in other processes. For example, over 200,000 metabolites are likely to be encountered in the plant kingdom [4].

Another fact to be considered is that metabolites differ greatly in structure, chemical properties and concentration. Consequently, these differences are going to difficult a complete analysis. For example, the variation in concentration is likely to go from major to trace level

metabolites, and can vary spatially and temporally [2,3]. The different polarities they have are an important point too, especially on the extraction step.

Even though the term metabolite includes from the most to the least hydrophilic compounds, it is often applied just to refer to the hydrophilic metabolites (such as amino acids or carbohydrates), while the hydrophobic metabolites are referred as lipids as a whole [5]. Consequently, the study focused on the analysis of lipids alone is called lipidomics. For the difficulties implied on studying metabolites of different polarity together, metabolomics and lipidomics are performed separately.

3.2. STRUCTURE OF A METABOLOMIC STUDY

A metabolomic study requires the acquisition of samples of the chosen organism. In case of a perturbation-induced study, there are needed at least two types of samples: control samples and some groups of perturbed samples, generally in different degrees. Control samples are organisms in which no perturbation is applied, and serve to compare with the perturbed samples in order to discriminate the potential biomarkers.

The organism studied is not chosen haphazardly, but it is likely to be a model organism. An organism considered model can be correlated with other organisms belonging to the same family but is sometimes easier to study. Thus, the amount of information that can be collected is considerably higher and useful.

3.2.1. Rice as a plant model

Rice is a member of the family *Poaceae* or *Gramineae*, commonly known as the grass family, and included in the genus *Oryza*. The cereal grain it produces is well known for being the most widely consumed staple food worldwide, and the third one more produced. There are several known species (over 300) both wild and domesticated, but rice production basically lies into the species *Oryza sativa* (Asian rice) and *Oryza glaberrima* (African rice), although the first one is the most widespread [6].

Oryza sativa (figure 2) contains two major subspecies. The *Indica* variety is characterized by growing almost submerged in flooded fields and producing non-sticky, long grains, while *Japonica* variety is usually cultivated in dry fields and its grains are sticky and short.

Figure 2. Different parts of an Oryza sativa L. specimen.
(Franz Eugen Köhler, 23/4/15 via Wikipedia, public domain)

Aside from its importance as crop, rice is considered to be a plant model in genetic studies. A plant model provides information that can be related to other plants, especially on those of the same family. The reasons that make *Oryza sativa* a perfect candidate for being a plant model are basically the easiness to modify genetically and its relatively short period of growth. Thus, the time involved in growing genetically modified plants is considerably reduced, and the amount of information that can be obtained is wider than using other plants.

Besides, the knowledge concerning such an important crop makes rice a perfect candidate to be studied. The effect of the environmental conditions in which the plant grow can alter the plant, if they differ from the considered optimal. A metabolomic study enables to detect how the plant is affected.

In this work, it will be investigated the effect of drought on *Oryza sativa L. Japonica*. In fact, water scarcity is the most problematic stressor on crops present in many regions of the world [7]. The study also has a temporal component. Metabolism in plants, as in almost every organism is rolled by the circadian rhythm, which adapt the different biological functions to the day period. There are many "omic" studies focused on the circadian clock and the integration to all the influenced pathways [8].

3.2.2. Metabolomic approaches

As explained before, metabolomics requires the analysis of all metabolites (or lipids in lipidomics), but the quantification of the whole set of metabolites in an organism simultaneously is not an easy task. There exist two approaches that seem as useful as a whole metabolome quantification, but easier to carry out, which are metabolite profiling and metabolite fingerprinting [1-3].

3.2.2.1. Metabolic profiling

Unlike a metabolomic study in its true sense, the aim of a metabolic profiling is the quantitative analysis of a selected group of metabolites. This approach is often called targeted metabolomics, because it is only focused on a few metabolites of interest discarding the rest. The selected metabolites may have in common a similar structure or be part of the same class of compounds, which facilitate a simultaneous analysis. In other cases, it can be interesting the analysis of some metabolites related with a specific metabolic pathway. The analyzed metabolites are therefore identified and quantified. Generally, it is done in presence of some variation or perturbation to correlate with their roles in the metabolic pathway. This approach is most useful for obtaining information about a specific metabolic pathway or linking it with other pathways, and thus creating a pathway map [1,3].

3.2.2.2. Metabolic fingerprinting

The metabolic fingerprinting or untargeted approach is closer to global metabolomics, because it aims to analyze as many metabolites as possible. However, the identification of the analyzed compounds is not the initial goal. The main purpose is generally based on evaluating the variations produced in the metabolome when changing the system. This is achieved by studying natural fluctuations on the concentration of metabolites or applying a perturbation, whether directly in the organism, like in a genetic alteration, or indirectly in its environment. The corresponding results are then compared, to find the common pattern and the distinct features between the normal and perturbed metabolome. The distinct features show the group of metabolites that actually change, and thus give information about the presence or absence of this specific perturbation on the organism. Those metabolites are identified as biomarkers. It also enables to improve biological understanding by correlating biomarkers with the effect they produce. Whereas a metabolic profiling study is in most cases derived from a hypothesis, metabolic fingerprinting is more a hypothesis-generating approach [1,3].

3.2.3. Extraction step

Depending on the kind of sample, a previous extraction of the desired metabolites is required. In order to perform the extraction, the sample is treated in an adequate solvent or mixture of solvents. It is in this step that lipids and metabolites are separated. Using a more polar solvent or mixture enables the extraction of metabolites, while lipids are extracted with a less polar solvent. Even so, a little amount of the non-desired class of compounds is going to be extracted. To discriminate the non-desired compounds, a liquid-liquid extraction is generally performed afterwards. This separation of lipids and metabolites in the study is necessary for the following step, since the complete analysis would be too complex and time consuming [3,9].

This study is focused on lipidomics, so the extracting solvent or mixture has to be rather hydrophobic.

3.2.3.1. Extraction of lipids from plants

Any extraction involved in an analysis procedure requires to be optimized before the lipidomic study itself, so an appropriated quantification can be done.

This study is based on untargeted lipidomics on rice, so it aims to extract as many lipids as possible on this specific organism. Therefore, the extracting method has to be quite versatile. Several methods are described to accomplish this goal, which seem suitable to be used on plants. The most popular methods basically use a mixture of chloroform and methanol as the extracting solvent, in different proportions. This mixture seems to be able to extract from the most to the least hydrophobic lipids. Those methods can be summarized into two, regarding the proportions of the extracting mixture, the Bligh-Dyer [9,10] and Folch [9].

An alternative method substitutes chloroform by a less hazardous solvent, MTBE, in a mixture with methanol [11,12].

Besides, in some works has been described a step previous to the extraction. It turns out that some plants have very active phospholipase D, which participates in the hydrolyzation of phospholipids to form phosphatidic acid and free fatty acids. The phospholipase seems to be somehow active especially in presence of chloroform. As a result, the lipid concentration may vary from the original one, distorting the results obtained. Thetreatment is based on the addition of hot isopropanol to the sample, to achieve the phospholipase inhibition. Therefore, the changes in phospholipids' concentration are minimized [13].

In order to find out which of these methods is most suitable in rice untargeted lipidomics, a testing on the different methods is included as a preliminary part of this work. The testing will evaluate the effect of three factors: the extraction method, the treatment with isopropanol and an additional one regarding the amount of sample used.

3.2.4. Instrumental techniques

Another important step on metabolomics is the analysis of the extracted metabolites or lipids in each sample. As said before, it is a complex task due to the amount of compounds that are to be analyzed. Currently, two different instrumental techniques are basically used in metabolomics and lipidomics: NMR and MS spectrometry [3,9].

NMR spectroscopy is a technique that exploits the magnetic properties of certain isotopes to their detection, identification and quantification. It provides spectrums that profile the samples metabolome. It is noted in metabolomics because it requires minimal sample preparation and is non-destructive. However, it is unable to detect metabolites in low concentration [3].

Mass spectrometry (MS) is a technique widely used in analytical chemistry, which enables to identify the type, structure and concentration of any compound present in the analyzed sample. The sample is submitted to an ionization source that ionizes the molecules, and depending on the source used, fragments them to different degrees. Those charged molecules or fragments are analyzed according to their mass to charge ratio. This enables their identification and quantification, performed by measuring the abundance of the ionized molecules or fragments.

In this work, two different MS detectors are going to be utilized. In the preliminary test, it will be used a Triple Quadrupole Detector (TQD) and in the lipidomic study, a Time of Flight (TOF) detector. TOF is an exact mass detector, which is necessary in order to identify the lipids eluted.

Especially in lipidomics, MS is the most utilized technique [5,9]. The direct injection of sample to the MS, commonly known as "shotgun", is a simple way to analyze the metabolites of a sample, which requires little treatment and provides useful information quite easily. However, it has some drawbacks. As all metabolites are analyzed at once, compounds with the same molecular mass, generally chemical isomers, cannot be differentiated. This is an inconvenient, particularly on lipidomics, where isomers are very common. Also, it can produce enhancement or suppression of signals and lead to misinterpretation [3].

In order to avoid those drawbacks, the mass spectrometer can be coupled to a chromatographic technique, typically GC or LC.

LC-MS separate the different metabolites or lipids regarding their different affinity towards the mobile phase and stationary phase, coated inside the column. While metabolomics requires more hydrophilic columns (typically HILIC) and hydrophobic solvents, in lipidomics are used hydrophobic-coating columns such as C-18 and C-8 and hydrophilic solvents [3,9].

GC-MS is less versatile in metabolomics, for only thermally stable and high vapor pressure substances can be analyzed, and sometimes require to be derivatized. It separates the metabolites regarding the affinity towards the stationary phase [3].

3.2.5. Data processing

The data obtained by analyzing the extracted metabolites or lipids contains all the information related to nature and concentration of every compound of the sample. It also enables the comparison of samples submitted to any perturbation in different degrees with those of control and between them, in order to find the corresponding biomarkers. However, the amount of data is exceedingly large to be treated with a normal procedure.

3.2.5.1. Representation of the acquired data

The use of an MS technique coupled with liquid chromatography will provide a matrix containing the retention times in rows and the whole set of mass to charge ratio values in the columns. This matrix is usually represented as a TIC chromatogram. It is obtained by adding the intensity of every compound eluted at a given time of retention, regardless of the mass to charge value. Hence, the initial matrix is simplified into a vector, which represents the intensity in each time.

It is a simple and illustrative way to represent the data, although it is not useful to identify single metabolites. However, it represents the sample profile, which allows to be compared with the other samples.

An alternative used to represent LC-MS data is obtained from the MetaboNexus software. MetaboNexus is an interactive analysis platform focused on the processing of data acquired from metabolomics. It combines pre-processing of raw peak data with statistical analysis and identification of metabolites.

Metabonexus does a data preprocessing, including peak alignment to correct small differences on the elution time due to the instrument, and baseline correction. Then, it provides the mass to charge ratio of every detected substance present on a significant amount of samples, the corresponding retention time and the peak intensity in each sample. Since each mass to charge value corresponds to an eluted metabolite, it is possible to identify some of interest. If any patron has been added to the samples and it is detected, intensity can be normalized by dividing the intensity value in each sample. Thus, any difference in concentration between samples due to the extraction or instrumental measurement can be eliminated. A scheme of data conversion using the Metabonexus software and TIC chromatogram is illustrated in figure 3.

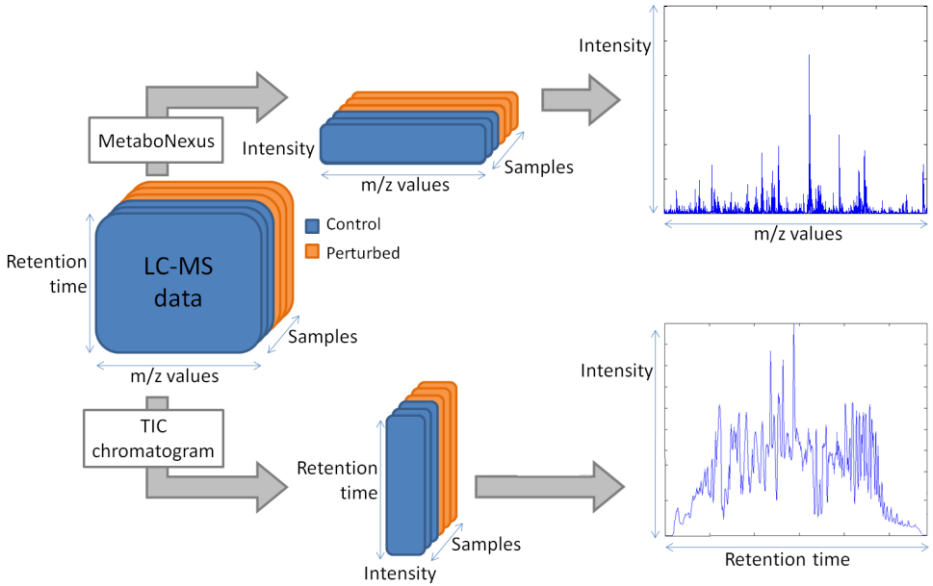

Figure 3. Data conversion and corresponding graphics from TIC and MetaboNexus.

In an untargeted lipidomic study, only the variables that differ when comparing the whole set of samples is going to be of interest [3]. The variables used are the retention times in a TIC chromatogram or the mass to charge ratios in the vectors obtained from MetaboNexus. So the variables susceptible to change when applying a perturbation allow finding informative peaks, and thus can be related with the metabolites affected. The process of discriminating the compounds that are not affected by the perturbation, which are of no interest in that particular study, from those which actually vary, is achieved using chemometrics.

Chemometrics is defined as a discipline that uses mathematics, statistics and logic integrated to chemical studies. It is used to select optimal experimental procedures or to obtain maximum relevant information on chemical data. It is a widely used tool to deal with huge amounts of chemical data in order to extract the aimed information [14].

3.2.5.2. Principal component analysis

Principal component analysis (PCA) is a basic method used in chemometrics. It allows to represent a huge amount of data in a reduced dimension plane, which otherwise would be too complex to deal with. This representation is done by converting the original variables into new ones, called principal components (figure 4). Each PC is orthogonal to the rest, because they are obtained from linear combinations of the original variables. Thus, no information is repeated, and the number of PC needed to explain the relevant variance between samples is much lesser. The amount variance explained by each PC is not equal, though, and are typically ordered from more to less variance described [14,15].

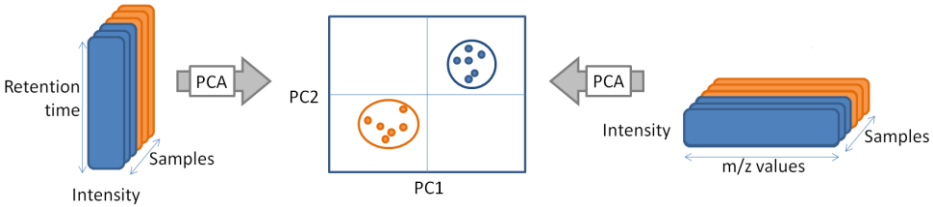

Figure 4. PCA data representation

Another function of PCA is to reveal potential similarities or differences among samples, and to classify them in different groups. It is also possible to discriminate which original variables describe better each group of samples and which ones are not related to anyone. Whereas samples are represented on the scores diagram, variables are represented on the loadings diagram, using the same PCs [14].

Concerning metabolomics, a PCA is likely to separate the samples between controls and perturbed. Besides, each group would be described by some variables, while the rest are not going to be of interest. The variables, or the metabolites involved, that better describe the variation between controls and perturbed samples are feasible biomarkers [3].

3.2.5.3. Partial least squares regression

Partial least squares (PLS) regression is another chemometric method typically used in multivariant calibration. In spite of its apparent resemblance to PCA, they work differently. In order to calculate the components, two matrixes are needed, **X** and **Y**. **X** contains the information, and it is correlated with a property contained in **Y**. The PLS components are obtained from the maximum amount of variance described by the samples that better correlates to **Y**. Thus, in contrast to PCA, which merely consider variance among samples to separate them in groups, PLS only evaluate the variance that can be used to separate samples basing on the desired property. It is very useful to do calibration models, aimed at finding the correlation between two elements even when it is not evident. It can therefore be used to predict **Y** property value of unknown samples [16].

Besides, PLS regression possesses other advantages. Considering the variables, it is possible to determine which ones are more important at correlating **X** to **Y**. There exist several indexes to evaluate so, but the most popular is VIP index (figure 5). It is calculated by adding the square weight at building the PLS model of each original variable, in every component, and pondered by each component variance. If the VIP value is greater than 1, the variable is considered relevant. If not, the variable is not related to that property. Moreover, the greater the VIP value is, the most important is considered [16].

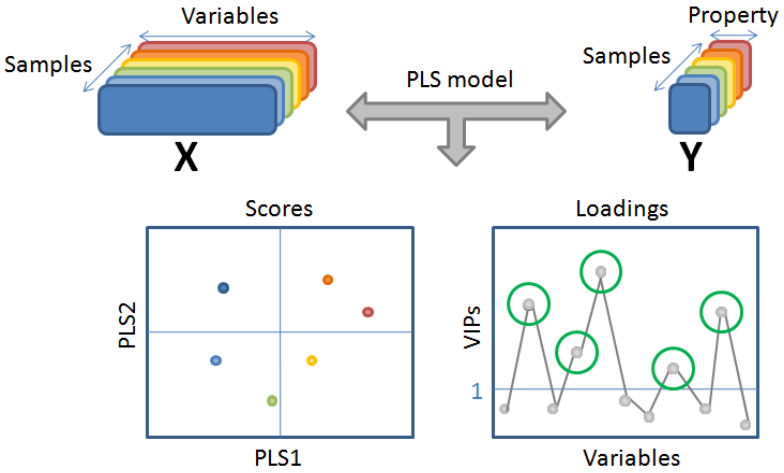

Figure 5. PLS proceeding and obtained diagrams: scores and loadings

Considering metabolomics, PLS can be used to correlate the factor studied to the acquired data, and thus obtaining a good approach on potential biomarkers for that factor. They are featured by their corresponding variables (mass to charge ratio), if their VIP is greater than 1.

In this work, it is studied the effect of drought on rice lipidome through time. Hence, PLS will be useful to correlate the acquired data with time after last irrigation, and the corresponding VIPs, at discerning the most affected lipids.

3.2.6. Identification of metabolites

The final goal on metabolomics or lipidomics is to identify the biomarkers that give information of the organism state, considering the studied perturbation. From the mass to charge ratio of the selected variables, it is possible to identify the involved metabolite. However, the ionized fragments on MS are not detected as themselves, but forming adducts with other ions present on the medium, so the mass detected is not exactly the metabolite nominal mass [3]. In order to find an adduct of a metabolite with a mass to charge ratio that matches the one detected, it has to be searched on a database. Several metabolite related databases have been created in order to sustain the immense amount of data amassed. Among those, there are METLIN, HMDB (Human Metabolome Database) and Lipid Maps.

4.OBJECTIVES

This work aims at the realization of an untargeted lipidomic study on rice to evaluate the effects of the harvesting hour and time after irrigation. These two factors are related to circadian rhythm and water scarcity as a physical stressor. The study included the analysis of the most affected lipids, considering each factors alone, and the occasional identification of some as potential biomarkers.

Additionally, a preliminary study was carried out to compare several methodologies of lipid extraction from plant tissue, in order to choose the most appropriate. The extraction optimization considered three factors: the extraction method, the presence of a previous treatment with isopropanol and the amount of sample weighted.

5. EXPERIMENTAL SECTION

This work comprised two different parts: the optimization of the extracting method and the lipidomic study. Since some of the experimental procedure is common for both parts, it is presented as a unique section, pointing out the differences or specifications when necessary.

5.1. SOWING AND SAMPLING OF RICE

Each part of the study required a whole culture. A culture comprised a total of 180 sowed seeds of *Oryza sativa L. Japonica*. They were planted and distributed in groups of 9, in 20 different containers, as seen in figure 6.

Figure 6. Distribution of rice in a container

In order to accelerate the growing of rice once planted, the seeds were germinated in damp and warm conditions during the two previous days. To do so, the rice seeds were put in a beaker filled with layers of soaked paper at 30ºC, in an oven.

The seeds were planted in a mixture of equal volume of soil and vermiculite, along with $CaCO_3$, necessary to regulate the earth acidity. Then, soil was irrigated twice, assuring that the base contained enough water.

5.1.1. Growing conditions

Since each part of the study aims at a different goal, the growth conditions of the rice need specific considerations.

5.1.1.1. Testing of the extraction methods

In case of the rice destined to the testing of the extraction method, the plants were grown in the laboratory conditions. The plants were not irrigated directly, but the water was deposited on the container, for the plants to drain it. Each container was filled with 100 mL of Milli-Q quality water, on alternated days thrice a week, and always at 4 pm. The container was expected to be almost dry before irrigating it again, because the *Japonica* variety does not need an excess of water to grow.

The rice was harvested three weeks after sowing, at the same time, without distinction of any sample.

5.1.1.2. Lipidomic study

The rice destined to the lipidomic study was grown in more controlled conditions. They were cultivated in an environmental test chamber (Panasonic MLR-352H), showed in figure 7, which simulated the ideal conditions for the rice to grow. That included the light intensity, moisture and temperature, regulated to reproduce the fluctuations attached to the different hours, from day to night. Besides, the containers were periodically repositioned inside the chamber following random numbers, in order to minimize deviation caused by differences of any condition between positions.

Figure 7. Rice culture in the environmental test chamber.

The plants were irrigated similarly to the ones destined to the testing part, although at dawn (6pm), until the 11th day after sowing. The samples were harvested from the 14th to the 16th day after sowing, at different hours.

5.1.2. Sampling

The sampling was done in the third week after sowing for the plants destined for the testing part and on the second for the ones destined to lipidomics. Besides, the harvest for the lipidomic study required specific timing. It is summarized in table 1.

Day	Hour	Samples/ Hours after irrigation			
11th	6 am	Irrigated with 100 mL			
14th	5:30 am	Irrigated with 100 mL		Not irrigated	
		Harvesting			
14th	6 am	Sample 1	0 h	Sample 1'	72 h
	9 am	Sample 2	3 h	Sample 2'	75 h
	12 am	Sample 3	6 h	Sample 3'	78 h
15th	6 am	Sample 4	24 h	Sample 4'	96 h
	12 am	Sample 5	30 h	Sample 5'	102 h
16th	6 am	Sample 6	48 h	Sample 6'	120 h
	12 am	Sample 7	54 h	Sample 7'	126 h

Table 1. Irrigation and harvesting schedule.

As can be seen, the samples were collected in pairs. One of them was irrigated on the 14th day after sowing, and the other was not. This means that, between them, there is a difference of 72 hours regarding the last time being irrigated. Thus, the study includes 14 samples from the moment being irrigated to 5 days later.

A problem that entails working with living organisms is that there exist a biological variation between individuals. Since "omics" aims to find variation between classes of samples rather than variation inside classes, which is likely to lead to erratic conclusions, it has to be minimized. To achieve that, several plants were used per class in order to obtain a pooled sample. Hence, each sample on the lipidomic study corresponds to all the plants in a container. Since not all plants grow sufficiently, two containers had to be discarded, and two samples included the plants corresponding to two containers. The sampling of the testing part did not involve different classes, so the pooling included every plant.

To carry out the harvest, the soil was removed from the plants. The leaves were cut apart from the roots, and wrapped separately. Although both parts can be analyzed, this study will be

centered only on the leaves, discarding the roots. After that, the packages were immediately submerged in liquid nitrogen to quench the metabolome. The quenching was achieved by an instant freezing (-78ºC), and is necessary in order to stop any biochemical activity on the plant, and thus assure that the metabolome is not going to change before analyzing it. The packages were then kept on a freezer at -80ºC to prevent any subsequent activity.

A day before the extraction, the leaves were grounded and homogenized as part of the pool. It was done in a mortar provided with a compartment to be filled with liquid nitrogen. It enabled to keep the sample frozen. The milling was performed for every class sample separately, and the mortar was cleaned between each one. The ground samples were put in separated falcon tubes and lyophilized during the night.The appropriate amount of dried samples is weighted in an analytical balance just before the extraction phase.

On the lipidomic study, another consideration was taken referred to the sampling. In order to do a quality control through the process, all the ground leaves left were put together and vortexed, and two more samples were weighted from there. Those samples were treated seemingly to the rest until their analysis.

5.2. EXTRACTION PROCEDURE

The study started with an optimization of the extraction method. It evaluates the effect of three factors: the method used, the addition or not of hot isopropanol and the amount of sample analyzed. One factor contains two levels (isopropanol), while the other two have three, as shown in table 2.

Extracting method	Isopropanol treatment	Amount (mg)	Kind of Sample
Bligh-Dyer	No	5	BN
		10	BN
		15	BN
	Yes	5	BY
		10	BY
		15	BY
Folch	No	5	FN
		10	FN
		15	FN
	Yes	5	FY
		10	FY
		15	FY
MTBE	No	5	MN
		10	MN
		15	MN
	Yes	5	MY
		10	MY
		15	MY

Table 2. Scheme of the three factor optimization and the corresponding samples

Two replicates were done, so a total of 36 samples were weighted from the pooled rice once lyophilized. All were weighted in glass vials except from the ones in which the MTBE extraction without the isopropanol treatment was performed, which were weighted in 2 mL Eppendorf.

One half of the samples were submerged into 1mL of isopropanol and kept in an oven at 75ºC for 15 minutes [13]. After that, they were dried in a current of nitrogen.

The experimental proceeding in all three extraction methods is very similar, only differing in some steps, and obviously in the extracting mixture utilized. In figure 8 is shown a scheme that describe the general proceeding.

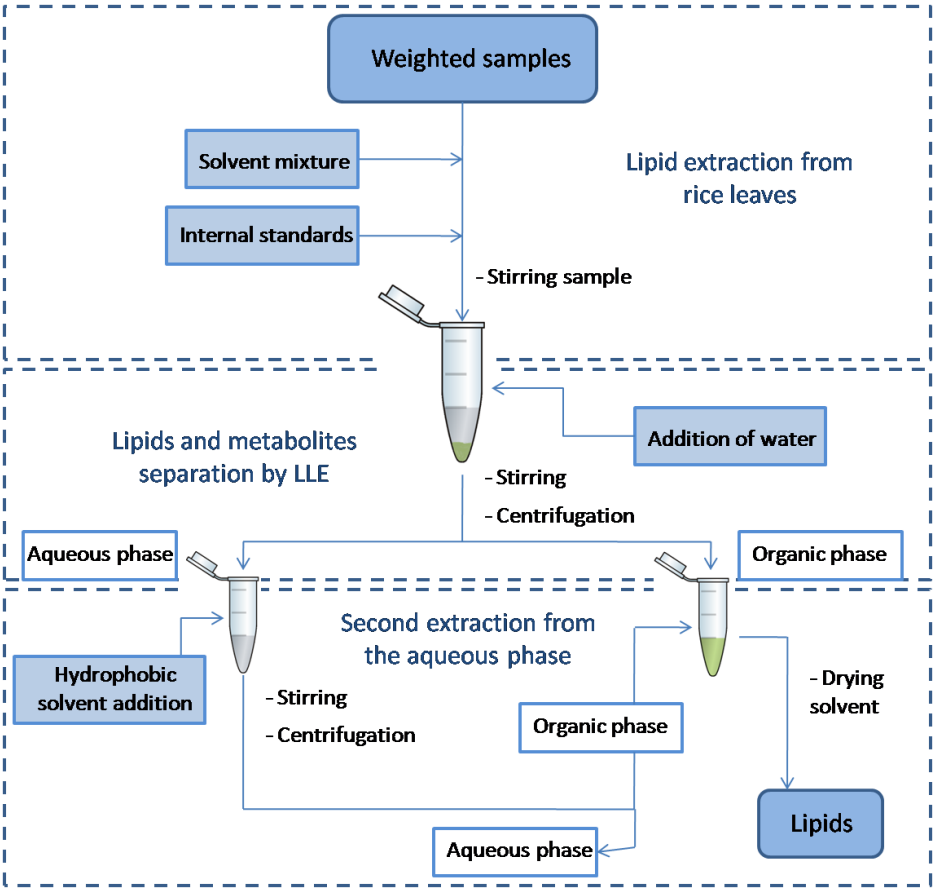

Figure 8. General scheme of lipid extraction methods.

The volumes and specific mixtures of solvents used for each method are shown in table 3. They were used as described in the literature, occasionally with some adaptations: Folch [9], Bligh-Dyer [10] and MTBE [11].

Method	Step	Solvents	Proportions	Volume (mL)
Bligh-Dyer	Extraction	$CHCl_3$:MeOH	1:2	1
	Separation	H_2O	-	0.41
	2nd extraction	$CHCl_3$:H_2O	1:1	1
Folch	Extraction	$CHCl_3$:MeOH	2:1	1
	Separation	H_2O	-	0.6
	2nd extraction	$CHCl_3$:MeOH	86:14	1
MTBE	Extraction	MTBE:MeOH	3:1	1
	Separation	H_2O:MeOH	3:1	0.5
	2nd extraction	MTBE:MeOH:H_2O	20:6:5	1

Table 3. Solutions used in each process of the extraction methods tested

Other specifications are described below:

- The sample with the extraction mixture was spiked with 10 µL of an internal standards solution. Its content is described in appendix 1. The solution was added to each vial or Eppendorf using an automatic micropipette.
- Each extraction solution contained a 0.01% of BHT, added previously by weighting.
- Once separated the two phases, the organic one was collected in an Eppendorf for each sample, using a 20-200 µL air displacement micropipette. On the MTBE method, the organic phase is the upper, so it can be directly collected. On Bligh-Dyer and Folch is the lower (chloroform), so the two phases were collected in different Eppendorfs. Every collecting required a new disposable tip, in order to avoid contamination or mixing between samples.
- The first stirring took one minute per sample, and was carried out using a vortex mixer, while the second and third one was shorter. In case of the Bligh-Dyer method, the second one was not necessary.
- After each stirring, the Bligh-Dyer and Folch samples were left at room temperature for about 10 minutes.
- Every centrifugation was done during 5 minutes at 2000 *g*.

- The samples of the MTBE method were sonicated in an ultrasonic cleaner, for 10 minutes, before adding the water mixture.
- The Bligh-Dyer samples required to be cooled at -18ºC in order to separate phases after the second extraction.

After the extraction, lipids were kept dry at -80ºC.

Once finalized the optimization, the most suitable methodology was chosen. Consequently, the extraction on the subsequent lipidomic study was carried out basing on MTBE method without the previous step. However, it had to be weighted 5 mg of sample instead, because of the little amount obtained per class. Since it was more suitable weighting 10 mg, all volumes used were reduced to a half, in order to maintain the ideal proportions. This led to the need of using insert vials on the LC injection, or otherwise the LC syringe cannot inject the sample properly.

5.3. INSTRUMENTAL PROCEDURE

The separation and analysis of lipids was performed using a LC-MS. Nevertheless, the MS analyzer used was different for each part.

5.3.1. Instrumentation

For the analysis of lipids, a different LC-MS instrumentwas used for each part. On the testing part, which did not needaccurate mass data, was used an UPLC-TQD (Waters Acquity UPLC System). Opposite to it, the lipidomic study required accurate mass values in order to identify potential biomarkers, so it was used an UPLC-TOF (Waters Acquity UPLC System) instead. The chromatographic column used (ACQUITY UPLC BEH C8, 100 x 2.1 mm, 1.7 µm) was common for both parts.

The data obtained from the LC-MS is acquired in RAW format, owned by Waters®. This is converted to CDF format using Databridge from the program MassLynx (Waters®). CDF format, in contrast of RAW, is compatible with Toolbox from MATLAB®, which was the program used in chemometrics.

Concerning the pre-treatment of the data before being analyzed by PCA or PLS, it depended on the vectors utilized.

On TIC chromatograms vectors, it was performed a baseline correction (Automatic Weighted Least Squares), Variable alignment (COW; 5, 50), which synchronizes spectrums, and mean centering, which eliminates the magnitude scale.

On MetaboNexus treated data, the baseline correction and alignment are already applied. So data was auto scaled, which eliminates magnitude and width scale.

5.3.2. Reconstitution of samples

In order to be injected into the chromatograph, the lipids of each sample have to be reconstituted. To do so, 250 mL of methanol were added to each Eppendorf, and after that, 10 µL of a second solution of internal standards (appendix 1). Each Eppendorf was stirred in a vortex mixer for about a minute. After that, 200 mL of the surfacing solution was transferred to a vial, using a 20-200 µL air displacement micropipette.

5.3.3. Mobile phase conditions

The chromatographic method used was already established, and it is suitable for the elution of lipids on untargeted lipidomics [17]. It is a non-isocratic method that took 22 minutes for the total elution of lipids (table 4). It consisted on the use of two main eluting solvents, prepared using MS quality solvents. The solvents and the utilized salts are going to for the most probable adducts in MS:

- Solution A: Methanol, 1 mM ammonium formate, 0.2% formic acid.
- Solution B: Water, 2 mM ammonium formate, 0.2% formic acid.

Time (min)	% Solvent A	% Solvent B	Elution mode
0 - 3	80 → 90	20 → 10	Lineal
3 - 6	90	10	Isocratic
6 - 15	90 → 99	10 → 1	Lineal
15 - 18	99	1	Isocratic
18 - 20	99 → 80	1 → 20	Lineal
20 - 22	80	20	Isocratic

Table 4. Eluting conditions.

The injection order was at random, in order to avoid systematic error for instrumental deviation.

6. EXTRACTION EVALUATION

The 36 samples of rice, extracted under different conditions, were converted into the corresponding TIC chromatograms. The retention times as variables and the corresponding intensities were used to make a PCA model. The scores diagram shows the samples grouped according to their similarity. Qualitatively, it is considered that a factor affect to the extraction phase if the different groups of samples are separated from the others. A factor is also considered relevantif the samples disposition is clearly influenced by this factor.

A first look on the scores diagram (figure 9) formed by the PC1 and PC2, which are the ones that describe most part of the samples variability, clearly shows that the amount of sample weighted is the most influent factor:

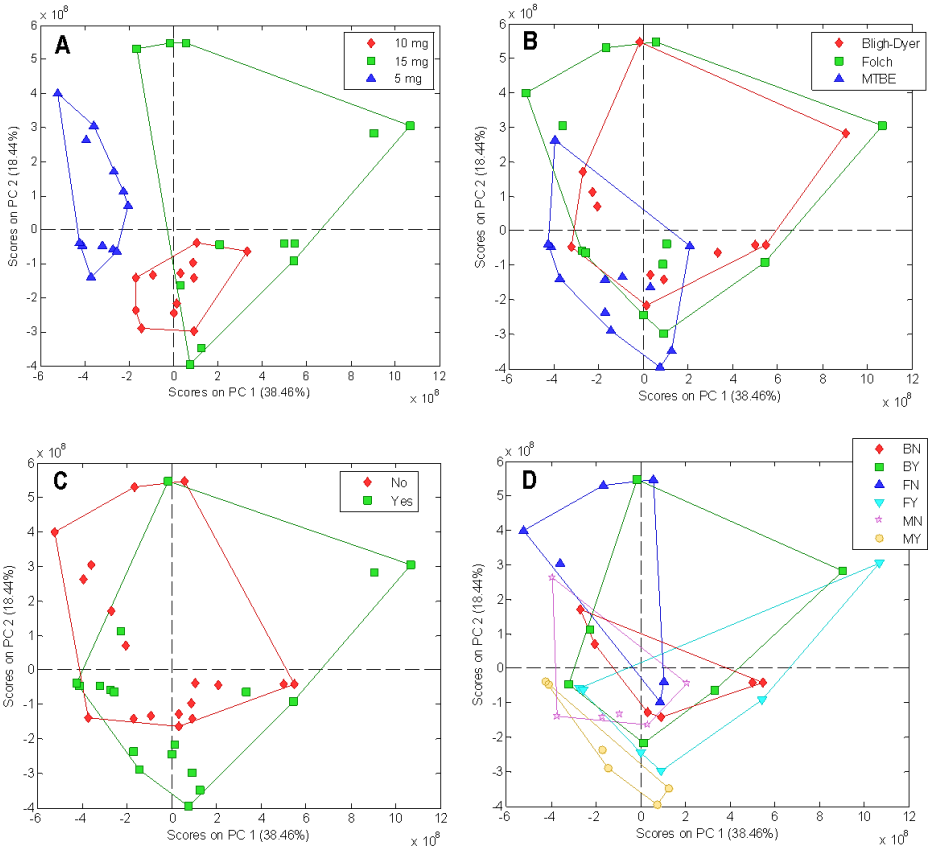

Figure 9. Samples represented on the scores diagram, showing: (A) amount of sample weighted; (B) extraction method; (C) isopropanol; (D) B and C combined.

It may appear obvious because the amount of sample is directly proportional to the metabolites concentration resulting to the extraction. However, the preprocessing applied on the model includes baseline correction and mean centering, which correct that differences. Therefore, this factor influences beyond the expected. According to the bibliography, the ideal proportions of solvents would correspond to 10 mg of sample weighted. It seems to be confirmed by their position. The corresponding samples are situated in the middle of the diagram and very close to each other, significant of little variance among the class. It is important concerning the method choice, because lipidomics goal is to find variance between controls and perturbed samples and the related metabolites. Hence, the methodology itself has to contribute the lower variance possible not to eclipse the aimed variance. It is similar to the 5 mg samples, quite close to each other. However, the 15 mg samples appear quite dispersed onto the diagram. A feasible explanation is due to solvent saturation in the extraction phase, which could have been a source of error, so it can be discarded.

Since the amount of weighted sample is the most influencing factor, the other ones are quite overshadowed. The separation of the corresponding classes is more visible when representing higher PCs, but it is not obvious. The effect of those factors is crucial to decide the how to plan the lipidomic study, so further analysis is needed. In order to determine whether the other two factors affect, the 15 mg samples are discarded, for they involve too much irrelevant variation, and a new PCA model is built.

In the new model scores, shown in figure 10, PC1 serves to explain the variance caused by the amount of sample once more. However, the representation of PC2 and PC3 in the scores diagram distinguishes the levels of each factor. The most visible is the separation due to the presence or not of the isopropanol treatment, so it seems that is also relevant.

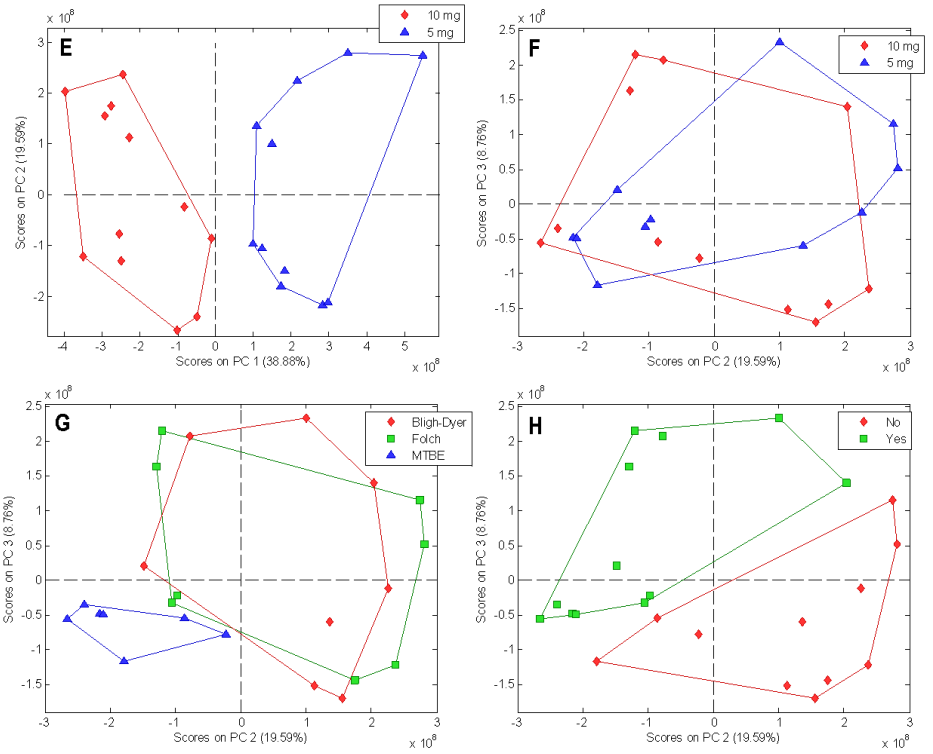

Figure 10. Scores diagram of the new model showing: (E) amount of sample in PC1 and PC2; (F) amount of sample in PC2 and PC3; (G) Method in PC2 and PC3; (H) isopropanol in PC2 and PC3.

Considering the extracting method, Folch and Bligh-Dyer cannot be set apart but they are separated from MTBE. So it seems that the extracting method is a factor to be considered apart from the easiness to carry out. A remarkable fact is that, whereas Folch and Bligh-Dyer methods clearly appear different whether the isopropanol treatment was applied or not, the MTBE samples are very close each other. It is more obvious if those two factors are considered together (figure 11), so 6 classes of samples are considered (named as in table 2).

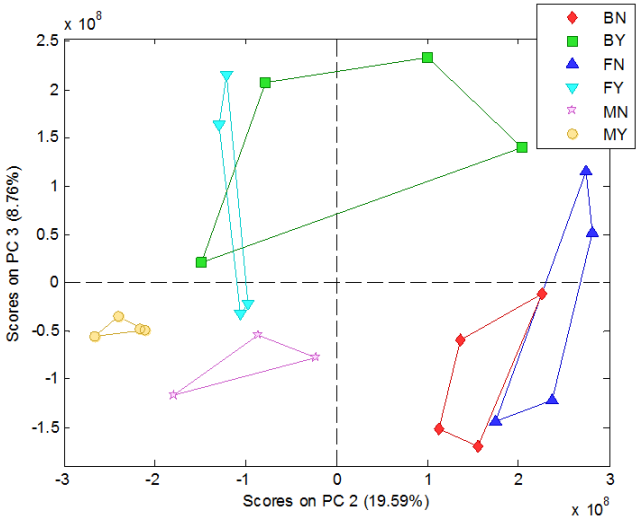

Figure 11. Scores diagram representing method and isopropanol treatment factors together.

PC2 distinguish horizontally the non-treated Folch and Bligh-Dyer samples (dark blue and red, respectively) from the rest. As PC2 is more important than PC3 at describing the samples variance, it can be concluded that the MTBE samples are more similar to the isopropanol treated Folch and Bligh-Dyer samples than the non-treated. According to the bibliography, $CHCl_3$ could result activate phospholipase D and lead to the hydrolyzation of some phospholipids. Nevertheless, as stated by the samples similitude, this did not appear to happen on MTBE extracted samples. So the use of the isopropanol treatment would be kept for $CHCl_3$ extractions (namely Folch and Bligh-Dyer). Therefore Folch and Bligh-Dyer extractions alone and MTBE along with the isopropanol treatment are discarded.

Examining similitude inside the left classes, Bligh-Dyer (green) is the one that seems to have more variance. This fact along with the difficulty attached at certain points of the experimental proceeding are evidences to discard it, too.

Concerning MTBE (pink) and pretreated Folch (cyan), the second one apparently leads to more variance, but the difference is attributed to the amount of sample, instead. Seeing that the use of those two methodologies does not bring much difference in the results, the choice is based on the easiness to carry out. This leads to consider the MTBE method more suitable, because the time-consuming isopropanol treatment is not necessary. An additional step on the

process generally increases the probability to make errors or sample loss. Besides, MTBE is known for being less hazardous than $CHCl_3$, so it is another advantage.

7. EFFECTS OF WATER AVAILABILITY AND HARVESTING HOUR

As said before, the lipidomic study was aimed at the evaluation of drought effects on periodically collected samples after a final irrigation. Moreover, the study contemplates an additional factor, which is the hour of harvesting. Whether those factors influence on the lipidome or not can be observed seemingly as done in the previous part. If the variance among samples comes from the effect of one of those factors, the samples can, therefore, be separated into the corresponding classes. However, unlike the previous part, it is just a preliminary study. The following considerations depend on the magnitude of each factor effect. The corresponding scores diagrams are shown in figure 12:

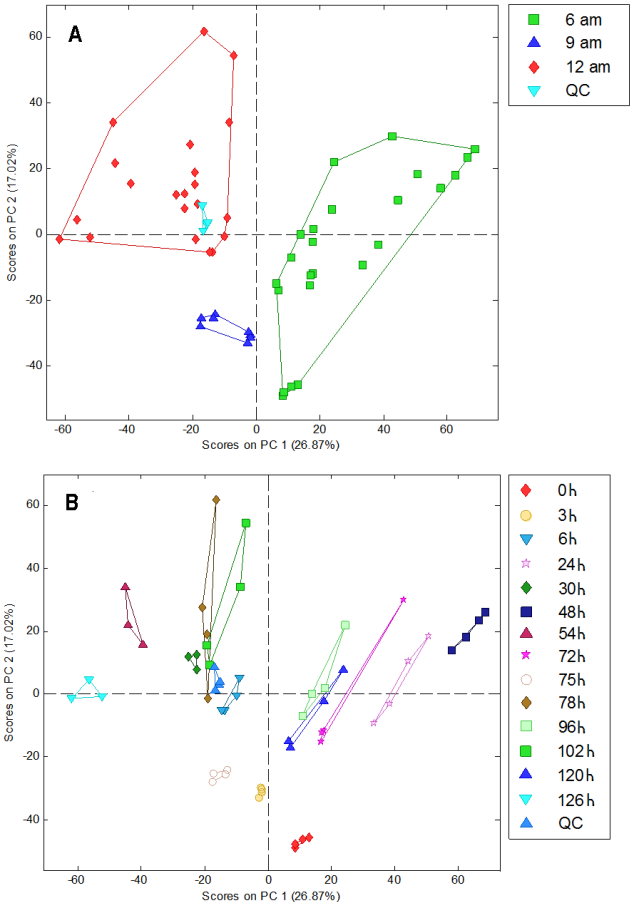

Figure 12. Scores diagrams showing: (A) harvesting hour and (B) time from last irrigation.

As can be seen, samples are separated into the corresponding classes in both factors. Besides, it is PC1, which is the one that explains more variance (26.87%), that separate from samples harvested at 6 am to 12 am. Inside each group, it is visible a separation too, considering the amount of hours after last irrigation. PC2 sets apart samples harvested at 9 am from the rest, along with the first sample harvested. It had been irrigated moments before, so its lipidome was likely to change during the initial draining. This difference caused by the process of absorbing water could also be applied to the sample harvested three hours later.

Another consideration is the position of the quality controls, injected along the process. They are very close to the center, which was the expected fact. Since it was obtained from a global pooled sample, the variance among individual classes was supposed to disappear. So, they had to represent all samples but any in particular at the same time.

Globally, samples appear very close to others of the same class. Along with the quality controls result, it proved the effectiveness of the extraction method chosen.

Thus, the effect of both factors is relevant onto the lipidome and further analysis can be done. Two PLS models were created: one using the harvesting hour values and the other using the number of hours from last irrigation. Each model correlated the data with the utilized factor, considering the variables that better describe the associated variability. The scores diagrams showing PLS components are similar to PCA, albeit they are conditioned by the factor information included in the model. For this reason, the first data analysis was performed with PCA instead as showed in figure 12. The quality controls were removed from PLS models, as they were specifically designed not to describe any particular class.

PLS was focused on the importance of the variables at describing the factor related variability, using VIP index.

The first factor considered was time after last irrigation, namely water availability. Its correlation to the data was good, and four components were used, with a description of **Y** variance of 93.14%. The loadings diagram representing the VIP values of each variable can be seen in figure 13. The variables that have a VIP value higher than 1, which appear above the red line, are relevant to describe the variance associated with the factor. There is a total of 3367 variables, corresponding to mass to charge ratios from 250 to 1800, but only 918 have a VIP value higher than 1 and, therefore, are important for the model, so it the amount is considerably reduced.

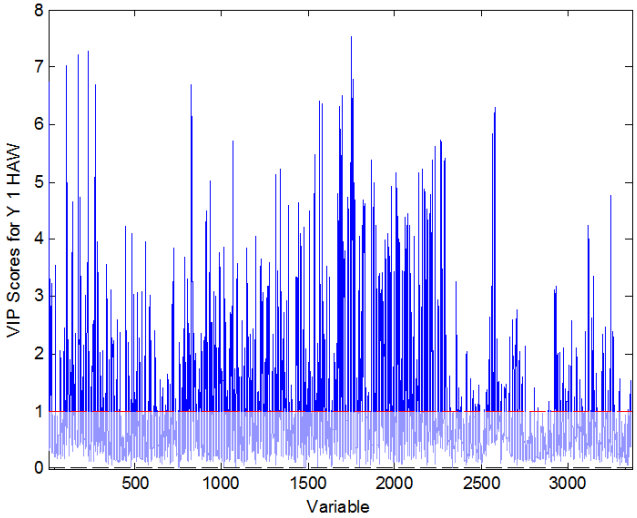

Figure 13. VIP loadings for water availability factor.

Even though there are variables more important than others among the ones with VIP higher than 1, all are important for the model. Thus, every one of these variables could be a potential biomarker. However, this study will consider just the most important ones. Several variables represent different isotopic forms of the same lipid, though.

The other factor studied was time of harvesting (figure 14).

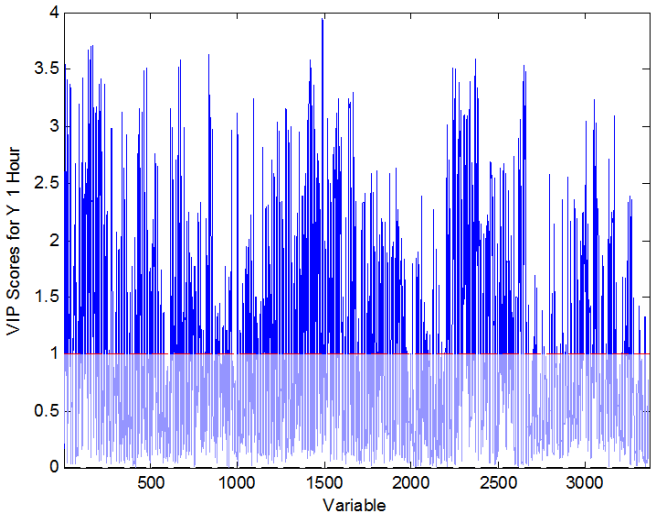

Figure 14. VIP loadings for harvesting time.

Its correlation to the data was even better than in water availability, and only three components were needed, with a description of **Y** variance of 96.57%.

The variables with a VIP value higher than 1 are 1251 from 3367. So there are more lipids influenced through time, and there is less difference of relevance. It is coherent, as many lipids are subjected to fluctuation over time, orchestrated by the circadian clock, albeit they can be related to many pathways. Another difference of time compared to water availability is that there are much more low mass lipids (the variables are ordered from lower to higher m/z ratio). In fact, there are only four variables with m/z value lower than 500 that are relevant to describe water availability.

The subsequent step was the identification of the lipids described by the relevant variables, which could be feasible candidates to be biomarkers. This term is more centered on water availability factor, because it is a physical stressor. Besides, as shown in figure 13, the variables of interest are very few and their VIP value were considerably higher compared to the rest. On the other hand, a study through time does not include any perturbation, so the term biomarker is not as descriptive.

Some of the most relevant variables in each factor have been summarized below, in table 5 for water availability factor and in table 6 for time. It has been proposed a possible lipid too, based on the most feasible encountered in HMDB. Some criteria were the presence of the corresponding adducts on the process and the discarding of hydrophilic metabolites. It has been searched the origin too, and the ones showed are likely to be encountered in plants.

Variable	m/z value	Elutiontime (min)	Error (ppm)	Proposed lipid
177	419,3159	4,15	<5	11'-Carboxy-alpha-chromanol
228	448,3516	4,91	<10	PE(22:5(7Z,10Z,13Z,16Z,19Z)/24:0)
277	486,3600	1,54	<5	Glabrolide
1314	779,5469	5,52	<5	1,28-Octacosanediol diferulate
1342	784,5476	5,10	<5	PC(20:5(5Z,8Z,11Z,14Z,17Z)/14:0)
1540	819,5131	3,62	<5	Quinquenoside F1
1567	826,5969	5,09	<5	PE(20:1(11Z)/20:4(8Z,11Z,14Z,17Z))
1737	860,7713	15,89	<5	TG(18:0/20:2n6/14:1(9Z))
1753	866,7254	14,24	<5	TG(20:3n6/14:1(9Z)/18:3(9Z,12Z,15Z))
1869	888,8008	16,53	<5	TG(18:0/16:1(9Z)/18:2(9Z,12Z))
2009	916,7249	11,42	<20	TG(18:3(9Z,12Z,15Z)/20:3n6/18:4(6Z,9Z,12Z,15Z))
2215	966,8200	17,07	<30	TG(22:1(13Z)/18:2(9Z,12Z)/18:3(9Z,12Z,15Z))
2228	970,8705	18,23	<10	TG(15:0/20:4(8Z,11Z,14Z,17Z)/24:0)
2265	984,8880	18,71	<10	TG(18:0/20:4(8Z,11Z,14Z,17Z)/22:0)
2286	992,8396	17,84	<25	TG(20:2n6/22:1(13Z)/20:4(5Z,8Z,11Z,14Z))

Table 5. Feasible identity of biomarkers for water scarcity stress.

As can be seen, there are several variables that have triglycerides (TG) as the most probable option. All of them formed adducts with ammonium, present in the chromatographic conditions. The retention time are close to the TG internal standard (appendix 1), which is about 17.6 minutes, and the mass error is low. Thus, it could be thought that a metabolic pathway related to the biosynthesis of TG could be affected by drought. However, this analysis would be beyond the objectives of this present work.

Variable	m/z value	Elutiontime (min)	Error (ppm)	Proposed lipid
13	277,2149	1,54	<5	Stearidonic acid
22	293,2112	1,04	<15	Hydroxyhexanoycarnitine
152	399,3651	5,36	<10	Campest-4-en-3-one
161	411,3633	4,74	<5	Stigmasta-4,6-dien-3-one
210	442,3677	4,09	<5	(22E,24R)-Stigmasta-4,22-diene-3,6-dione
212	443,3526	3,25	<5	Camellenodiol
466	576,5117	13,43	<5	1,2-Epoxy-1,2,7,7',8,8',11,12-octahydro-psi,psi-carotene
662	615,5021	6,37	<10	DG(16:1(9Z)/20:4(5Z,8Z,11Z,14Z)/0:0)
1413	794,5828	6,37	<10	PG(18:1(9Z)/18:0)
1423	796,5913	6,36	<10	PE(20:3(5Z,8Z,11Z)/20:1(11Z))
2340	1006,8022	15,59	<20	TG(22:6(4Z,7Z,10Z,13Z,16Z,19Z)/20:2n6/22:6))

Table 6. Feasible identity of lipids which concentration is affected through time.

Concerning the study of the harvesting hour, the assignment of lipids corresponding to the most relevant variables in the PLS model, did not lead to any hypothesis. Even so, there is no evidence that proves their identity, but they are just proposed structures. Seemingly, their accurate identification is beyond this work objectives.

8. CONCLUSIONS

After the realization of this work, each part results were considered. The extraction optimization led to the following conclusions:

- The MTBE method without the previous step was the most appropriate extraction method for the subsequent lipidomic study.
- Bligh-Dyer and Folch methods provided similar results but their effectiveness was conditioned by the isopropanol treatment, while MTBE was not affected by it.
- Even though the extraction method and isopropanol treatment were relevant, the amount of sample weighted was the most influent factor.

Concerning the main lipidomic study, it was concluded that:

- Time of harvesting was the most conditioning factor. However, water scarcity affected rice lipidome as well. Thus, PLS models could be built.
- VIP values on both models revealed which variables were more important to describe water availability and time of harvest, respectively. Thus, their mass to charge values led to potential biomarkers.
- Based on the results, a future work could be focused on the proper identification of biomarkers and the association to metabolic pathways.

9. REFERENCES AND NOTES

1. Patti, G. J.;Yanes, O. Metabolomics: The apogee of the omics trilogy. Nature Rev Mol Cell Biol. **2012**, 13(4), 263-269.
2. Dettmer, K.; Hammock, B. D. Metabolomics - A new exciting field within the "omics" sciences. Environ Health Perspect. **2004**, 112(7), 396-397.
3. Dettmer, K.; Aronov, P. A.; Hammock, B. D. Mass spectrometry-based metabolomics. Mass Spectrom Rev. **2007**, 26(1), 51-78.
4. Weckwerth, W. Metabolomics in systems biology. Annu Rev Plant Biol. **2003**, 54, 669-689.
5. Checa, A.; Bedia, C.; Jaumot, J. Lipidomic data analysis: Tutorial, practical guidelines and applications. Analytica Chimica Acta. **2015**.
6. Food and Agriculture Organization of the United Nations: Statistics division. http://faostat3.fao.org (accessed April 26, 2015).
7. Majeed, A.; Salim, M.; Bano, A.; Asim, M. Physiology and productivity of rice crop influenced by drought stress induced at different developmental stages. African Journal of Biotechnology. **2011**, 10(26), 5121-5136.
8. Sato, S.; Arita, M.; Soga, T. Time-resolved metabolic modulation in rice foliage. BMC Systems Biology. **2008**, 2:51.
9. Catja, T.; Fiehn, O. Comprehensive analysis of lipids in biological systems by liquid chromatography-mass spectrometry. Trends in Analytical chemistry. **2014**, 61, 192-206.
10. Okazaki, Y.; Kamide, Y.; Hirai, M.Y.; Saito, K. Plant lipidomics based on hydrophilic interaction chromatography coupled to ion trap time-of-flight mass spectrometry. Metabolomics. **2013**, 9 (SUPPL.1), 121-131.
11. Hummel, J.; Segu, S.; Li, Y.; Irgang, S.; Jueppner, J.; Giavalisco, P. Ultra Performance Liquid Chromatography and High Resolution Mass Spectrometry for the Analysis of Plant Lipids. Front Plant Sci. **2011**, 2(October), 1-17.
12. Matyash, V.; Liebisch, G.; Kurzchalia, T. V.; Shevchenko, A.; Schwudke, D. Lipid extraction by methyl-tert -butyl ether for high-throughput lipidomics. J Lipid Res. **2008**, 49(17), 1137-1146.
13. Shiva, S.; Hieu, S. V. Lipidomic Analysis of Plant Membrane Lipids by Direct Infusion Tandem Mass Spectrometry. Methods in Molecular Biology. 2013, 1009, 79-91.
14. Massart, D. L.; Vandeginste, B. G. M.; Buydens, L. M. C.; De Jong, S.; Lewi, P. J.; Smeyers-Verbeke, J. Handbook of Chemometrics and Qualimetrics: Part A; Elsevier: Amsterdam, **1997**.
15. Esbensen, K. H.; Guyot, D. Multivariate Data Analysis-in practice. 4th ed., Camo: Trondheim, **2000**.
16. Wold, S.; Sjöström, M.; Eriksson, L.PLS regression: a basic tool of chemometrics. Chemometrics andIntelligent Laboratory Systems. **2001**, 58(2), 109-130.
17. Merrill, A. H. Jr.; Sullards, M. C.; Allegood, J. C; Kelly, S.; Wang, E. Sphingolipidomics: high-throughput, structure-specific, and quantitative analysis of sphingolipids by liquid chromatography tandem mass spectrometry. Methods. **2005**, 36(2), 207-224.

10. ACRONYMS

BHT	Butylated hydroxytoluene
ESI	Electrospray Ionization
GC	Gas Chromatography
HILIC	Hydrophilic Interaction Liquid Chromatography
HMDB	Human Metabolic Database
LC	Liquid Chromatography
MS	Mass Spectrometry
MTBE	Methyl tert-butyl ether
NMR	Nuclear Magnetic Resonance
PC	Principal Component
PCA	Principal Component Analysis
PLS	Partial Least Squares
TIC	Total Ion Current
TOF	Time of Flight
TQD	Triple Quadrupole Detector
UPLC	Ultra Performance Liquid Chromatography
VIP	Variable Important in Projection

APPENDICES

APPENDIX 1: INTERNAL STANDARDS

Lipid family	Standards	pmoladded[a]
Neutral glycerolipids	1,2,3-17:0 TG	200
	1,3-17:0 D5 DG	166
	17:0 MG	145
Sterols	17:0 cholesterol ester	188
Phospholipids	16:0 D31-18:1 PA	411
	16:0 D31-18:1 PC	126
	16:0 D31-18:1 PE	133
	16:0 D31-18:1 PS	123
	16:0 D31-18:1 PG	125
Lysophospholipids	17:1 Lyso PE	200
	17:1 Lyso PG	200
	17:1 Lyso PS	200
	17:0 Lyso PA	140
	17:0 Lyso PC	120

(a) pmoladded to thesamples (corresponding to 10 µL.

List of internal standards added in the extraction phase.

PA	Phosphatidic acid
PC	Phosphatidylcholine
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PS	Phosphatidylserine