

# Integrated Masters in Bioengineering

## *Nuclear magnetic resonance metabolomics applied to agriculture and food research*

### Masters Dissertation

of

Francisco Castro Ribeiro Couto

Developed within the scope of the Dissertation course

Carried out at

Department of Chemistry and Physics - Universidad de Almería



Supervisor at FEUP: Prof. Manuel Simões

Supervisors at UAL: Dr. Ana Cristina Abreu; Prof. Ignacio Fernández

Department of Chemical Engineering

22<sup>nd</sup> of February 2021





# Acknowledgments

First, I would like to thank Professor Manuel Simões for being so attentive to my interests and for opening up the possibility for me to develop this thesis, adding the availability to listen and constant help. His passion and knowledge for science clearly influenced me.

A special thanks to Doctor Ana Abreu for the closeness, constant happiness and positive energy at work. Without her guidance and knowledge transmission this project would not be possible. The high standards and critical thinking really made me adopt a certain work ethic with a rigorous attitude, that I will take with me to future challenges. Also, it felt really good to communicate in Portuguese sometimes. I greatly appreciate Professor Ignacio Fernández de las Nieves for giving me the opportunity to work in this project in Almería and for welcoming me so kindly. The close relation and concern with everyone on the group was something impressive and admirable. Thank you for the encouragement and accessibility. I cannot forget my colleagues in the laboratory, which besides making me feel so comfortable, were always ready to give a hand and uplift me.

I must express my deep gratitude to my family. Specially, to my mother, father and brother for the constant support, patience and company along these years. My mother kept incentivising my strength of will and never let me doubt my abilities. This work is dedicated to my grandfather, Bernardino, the strongest and sweetest person in the world. His passion for agriculture was decisive for my choice and has a profound impact in my goals.

I have to mention Ronja, the bright full moon. Her desire for constant learning coupled with the curiosity and attention to detail were determinant for my motivation and success. Thank you for the support and balance.

I am very pleased with the last 3 or 4 years of my studies. I have a lot of good memories with my colleagues. The hard times we have shared led to the creation of strong, unbreakable friendships. Thanks for your presence in this journey, for the help and critical opinions in this project. I'm glad to have met you and that I can count on you. Much more is yet to come.

I also want to acknowledge my friends outside of the University. Distinct educational paths only made us realize we are tightly committed to our friendship. I am truly grateful for your concern, support and continuous boost.

## Resumo

O tomate (*Solanum lycopersicum*) é um dos produtos hortícolas mais consumidos mundialmente e a sua produção e comercialização é de extrema importância para a economia da província de Almería. A competição extrema com outros países e a disseminação de pestes, como a *Tuta absoluta*, está a incentivar a investigação de métodos de produção mais eficientes que garantam frutos saborosos e nutricionalmente ricos. Neste estudo, plantas de tomate cultivadas durante o inverno em estufas empregando tetos duplos com plásticos de diferentes cores foram estudadas através de uma abordagem metabolómica utilizando como plataforma analítica a Ressonância Magnética Nuclear (RMN), com o objetivo de compreender o efeito dos diferentes tetos na qualidade nutricional e organolética dos tomates, e ainda o seu impacto na incidência da *Tuta absoluta*. A metabolómica tem como objetivo fornecer um retrato global dos metabolitos de pequeno tamanho molecular que estão presentes num determinado sistema biológico, livres de uma observação tendenciosa inerente a outros estudos focados no metabolismo. Para obter uma caracterização ampla do perfil metabólico dos tomates, estes foram extraídos com uma mistura de metanol/tampão fosfato (50:50, v/v) e os metabolitos foram identificados através da análise de medidas unidimensionais e bidimensionais de  $^1\text{H}$  RMN. Posteriormente, foram aplicadas várias técnicas de análise multivariada de dados aos espetros de  $^1\text{H}$  RMN dos tomates com o objetivo de: (1) estudar as modificações metabólicas nos frutos ao longo de cinco semanas de colheita, (2) avaliar o impacto dos tetos duplos de plástico azul, vermelho e transparente no metaboloma do tomate. De igual modo, a capacidade antioxidante dos tomates foi estudada através do método *Oxygen Radical Absorbance Capacity* (ORAC). Os dados de  $^1\text{H}$  RMN dos tomates foram analisados, em primeiro lugar, através da aplicação de um modelo exploratório de Análise de Componentes Principais (ACP), com o objetivo de identificar a variável neste estudo que mais influenciou o metaboloma destes frutos. De facto, não foi o tipo de teto que mais influenciou o tomate, uma vez que a análise ACP revelou uma maior discriminação dos tomates de acordo com a semana de colheita. As alterações metabólicas dos tomates ao longo da colheita foram evidentes através da aplicação de um modelo de Mínimos Quadrados Parciais (MQP): Valores mais elevados de ácidos orgânicos e aminoácidos foram discriminantes para as primeiras colheitas, enquanto um aumento de glicose e de compostos antioxidantes benéficos para a saúde caracterizaram os frutos colhidos mais tarde. Ambos os modelos de Mínimos Quadrados Parciais com Análise Discriminante (MQP-AD) e Projeção Ortogonal para Estruturas Latentes com Análise Discriminante (OPLS-AD) foram aplicados para identificar se existiam diferenças metabólicas devido ao tipo de cobertura plástica. De facto, o modelo OPLS-AD forneceu um modelo muito claro com os espetros agrupados de acordo com a sua cobertura. Em seguida, o metaboloma dos frutos foi comparado através da construção e análise de um mapa de calor. Foi claro o efeito positivo da aplicação de um teto duplo, uma vez que tomates de teto simples apresentaram um conteúdo metabólico global muito pobre em termos quantitativos. Os tomates cultivados sob o duplo teto transparente revelaram uma maior quantidade de açúcares e compostos antioxidantes, tais como ácidos hidroxicinâmicos, flavonoides e vitamina C, estando de acordo com os valores mais elevados

de capacidade antioxidante obtidos com estes frutos. De igual modo, observaram-se níveis elevados de aminoácidos. Os frutos cultivados sob o teto duplo de plástico azul revelaram a maior quantidade de aminoácidos essenciais e foram os segundos em termos de teor fenólico. Os tomates de teto duplo vermelho apresentaram o maior conteúdo em ácidos orgânicos e quantidades muito baixas de compostos antioxidantes. Os resultados obtidos pelo método ORAC foram proporcionais à quantidade de compostos antioxidantes encontrados para cada cobertura no mapa de calor. De modo geral, a aplicação de tetos duplos azul e transparente levou à obtenção de frutos mais saborosos e com uma maior capacidade antioxidante. Por último, pretendeu-se estudar o efeito dos revestimentos plásticos sobre a incidência de *Tuta absoluta* nas plantações. Deste modo, folhas de todas as plantas foram também colhidas, extraídas através de uma mistura de metanol/tampão fosfato (80:20, v/v) e analisadas por  $^1\text{H}$  RMN. Um modelo MQP-AD foi aplicado aos dados de  $^1\text{H}$  RMN das folhas, e permitiu concluir que as plantas cultivadas sob teto simples mostraram uma maior quantidade de metabolitos secundários associados à defesa das plantas, devido provavelmente a uma maior infestação observada nestas plantas por este inseto fitofágico. A prática agronómica de duplo teto levou a uma menor infestação pela praga.

**PALAVRAS-CHAVE:** *Solanum lycopersicum*, estufa, duplo teto, metabolómica por RMN, análise multivariada de dados, qualidade organolética, capacidade antioxidante, *Tuta absoluta*

## Abstract

Tomato (*Solanum lycopersicum*) is one of the most consumed crops worldwide and its production and commercialization is of highest importance to the economy of the Almería province. Competition from other countries and the spread of pests, such as *Tuta absoluta*, is pushing towards efficient production methods that guarantee tasty and nutritionally rich tomato fruits. In this project, tomato plants grown in the winter season under greenhouse double-roofs of distinct plastic colours were studied through a Nuclear Magnetic Resonance (NMR)-based metabolomics approach, to understand their effect on the crop performance, the organoleptic and nutritional value of tomato fruits, in addition to its impact on the incidence of *Tuta absoluta*. Metabolomics aims to provide a global snapshot of small-molecule metabolites in cells or biological systems, free of observational biases inherent to more targeted studies of metabolism. To obtain a wide coverage of the metabolic profile of tomato fruits, these were extracted with a solvent system of methanol/phosphate buffer (50:50, v/v) and compounds were identified by the analysis of one- and two-dimensional  $^1\text{H}$  NMR experiments. Then, several multivariate data analysis techniques were applied to the  $^1\text{H}$  NMR data to evaluate (1) changes in fruits along five weeks of harvests, and (2) the impact of uncoloured, blue and red plastic double-roofs in the tomato metabolome. Apart from the metabolomic analysis, the antioxidant activity of all fruits was measured by the Oxygen Radical Absorbance Capacity (ORAC). By means of a Principal Component Analysis (PCA), the  $^1\text{H}$  NMR data was explored to understand which variable most influenced the tomato population. In fact, it was not the type of covering that impacted most the tomatoes, as the PCA analysis revealed a higher discrimination according to the week of harvesting. Tomato metabolic changes along harvesting were evident by the application of a Partial Least Squares (PLS) model. A higher content in organic acids and amino acids was verified in tomatoes from the first harvest, while an increase in glucose and health-benefiting antioxidant compounds was found in later harvested fruits. Both Partial Least Squares Discriminant Analysis (PLS-DA) and Orthogonal Projection to Latent Structures Discriminant Analysis (OPLS-DA) models were applied to identify metabolic differences due to the kind of plastic covering. In fact, the OPLS-DA provided a much clearer model with the spectra grouped according to their cover. Then, the fruits' metabolome was compared by the construction and analysis of a clustered heatmap. It was clear the positive effect of the application of a double-roof, as tomatoes from a single-roof had an overall metabolic content of much poorer nutritional quality. Tomato fruits grown under the uncoloured double-roof revealed the highest amounts of major sugars, as well as of antioxidant compounds, such as hydroxycinnamic acids, flavonoids and vitamin C. These results were coherent with the ORAC assay, since these tomatoes showed the highest antioxidant activity. In addition, high levels of amino acids were also found. Fruits grown under blue plastic double-roof displayed the highest amount of essential amino acids and were second in terms of phenolics content. Tomatoes from red double-roof had the highest quantities in organic acids and very low amounts of antioxidant compounds. The results obtained by the ORAC method were proportional to the quantity of antioxidant compounds found for each covering in the heatmap.

Thus, the application of uncoloured and blue double-roofs rendered the production of tastier fruits with an enhanced antioxidant capacity. To study the effect of the plastic coverings on the *Tuta absoluta* incidence in the plants, leaves from all plants were also extracted with methanol/phosphate buffer (80:20, v/v) and measured by  $^1\text{H}$  NMR. A PLS-DA model was applied to the leaves  $^1\text{H}$  NMR data, which showed that leaves from plants grown in single-roof produced more secondary metabolites, which are linked to the plant defence. These results are coherent with the fact that these plants were the most attacked by this phytophagous insect. On the contrary, the double-roof agronomical practice led to a smaller infestation by this pest.

**KEY WORDS:** *Solanum lycopersicum*, greenhouse, double-roof, NMR-based metabolomics, multivariate data analysis, organoleptic quality, antioxidant capacity, *Tuta absoluta*

## Declaration

I, Francisco Castro Ribeiro Couto, declare, under honour, that this work is original and that all non-original contributions were duly referenced with identification of the source.

22<sup>nd</sup> of February 2021

Francisco Castro Ribeiro Couto

# Table of contents

|          |                                                                                       |           |
|----------|---------------------------------------------------------------------------------------|-----------|
| <b>1</b> | <b>Introduction .....</b>                                                             | <b>1</b>  |
| 1.1      | Almería as a leading horticultural producer .....                                     | 1         |
| 1.2      | Major throwbacks in tomato production .....                                           | 3         |
| 1.3      | The introduction of greenhouses .....                                                 | 5         |
| 1.4      | Tomato organoleptic and nutritional value .....                                       | 6         |
| 1.5      | Metabolomics and its suitability to study agronomical practices .....                 | 8         |
| 1.6      | NMR principles and features .....                                                     | 12        |
| 1.7      | Experimental steps in a metabolomic project .....                                     | 15        |
| 1.8      | Examples of the application of NMR-based metabolomics in the agricultural field<br>21 |           |
| <b>2</b> | <b>General aims .....</b>                                                             | <b>22</b> |
| <b>3</b> | <b>Materials and methods .....</b>                                                    | <b>23</b> |
| 3.1      | Plant material and greenhouse experiment layout .....                                 | 23        |
| 3.2      | Sample preparation and extraction .....                                               | 27        |
| 3.3      | NMR experiments .....                                                                 | 28        |
| 3.4      | Oxygen radical absorbance capacity (ORAC) assay .....                                 | 28        |
| 3.5      | Statistical analysis .....                                                            | 30        |
| <b>4</b> | <b>Results and Discussion .....</b>                                                   | <b>31</b> |
| 4.1      | Metabolite profiling on tomato fruits .....                                           | 31        |
| 4.2      | Coloured double-roof effect on tomatoes antioxidant activity .....                    | 36        |
| 4.3      | Tomato production data .....                                                          | 39        |
| 4.4      | Multivariate data analysis .....                                                      | 41        |
| 4.4.1    | Exploratory analysis (Principal Component Analysis) .....                             | 41        |
| 4.4.2    | Tomatoes metabolic evolution along the harvests .....                                 | 43        |
| 4.4.3    | Tomatoes metabolic profile as a function of the cover colouring .....                 | 47        |
| 4.4.4    | Plastic cover colouring effect on leaves defence metabolites .....                    | 51        |
| <b>5</b> | <b>Conclusion and future recommendations .....</b>                                    | <b>57</b> |

|          |                                                                        |           |
|----------|------------------------------------------------------------------------|-----------|
| <b>6</b> | <b>References .....</b>                                                | <b>59</b> |
| <b>7</b> | <b>Annexes .....</b>                                                   | <b>74</b> |
|          | <b>Annex A - Metabolite profiling on tomato fruits .....</b>           | <b>74</b> |
|          | <b>Annex B - Oxygen radical absorbance capacity (ORAC) assay .....</b> | <b>80</b> |
|          | <b>Annex C - Multivariate Data Analysis .....</b>                      | <b>84</b> |

# Notation and glossary

## *List of abbreviations*

|                  |                                                                                       |
|------------------|---------------------------------------------------------------------------------------|
| 1D               | One-dimensional                                                                       |
| 2D               | Two-dimensional                                                                       |
| AAPH             | 2,2'-Azobis(2-amidinopropane)dihydrochloride                                          |
| ANOVA            | Analysis of Variance                                                                  |
| AUC              | Area Under Curve                                                                      |
| AUC <sub>c</sub> | Area Under Curve Clean                                                                |
| ChEBI            | Chemical Entities of Biological Interest                                              |
| COSY             | <sup>1</sup> H- <sup>1</sup> H Correlation Spectroscopy                               |
| DA               | Discriminant Analysis                                                                 |
| DDS              | 2,2-Dimethyl-2-silapentane-5-sulfonic Acid                                            |
| FA               | Fatty Acid                                                                            |
| FAO              | Food and Agriculture Organization                                                     |
| GABA             | γ-Aminobutyric Acid                                                                   |
| GC-MS            | Gas Chromatography-Mass Spectrometry                                                  |
| HMBC             | <sup>1</sup> H- <sup>13</sup> C Heteronuclear Multiple Bonds Coherence Spectroscopy   |
| HMDB             | Human Metabolome Database                                                             |
| HSQC             | <sup>1</sup> H- <sup>13</sup> C Heteronuclear Single-quantum Correlation Spectroscopy |
| LC-MS            | Liquid Chromatography-Mass Spectrometry                                               |
| LDPE             | Low Density Polyethylene                                                              |
| LSD              | Least Significant Difference                                                          |
| KEGG             | Kyoto Encyclopaedia of Genes and Genomes                                              |
| MetPa            | Metabolic Pathway Analysis                                                            |
| MI               | Myo-inositol                                                                          |
| MS               | Mass Spectroscopy                                                                     |
| MSEA             | Metabolite Set Enrichment Analysis                                                    |
| NMR              | Nuclear Magnetic Resonance                                                            |
| noesy            | <sup>1</sup> H-Nuclear Overhauser Effect Spectroscopy                                 |
| OPLS             | Orthogonal Projection to Latent Structures                                            |
| ORAC             | Oxygen Radical Absorbance Capacity                                                    |
| PA               | Photosynthetic Activity                                                               |
| PAR              | Photosynthetic Active Radiation                                                       |
| PCA              | Principal Component Analysis                                                          |
| PLS              | Partial Least Squares                                                                 |
| PUFA             | Poly Unsaturated Fatty Acid                                                           |
| SSB              | Single Side Band                                                                      |
| TE               | Trolox Equivalent                                                                     |
| TOCSY            | <sup>1</sup> H - <sup>1</sup> H Total Correlation Spectroscopy                        |
| TSP              | 3-(trimethylsilyl)-2,2',3,3'-Tetradeuteriopropionic Acid                              |
| TTMS             | Tetrakis(trimethylsilane)silane                                                       |
| UFA              | Unsaturated Fatty Acid                                                                |
| VIP              | Virtual Importance to Projection                                                      |

# List of Figures

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |           |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| <i>Figure 1 - Satellite point of view from greenhouse concentration in the district of Campo de Dalias in the city of Almería. Image obtained from Google Maps (December 13, 2020). .....</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | <i>2</i>  |
| <i>Figure 2 - Tomato leaf miner (Tuta absoluta) moth in adult phase (top), larvae (bottom left) and its deteriorating effect on tomato leaves and fruit. Images obtained from Hortidaily (2019) and Albida Agriculture (2018).....</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | <i>3</i>  |
| <i>Figure 3 - Evolution of publications on metabolomics applied to agricultural and food sciences from 2010 to 2019. The data was generated from an advanced research in PubMed database with the following keywords: (("Food Technology"[Mesh]) OR "Agriculture"[Mesh]) AND "Metabolomics"[Mesh]). There is an official list of Medical subject headings (MeSH), with which the indexers label an article regarding its content. Applying MeSH terms can lead to a more efficient research, as the resulting articles will be necessarily about the suggested topic instead of just mentioning it (Baumann, 2016). .....</i>                                                                                     | <i>10</i> |
| <i>Figure 4 - 1D <sup>1</sup>H-NMR spectra of a tomato fruit sample. Spectra was obtained in a Bruker Avance III 600 MHz spectrometer and visualized using the Bruker TopSpin (3.6.2) software. ....</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | <i>14</i> |
| <i>Figure 5 - Scheme of the successive steps carried out on a metabolomic study applied to plant sciences study. ....</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | <i>16</i> |
| <i>Figure 6 - U9 and U11 greenhouses dimensions (a); U12 greenhouse dimensions. ....</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | <i>23</i> |
| <i>Figure 7 - Greenhouses positioning (Latitude: 36°51' "N; Longitude: 2° 17' W; Altitude: 87 m). N-north; O-west; E-east; S-south. ....</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | <i>24</i> |
| <i>Figure 8 - Images from tomato culture inside the experimental greenhouses. U9 E with standard Agripolyane double roof (a); U9 O with photo-selective blue plastic double roof (b); U11 E with photo-selective red plastic double roof (c) ; U11 O without double roof (d) ; U12 E with standard Agripolyane double roof (e) and U12 O with photo-selective red plastic double roof (f). ....</i>                                                                                                                                                                                                                                                                                                               | <i>25</i> |
| <i>Figure 9 - Tuta absoluta butterflies in the insect net inside the side window of the U11E (a), black chromatic trap on the substrate container (b) and galleries in tomato leaves (c) on March 3 2020. ...</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | <i>26</i> |
| <i>Figure 10 - <sup>1</sup>H NMR (600 MHz) spectrum of a tomato fruit CD<sub>3</sub>OD/D<sub>2</sub>O KH<sub>2</sub>PO<sub>4</sub> buffer (50:50, v/v) extract. Aden - adenosine, aden-like - adenosine like, Ala - alanine, Arg - arginine, Asn - asparagine, Asp - aspartate, cho - choline, FA - fatty acids, fru - fructose, glc - glucose, Glu - glutamate, Gln - glutamine, Ile - isoleucine, Leu - leucine, Lys - lysine, MeOH - methanol, MI - myo-inositol, Phe - phenylalanine, querc(deriv) - quercetin derivative, suc - sucrose, thr - threonine, trp - tryptophane, tyr - tyrosine, urid - uridine, val - valine. Chemical shifts, coupling constants and multiplicity reported in Table 2.....</i> | <i>32</i> |
| <i>Figure 11 - Trolox fluorescence evolution along time, at different concentrations: 200 µM, 100 µM, 50 µM, 25 µM and 12,5 µM. Fluorescence values were successively measured every 4 minutes in a microplate reader.....</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | <i>37</i> |
| <i>Figure 12 - Visual illustration of the method applied to calculate clean/net AUC of Trolox and every tomato extract sample, through the ORAC method.....</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | <i>37</i> |

Figure 13 - Example of a Trolox calibration curve used in one of the four independent assays, which enables the quantification of Trolox equivalent micromoles per gram of lyophilized tomato in each sample. .... 38

Figure 14 - Antioxidant activity measured by ORAC methods for extracts of tomato fruits from 1st to 5th harvest grown under different coloured plastics. Uncoloured standard plastic double-roof (UC), blue photo-selective plastic double-roof (BL); without double plastic covering (NP) and red photo-selective plastic double-roof (RD). Results are expressed as Trolox equivalent (TE,  $\mu\text{M}$ ) per gram of dried sample. Bar with \* mean the antioxidant activities is significantly higher compared with the other conditions ( $p < 0.05$ ). .... 38

Figure 15 - (a) PC2/PC3 PCA scores labelled by harvesting time; (b) loadings plots correspondent to the scatter plot in (a). Blue and red squares enclose the buckets which contribute more to the discrimination, for later and first harvests, respectively, and (c) PC1/PC2 PCA scores labelled by plastic covering obtained for  $^1\text{H}$  NMR data of a total of 120 tomato fruit  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (50:50, v/v) extracts covered with different coloured plastics and harvested at five different days (9/01/2020, 22/01/2020, 06/02/2020, 20/02/2020 and 02/03/2020 Almería, Spain). Scaling was done to Univariate. BL- blue plastic double-roof; NP- no double-roof; RED- red plastic double-roof; TR- uncoloured, standard plastic double-roof. .... 43

Figure 16 - (a) PLS scores (with  $R^2\text{X}=0.592$ ,  $Q^2= 0.92$ ), and (b) loadings plot obtained for  $^1\text{H}$  NMR data of a total of 120 tomato fruits  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (50:50, v/v) extracts covered with different coloured plastics and harvested at five different days (9/01/2020, 22/01/2020, 06/02/2020, 20/02/2020 and 02/03/2020 Almería, Spain). Scaling was done to Univariate. Blue and red dots in (b) are buckets that scored  $> 1$  in the VIP, for the first and last weeks of harvesting, respectively. .... 44

Figure 17 - (a) PLS-DA scores (with  $R^2\text{X}=0.828$ ,  $Q^2= 0.582$ ), and (b) OPLS-DA (with  $R^2\text{X}=0.934$ ,  $Q^2= 0.666$ ) obtained for  $^1\text{H}$  NMR data of a total of 120 tomato fruits  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (50:50, v/v) extracts covered with different coloured plastics and harvested at five different days (9/01/2020, 22/01/2020, 06/02/2020, 20/02/2020 and 02/03/2020 Almería, Spain). Scaling was done to Univariate for PLS-DA model and Pareto for OPLS-DA model. BL- blue plastic double-roof; NP- no double-roof; RED- red plastic double-roof; TR- uncoloured, standard plastic double-roof. .... 48

Figure 18- Heatmap showing z-scores of 53 discriminatory buckets from  $^1\text{H}$  NMR data of a total of 120 tomato fruits  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (50:50, v/v) extracts and their assigned metabolites for the four different experimental plastic coverings. The heatmap construction was based on Euclidean distances and Ward clustering. The X axis represents the average of thirty biological replicates for each plastic experiment (BL-blue plastic double-roof; NP-no plastic; RED-red plastic double-roof; TR-uncoloured, standard plastic double-roof) from all harvesting stages. The colour scheme represents the variation on the metabolic content among the four conditions in a scale from dark blue (lowest content) to dark red (highest content). In this case, spectroscopic data underwent variable size bucketing. Abbreviations: acet, acetate; aden, adenosine; ala, alanine; asc, ascorbate; asn, asparagine; aasp, aspartate; cit, citric acid; FA, fatty acids; flav, flavonoids; formic, formic acid; fru, fructose; fumaric, fumaric acid; GABA,  $\gamma$ -aminobutyric acid; glc, glucose; HCA, hydroxycinnamic acid derivatives; his, histidine; ile, isoleucine;

lac, lactate; leu, leucine; lys, lysine; mal, malic acid; MI, myo-inositol; phe, phenylalanine; pro, proline; sterols; suc, sucrose; thr, threonine; trig, trigonelline; trp, tryptophan; tyr, tyrosine; urid, uridine; val, valine. .... 49

Figure 19 - PLS-DA scores (with  $R^2X = 0.907$ ,  $Q^2 = 0.926$ ) obtained for  $^1\text{H}$  NMR data of a total of 24 tomato leaves  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (80:20, v/v) extracts covered with different coloured plastics from the same harvest (02/03/2020 Almería, Spain). Scaling was done to Univariate. BL- blue plastic double-roof; NP- no double-roof; RED- red plastic double-roof; TR- uncoloured, standard plastic double-roof..... 52

Figure 20 - PLS-DA contribution scores (with  $R^2X = 0.907$ ,  $Q^2 = 0.926$ ) obtained for  $^1\text{H}$  NMR data of a total of 24 tomato leaves  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (80:20, v/v) extracts, showing the loadings containing metabolites that increased (upper bars) and decrease (down bars) in leaves from single-roof covering respect to the other plastic-covers. Buckets from the flavonoid and hydroxycinnamic acid derivatives are highlighted, in dark and light blue, respectively. From left to right, in descendent order, these eight coloured buckets scored  $> 1$  in the VIP. .... 53

Figure 21 - 1D  $^1\text{H}$ -NMR spectra of tomato leaf samples zoomed in the phenolic region. From top to bottom, displayed spectra are from leaves collected in uncoloured double-roof greenhouses, without double-roof, red double-roof and blue double-roof. Spectra were obtained in a Bruker Advance III 600 MHz spectrometer and visualized in Bruker TopSpin 3.6.2. .... 55

Figure 22 - Average values of *Tuta absoluta* individuals captured in the chromatic trap for all plastic coverings. Uncoloured standard plastic double-roof (UC), blue photo-selective plastic double-roof (BL); without double plastic covering (NP) and red photo-selective plastic double-roof (RD). Insects were counted in 29/01/2020, 05/02/2020 and 12/02/2020. .... 56

Figure A1- 1D  $^1\text{H}$ -NMR spectra comparison of tomato fruit samples extracted with different solvents. Top: spectrum obtained with  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (80:20, v/v) fruit extracts. Bottom: spectrum obtained with  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (50:50, v/v) fruit extracts. Spectra were obtained in a Bruker Advance III 600 MHz spectrometer and visualized in Bruker TopSpin 3.6.2. .... 74

Figure A2 - 2D and 3D chemical structure of  $\gamma$ -Aminobutyric acid. Carbons mentioned in the text are indexed with a,b and c. Images obtained from: National Center for Biotechnology Information (2021). PubChem Compound Summary for CID 119, gamma-Aminobutyric acid. Retrieved January 31, 2021 from <https://pubchem.ncbi.nlm.nih.gov/compound/gamma-Aminobutyric-acid>. .... 75

Figure A3 - GABA three characteristic peaks identified in the 1D  $^1\text{H}$ -NMR spectrum of tomato  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (50:50, v/v) fruit extracts. Peak center at 3,01 ppm (t), 2,38 ppm (t) and 1,91 ppm (m) from top to bottom, respectively. Spectrum was analyzed in CHENOMX software. .... 77

Figure A4 - 2D  $^1\text{H}$ - $^1\text{H}$  TOCSY NMR experiment from tomato  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (50:50, v/v) fruit extracts. TOCSY can provide correlations between all the protons that make up a spin system. Red circles show the correlation between the three type of protons that give rise to the characteristic resonance signals of GABA. .... 78

Figure A5 - 2D  $^1\text{H}$ - $^{13}\text{C}$  HSQC NMR experiment from tomato  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (50:50, v/v) fruit extracts.  $^1\text{H}$ - $^{13}\text{C}$  Heteronuclear Single Quantum Coherence Spectroscopy (HSQC) shows which hydrogens

are directly attached to which carbon atoms. The  $^1\text{H}$  spectrum is shown on the horizontal axis and the  $^{13}\text{C}$  spectrum is shown on the vertical axis. The HSQC spectrum is most valuable when protons have already been assigned. Red arrows indicate the carbon signal attached to each proton that gives resonance in GABA. With a look at HMDB databases it can be confirmed these atoms belong to the molecule..... 79

Figure B1 - Trolox fluorescence evolution plotted along time, at different concentrations. Fluorescence values were successively measured every 4 minutes in a microplate reader..... 81

Figure B2 - Visual illustration of the method applied to calculate clean/net AUC of Trolox and every tomato extract sample, through the ORAC method..... 82

Figure B2 - Visual illustration of the method applied to calculate clean/net AUC of Trolox and every tomato extract sample, through..... 82

Figure B3 - Example of a Trolox calibration curve used in one of the four independent assays, which enables the quantification of Trolox equivalent micromoles per gram of lyophilized tomato in each sample. Table B 3 - Trolox solutions and blank AUC obtained for one independent assay and corresponding clean/net AUC obtained for the Trolox solutions. .... 83

Figure B3 - Example of a Trolox calibration curve used in one of the four independent assays, which enables the quantification of Trolox equivalent micromoles per gram of lyophilized tomato in each sample. .... 83

Figure C1 - Heatmap showing z-scores of 53 discriminatory buckets from CD3OD/D2O KH<sub>2</sub>PO<sub>4</sub> buffer (50:50, v/v) extracts of tomato fruits and their assigned metabolites for the four different experimental plastic coverings and each harvesting week. The heatmap construction was based on Euclidean distances and Ward clustering. The X axis represents the average of six biological replicates for each plastic experiment (BL-blue double-roof; NP-no plastic; RED-red double-roof; TR-uncoloured double-roof) from each harvesting stage. The colour scheme represents the variation on the metabolic content among the four conditions in a scale from dark blue (lowest content) to dark red (highest content). In this case, spectroscopic data underwent variable size bucketing. Abbreviations: acet, acetate; aden, adenosine; ala, alanine; asc, ascorbate; asn, asparagine; aasp, aspartate; cit, citric acid; FA, fatty acids; flav, flavonoids; formic, formic acid; fru, fructose; fumaric, fumaric acid; GABA,  $\gamma$ -aminobutyric acid; glc, glucose; HCA, hydroxycinnamic acid derivatives; his, histidine; ile, isoleucine; lac, lactate; leu, leucine; lys, lysine; mal, malic acid; MI, myo-inositol; phe, phenylalanine; pro, proline; sterols; suc, sucrose; thr, threonine; trig, trigonelline; trp, tryptophan; tyr, tyrosine; urid, uridine; val, valine. .... 86

# List of Tables

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |           |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| <i>Table 1 - Summary of NMR, MS advantages and disadvantages (Amberg et al., 2017; Emwas, 2015; Segers et al., 2019).....</i>                                                                                                                                                                                                                                                                                                                                                                                                                | <i>11</i> |
| <i>Table 2 - Peak assignment of metabolites identified in the tomato fruits CD<sub>3</sub>OD/D<sub>2</sub>O KH<sub>2</sub>PO<sub>4</sub> buffer (50:50, v/v) extracts by NMR .....</i>                                                                                                                                                                                                                                                                                                                                                       | <i>34</i> |
| <i>Table 3 - Average values for tomato fruit total production, marketable production and quality parameters, obtained for the experiment carried on the growth of Solanum lycopersicum L. cv Ventero, under single-roof and double-roof plastic coverings, developed during the autumn-winter season 2019-2020. Fruit weight - Fw; fruit equator diameter - Fd; fruit hardness - Fh; sucrose fraction - Sf; fruit acidity - pH and dry matter - Dm. Colorimetric coordinates: Lightness - L* and chromaticity a*/b*..</i>                    | <i>40</i> |
| <i>Table 4 - Average values of photosynthesis measurements carried out on the leaves and of the morphological parameters of plants cultivated under single-roof and double-roof with different plastic films. Photosynthetic Activity - PA; Photosynthetic Active Radiation - PAR. / Total stem length - Sl ; stem growth - ΔSg ; length of internodes below last physiologically mature leaf - Li ; stem diameter - Sd ; number of nodes below last physiologically mature leaf - Nn and length of the last leaf at maturity - Lf. ....</i> | <i>40</i> |
| <i>Table 5 - Harvest discriminating metabolites correspondent to the loadings plot in Figure 15b. These metabolites represent buckets that scored &gt; 1 in Variable Importance to Projection. ....</i>                                                                                                                                                                                                                                                                                                                                      | <i>45</i> |
| <i>Table 6 - Peak assignment of metabolites (limited to the aromatic region) identified in the tomato leaves CD<sub>3</sub>OD/D<sub>2</sub>O KH<sub>2</sub>PO<sub>4</sub> buffer (80:20, v/v) extracts by NMR. ....</i>                                                                                                                                                                                                                                                                                                                      | <i>52</i> |
| <i>Table B1 - Trolox and Blank fluorescence decay values along time. Fluorescence was measured in cycles of 4 minutes.....</i>                                                                                                                                                                                                                                                                                                                                                                                                               | <i>80</i> |
| <i>Table B2 - Tomato fruit extracts fluorescence decay values along time. Fluorescence was measured in cycles of 4 minutes.....</i>                                                                                                                                                                                                                                                                                                                                                                                                          | <i>81</i> |
| <i>Table B 3 - Trolox solutions and blank AUC obtained for one independent assay and corresponding clean/net AUC obtained for the Trolox solutions. <math>AUC_c = AUC_{sample} - AUC_{blank}</math> .....</i>                                                                                                                                                                                                                                                                                                                                | <i>82</i> |
| <i>Table B3 - Trolox solutions and blank AUC obtained for one independent assay and corresponding clean/net AUC obtained for the Trolox solutions. ....</i>                                                                                                                                                                                                                                                                                                                                                                                  | <i>83</i> |
| <i>Figure B3 - Example of a Trolox calibration curve used in one of the four independent assays, which enables the quantification of Trolox equivalent micromoles per gram of lyophilized tomato in each sample. Table B 3 - Trolox solutions and blank AUC obtained for one independent assay and corresponding clean/net AUC obtained for the Trolox solutions. ....</i>                                                                                                                                                                   | <i>83</i> |
| <i>Table B4 - AUC, clean/net AUC and Trolox equivalent (μM/ g of dried sample) obtained for every tomato fruit extract in one independent assay. ....</i>                                                                                                                                                                                                                                                                                                                                                                                    | <i>83</i> |

*Table C1 - Full list of harvest discriminating buckets and corresponding metabolites, indicated by the loadings plot relative to the PLS model (Figure 13) applied to the spectra data (X) with harvesting week (Y) as secondary observation. VIP- Variable Importance to Projection. .... 84*

# 1 Introduction

## 1.1 Almería as a leading horticultural producer

Nowadays, our society is facing a major challenge in achieving the food production and supply required to feed the growing world population. Although the rate of increase is low, it is estimated that the global population will stand between 8.5 to 8.6 billion in 2030 and between 9.4 and 10.1 billion in 2050, which represents a defying element towards a sustainable development (ONU, 2019).

Clearly, agriculture plays a key role in the food chain, and with the ongoing increases of population, per capita incomes and urbanization, its yield and productivity will have to attain unprecedented values (Hertel, 2011). The Food and Agriculture Organization (FAO) projected that the overall food production will have to raise around 70 percent to feed the world population by 2050, while another model forecasts that there will be a 100 to 110% increase in crop demand worldwide (Doering & Sorensen, 2018).

It is highly relevant that the future food policies encourage both sustainable food production as well as sustainable food consumption patterns, encouraging a healthy, varied and safe diet, instead of energy-dense and nutrient-poor foods. Moreover, cooperating in order to build locally-based and reliant food economies - where sustainable production, processing and distribution are integrated - is a strategy to improve economic, environmental and health aspects. To benefit agriculture and health, investment and food policies have to be well aligned with the know-how and knowledge gathered through scientific research by both public and private institutions (Kearney, 2010; Lainez et al., 2018).

The province of Almería, located in the east of Andalusia, exhibits the results of an integrated strategy based upon the local climatic characteristics towards socio-economic development. Almería has a semi-arid Mediterranean climate, where the agricultural system benefits from high solar radiation (3000-3600h of sunlight per year), mild temperatures mid-winter (with an annual temperature average of 18°C) and windy conditions, which decrease humidity (Reca et al., 2018). This favourable climate allows to fulfil the requirements of nowadays' consumers, and to obtain high-quality products with the lowest possible environmental consequences. In fact, Almería's greenhouses require 22 times less energy inputs compared to those of the Netherlands. Additionally, more kilograms of fruits and vegetables are produced per unit of energy with higher yields per hectare, in comparison with other, more technologically sophisticated, production systems (Valera et al., 2017). Indeed, in the last 50 years Almería went from being one of the poorest provinces in Spain to holding the largest concentration of

greenhouses in the world, Figure 1, leading to very significant social and economic benefits for the region (Cajamar, 2016). In 2018, there were 31,614 ha of protected cropping in Almería, with 86% being under biological control production, representing around 50% of the surface destined to greenhouses in all Spain (CADPR, 2018; Cajamar, 2016). This highly organized agro-industrial complex is clustered within a maximum radial distance of 65km which eases the connections of all commercial infrastructures, from harvesting to shipment. Season after season, their production and value are the central core of the city's economy. Annually, around 3.3 million of fruits and vegetables are produced in the province of Almería with a market value of 2.2€ billion. Around 70% of the agricultural production is sold in foreign markets, mainly in Germany, France, the Netherlands and the United Kingdom (Cajamar, 2016; Egea et al., 2018). It is not surprising that while, on average, agriculture contributes to a mere 2,5% of the gross domestic product (GDP) of Spain, in the case of Almería greenhouse production, it accounts for 13% of the local GDP. Moreover, the whole farming system economy accounts for 40% of the GDP of Almería (Castro et al., 2019).



Figure 1 - Satellite point of view from greenhouse concentration in the district of Campo de Dalías in the city of Almería. Image obtained from Google Maps (December 13, 2020).

Alongside pepper, tomato (*Solanum lycopersicum*) is the most produced vegetable in Almería's greenhouses. Tomato is of Andean origin, but its domestication seems to have occurred in Mexico. It is a multiannual herbaceous plant, although it is grown as an annual one, in short cycles of 4-5 months or 8-9 months long cycles. Even though tomato has a crawling stem, it is commonly grown on trellises, forcing its vertical growth. Its planting arrangement depends on the greenhouse type, crop cycle, soil and species variety (Cajamar, 2016). In the 2018/2019 campaign, 833 92 tons of tomatoes were produced, which translated into a market value of

58229€ million, although there has been a consecutive decrease in the production surface in Almería from 11,081 ha in 2013/2014 to 9,535 in 2018/2019 (Agrónoma, 2019). For the 2019/2020 campaign, a 8% downfall in the production was recorded (Fresh Plaza, 2020).

Tomato growing in Almería is struggling due to the increase in production costs, growing competition from north African countries and by the spread of pests such as *Tuta absoluta*, which altogether lead to a fall in revenue and profit that backs off the farmers' interest (Fresh Plaza, 2019). It is consumed more in the United States and in Europe, so there is a clear driving force for a country like Spain to reach high production numbers with top tier quality to achieve a big share in tomato exports to Europe's big countries (Nicola et al., 2009). In fact, tomato and tomato-derived products (canned tomatoes, sauce, paste and juice) fulfil the basic consumer's requirements, such as availability, suitability, variety, low cost and taste while it is a healthy food option.

## 1.2 Major throwbacks in tomato production

Undoubtedly, tomato crops as like other horticultural species are susceptible to a myriad of abiotic and biotic stresses, which result in yield-loss threatening damages (Mazzei et al., 2016). Phytophagous insects represent a huge problem in worldwide crop cultivation, and, nowadays, the economic and ecological effects of invasive species, mainly pests, is widely recognised. These cause yield reductions and significantly increase control measures' cost. Among pathogens, *Tuta absoluta* Meyrick (Lepidoptera: Gelechiidae), Figure 2, is a harmful, devastating threat with high propensity to use Solanaceae plants as secondary hosts, although it prefers feeding on tomato plants.



Figure 2 - Tomato leaf miner (*Tuta absoluta*) moth in adult phase (top), larvae (bottom left) and its deteriorating effect on tomato leaves and fruit. Images obtained from Hortidaily (2019) and Albida Agriculture (2018).

Native to South America, this pest was first detected in eastern Spain in 2006. *T. absoluta* spreads and colonises new areas rapidly, apparently without human help, as it has successfully invaded European and North African Mediterranean Basin countries, going suddenly from a South American tomato pest to a major threat to global tomato production. In fact, up to 2011, 21.5% of surface cultivated (0.95 M ha) and 27.2% of tomato production (41 M tons) are now a dwelling space to *T. absoluta*. Upon that time, a worrying forecast predicted that 84.9% and 87.4% of tomato cropped surface and production, respectively, were prone to the invasion of this pest in a near future (Desneux et al., 2011). In addition, the presence of this insect can cause yield losses from 80-100% if there are no control measures. Plants are attacked in every developmental stage and females commonly oviposit on leaves. After hatching, larvae penetrate the leaves, stems or tomato fruits, where they feed and grow, originating mines or galleries. In leaves, as the feeding larvae form mines within the mesophyll, the photosynthetic activity of the plant is considerably reduced, impacting growth and tomato yield. In stems, the galleries have an overall development impact that can lead to necrosis. In the tomato fruit, secondary pathogens can penetrate inside the galleries and cause fruit rot. Additionally, the insect often feeds the on growing tip, what hinders plant stable development. Clearly, the visual aspect of harvested fruits is not the desired one, what turns its commercialization impracticable. This way, and coupling it to the outstanding ease of spread and invasion, there is an imminent need for efficient, sustainable management agricultural practices such as the optimization of greenhouse structure (Desneux et al., 2010).

In South America, *T. absoluta* control was commonly performed by the repeated use of insecticide treatments, which lead to development of resistance. Indeed, undesirable effects, such as damage to the pest's natural enemies and other beneficial insects that provide biological control can arise, and chemical control is often too expensive and, in some cases, ineffective. Furthermore, pesticides' use has been reduced due to environmental and consumer constraints, which triggered the application of more environmentally friendly techniques, such as biological control with parasitoids and predators, light traps or mating disruption by pheromones (Cocco et al., 2013). Despite being a promising alternative, the importation of biological control applied in South America, which involves the introduction of exotic enemies, has unknown environmental risks associated with it. There is the possibility of attacking non-target organism and other deleterious unintended side-effects, what demands a baseline research in several fronts to make up its full potential (Desneux et al., 2010).

### 1.3 The introduction of greenhouses

From the beginning of mankind's domestication of plants, agriculture has been primarily conducted outdoors. It is a weather dependent industry with the most important factor driving horticultural crop productivity being its climacteric conditioning. Since greenhouses were introduced, protected cultivation, a specialised agricultural system where humidity, wind, temperature, water, nutrients, light intensity and atmospheric composition can be controlled, has been applied on a global level, providing a deep insight into the growth factors and inputs that improve crop productivity (Wittwer & Castilla, 1995). The aim of protected cultivation is to grow crops where naturally they could not survive, by modifying the environment to prolong the harvest period, commonly with earlier maturity, increasing yields and improving quality. Moreover, it enables year-round horticultural production, as well as the application of an integrated crop protection strategy to enhance pests and disease handling (Baudoin et al., 2013; Cástilla, 2012).

Greenhouses can be divided into two basic types: one that provides maximum climatic control in an environment that optimises productivity, and the other with minimal climatic control, allowing plant survival and producing an economical yield. This contrast is evident between the northern and southern European greenhouse industries. In the northern glass greenhouses of Netherlands, United Kingdom and Scandinavia, cutting edge technology is present in dynamic climate control systems, where several variables are 24h computer programmed. On the other hand, Mediterranean greenhouse agrosystems' trend is to adapt the plant to a suboptimal environment, and conversely these are characterized by the usage of simple plastic structures and a low level of technology, which creates a very strong dependence of the greenhouse microclimate in external conditions (Castilla & Hernandez, 2007).

In fact, over the last 30 years, there was a huge growth in protected horticulture and Mediterranean countries became increasingly competitive greenhouse vegetables' producers (Pardossi et al., 2004). During that time, a revolution in plastic greenhouse production technology has been occurring since the strategy to expose plants to suboptimal conditions instead of improving greenhouses to provide a favourable micro-climate consecutively leaves to suboptimal productivity, yield and low-quality vegetables in certain cropping seasons. Market competition and the increasing Europe-wide demand for high-quality vegetables required a better climate control. Indeed, well-designed and well-operated Mediterranean greenhouses secure a positive economic balance between the investment and operation costs and their agronomic performance (Castilla et al., 2004).

In the warm city of Almería, the growing period is often not year-round due to the absence of proper climate control. This area demands reducing excess temperature and improper air

humidity from late spring to summer and increasing inside temperature in winter (Castilla & Hernandez, 2007; Castilla et al., 2004).

With numerous greenhouses layout experiments and its study along the years, some preferred or recommended choices regarding structure and management are nowadays applied in Spanish greenhouses. Structurally, multi-span arched-roof is the prevalent industrial-type greenhouse design, being more airtight and easier to equip than the *parral* local-type greenhouse. In addition, roof slope has increased from nearly zero to roof angles around 30°, improving winter light transmission. Spatially, a E-W greenhouse is favourable with a N-S crop row orientation to optimize light uniformity. In terms of equipment, less than 1% of Almería's greenhouses have heating apparatus and double roofs are preferred in the winter season. Cooling is often performed with shading nets to limit solar irradiance, and decrease air temperature, when natural ventilation is sufficient, instead of using evaporating cooling devices. In fact, lack of proper ventilation is a typical problem in these greenhouses, a consequence of the limited venting area due to the lateral covering with insect-proof nets. Nowadays, greenhouses are characterized by larger openings, usually oriented to prevailing winds, in order to optimize this problem (Baudoin et al., 2013; Castilla & Montero, 2008).

One of the most common cladding materials in Mediterranean greenhouses is the low-density polyethylene (LDPE) and, by applying functional additives, special films can be synthesised, with the aim of achieving properties that benefit vegetable production. Of main interest are the photo-selective plastics, which selectively filter spectral components of the solar radiation such as UV and photosynthetic active radiation (PAR) and also transform direct light into scattered light, in order to induce desired physiological responses. (Cástilla, 2012).

## 1.4 Tomato organoleptic and nutritional value

Tomatoes are one of the most studied fruits, due to its high consumption and importance as component of the human diet. Compositionally, it has a precious nutritional and phytochemical profile as it is a low energy dense food that contains vitamin C, vitamin E, several carotenoids, soluble fibers, polyphenols and potassium. Raw tomato and processed products can provide a significant amount of antioxidants contributing to its free radical scavenging effects (Burton-Freeman & Reimers, 2011). Indeed, the epidemiological studies regarding tomato and lycopene intake, its main bioactive component and responsible for its colouring, granted them with disease prevention capabilities. There is a strong correlation between tomato intake and the reduced risk of cardiovascular disease, prostate cancer, osteoporosis, light-induced skin damages, among other disease conditions. Due to its widespread high consumption, there is an

overall interest in its metabolic study in order to provide nutritionally rich and tasteful tomatoes with an efficient production method (Cameron & Geana, 2005).

The functional quality and phytochemical content of tomatoes are substantially influenced by its genotype's interaction with environmental factors derived from the agricultural practices. Tomato fruit chemical composition is a function of its cultivar, climatic and environmental conditions (light, humidity, CO<sub>2</sub>, air pollutants, temperature), agronomic practices (water quality, soil properties, salinity, pruning, growing systems), stage of ripeness at harvest, and all post-harvest handling till it reaches the consumer. The understanding of preharvest factors is of great importance to improve tomato organoleptic and nutritional properties which directly affect postharvest quality. Clearly, postharvest handling is also important to preserve the fruit compositional status or at least to minimize its deterioration in order to fulfil consumers' demands (Shimeles et al., 2017).

The fruit development stage upon its harvest is a key aspect regarding its nutritional quality and taste, as during ripening several biochemical processes alter the tomato composition leading to changes in colour, flavour and nutritional properties (Dorais et al., 2008). The onset and progression of ripening in tomatoes is typically associated with changes in the external colour of the pericarp, reflecting the accumulation of antioxidant carotenoids and flavonoid pigments (Shinozaki et al., 2018).

Lycopene is the predominant lipid soluble compound, which is responsible for the deep red colour of ripe tomatoes and makes up to more than 80% of total carotenoids in fully ripe red fruits.  $\beta$ -carotene content is usually around 7% of carotenoids (Shimeles et al., 2017). Due to its molecular structure, symmetrical linear chain with 11 double bonds, lycopene is the most powerful antioxidant among the carotenoids thus highly reactive with singlet oxygens and peroxy radicals, scavenging these. The protective and beneficial health effects associated with tomato and tomato processed products consumption is primarily attributed to its high content in lycopene, which biosynthesis is influenced by temperature. In fact, temperatures above 30°C lead to a decrease in lycopene while increasing other carotenoids, resulting in yellow to orange tomatoes (Brandt et al., 2006)

Fruits' cell walls act as protective barriers against predators and environmental stress, allowing proper cell functioning. Enzymes synthesized during maturation disrupt cell wall glycosidic links of polysaccharides, hemicellulose and pectin. Eventually, the controlled weakening of the polymeric networks of the cell wall leads to a decrease in fruit firmness and increase in water content, resulting in a softer appealing tomato (Chylińska et al., 2017).

Besides physical characteristics as firmness and colour, another predominant aspect in tomato fruit quality is its perceived flavour. Nowadays, it is recognized that intensive breeding has

favoured crops yield which negatively impacted the fruits and vegetable flavour and nutritional content. Due to it, effort has been placed in understanding the genetic and chemistry underlying tomato sensorial stimulus in order to improve its quality and achieve consumer satisfaction (Klee & Tieman, 2013; Tieman et al., 2012).

An optimal ratio between sugars and acids, also viewed as sweetness/acidity is fundamental for tomatoes' taste. Additionally, a small set of volatile compounds really determines their unique flavour and aroma. In fact, a group of 20-30 volatile chemicals are usually present in a quantity capable of stimulating the human olfactory system (Klee & Giovannoni, 2011). In the end, it is the balance between sugars, organic acids and volatiles that determines the tomato fruit flavour.

Overall, there is a complex and very dynamic metabolism in the tomato's development, where multiple genetic and biochemical pathways lead to the appreciated organoleptic and nutritional value of a fully ripened fruit. In terms of primary metabolism, malic and citric acid decrease, while fructose and glucose accumulate from the breaker to the deep red stage. Secondary metabolites, such as total ascorbate, lycopene,  $\beta$ -carotene, flavonoids and phenylpropanoids reach maximum levels in the deep red tomatoes. Moreover, there is an increase in amino acid content and redox activity (Gautier et al., 2008).

## 1.5 Metabolomics and its suitability to study agronomical practices

Plants as sessile organisms evolved in a very peculiar and overwhelming fashion to achieve every living organism basic objective with great success, its survival and continuous reproduction. In fact, predators' attacks and environmental stress gave rise to an impressively dynamic and diverse plant biological system with a vast complex biochemical network, representative of the phenotype outcome of the genomic background, transcript expression and protein function (Fang et al., 2019). For hundreds of years, humanity sought the understanding of plants' metabolic pathways biochemistry and regulation. However only in the last two decades, with the outbreak of the Omic Sciences combined with the development and refinement of analytical and data processing technologies, it was possible to gather detailed information regarding an organism's metabolic status which significantly enhanced the knowledge of plants functioning (Hall, 2006).

In the beginning of this millennium, the international Human Genome Project accomplished its main goal with the entire sequencing of the human genome. This achievement gave birth to the genomic era or omic sciences which considerably improved the understanding of biology in a super quick unprecedented manner (Guttmacher & Collins, 2003). Nowadays, with the

available technology and advanced equipment, it is possible to generate massive amounts of data that surpass human analysis capabilities. In order to interpret them, powerful computational and statistical techniques are applied (Dra & Ruiz, 2015).

In the first place, genomics is the genetics field responsible for figuring out the content, organization, function and evolution of the DNAs molecular information embodied in the complete genome (Lesk, 2017). The totality of transcripts in a given cell at a specific physiological state constitutes its transcriptome, which is the template for protein synthesis. Transcriptomics main focus is labelling each species of transcripts, mRNAs, small RNAs, non-coding RNAs in order to understand the genome output under different conditions (Chambers et al., 2019; Z. Wang et al., 2009). Proteins play a key role in any living organism as they are crucial components in every cellular metabolic route. Their large-scale study is named Proteomics, and as these are much closer to functionality than genes, this branch of omic sciences frequently leads to scientific hypothesis or discoveries (Pandey & Mann, 2000).

In the same line as the previous omic terms, the set of small molecular metabolites synthesized by an organism constitutes its metabolome. Soughting the biochemical functioning of a biological system through gene and protein content can be misleading as these are submitted to epigenetic changes and post-translational modifications, respectively. Conversely, metabolites are the final output of the biological complex so these can provide a clearer mirroring of the biochemical outcomes that arise from the variables under study. Essentially, metabolomics is commonly recognized as the closest omic science to the phenotype (Patti et al., 2012).

Plants are commonly termed cell factories as a single species can contain from 5,000 to tens of thousands of distinct compounds in their metabolome. In addition to its enormous quantity, these molecular entities are very diverse chemically and divergent in terms of concentration, which turns the complete metabolome characterization a very challenging task (Wang et al., 2019). In fact, plant core primary metabolism is highly similar to that of non-plant species, however it's their additional secondary metabolism or specialized metabolites which represents the majority of the metabolome. These are not directly involved in growth and reproduction but they play a key role as an effective mean to resist adverse environmental conditions, which can either be biotic or abiotic stresses. (Fang et al., 2019; Singh et al., 2020). Indeed, specific metabolic routes are activated or inhibited relative to the set of conditions which are influencing the life-cycle, from planting to harvesting (Singh et al., 2020).

Metabolomics can be seen as the ID or profile status of an organism, providing a non-biased identification and quantification of the metabolites present in a given matrix at a given time. This way, plant metabolic profiling of primary and secondary metabolites is raising interest due

to its suitability to access, between many others, fruit production methods and correspondent quality with direct relation to their visual characteristics, the phenotype (Ghatak et al., 2018).

Applied to food production, metabolomic studies can provide compositional information pointing towards organoleptic and nutritional properties which assesses the crop quality and its commercial value (Hall, 2006).

A metabolomic study can be applied either with a targeted or untargeted approach. The targeted approach is usually hypothesis-driven and aims at the quantification of specific, already known, metabolites in biological samples. On the other side, untargeted studies pursue the generation of a hypothesis through the qualitative identification and relative quantification of the global metabolic profile between biological samples (Schrimpe-Rutledge et al., 2016).

In its early steps, metabolomics revealed huge and broad potential in agricultural and food sciences rendering safer, healthier and nutritive food through the application of powerful analytical tools such as nuclear magnetic resonance (NMR) and mass spectrometry (MS) alongside bioinformatic software (Dixon et al., 2006). Indeed, the successive increase in the last decade of publicised studies regarding metabolomics application in food sciences and agriculture, Figure 4, exhibits the value recognized by the scientific community to the metabolomic approach.

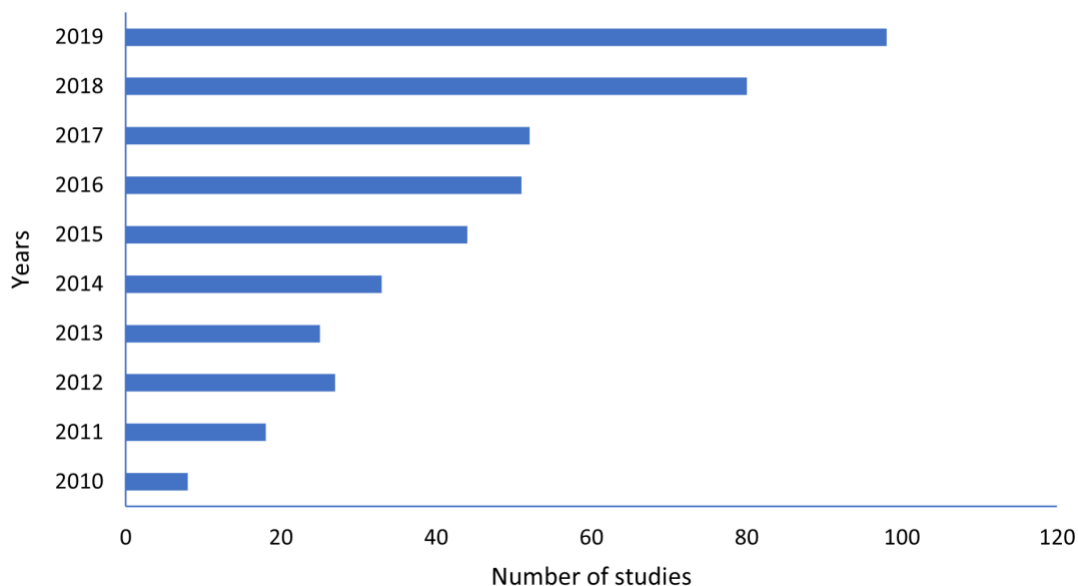


*Figure 3 - Evolution of publications on metabolomics applied to agricultural and food sciences from 2010 to 2019. The data was generated from an advanced research in PubMed database with the following keywords: ("Food Technology"[Mesh]) OR "Agriculture"[Mesh]) AND "Metabolomics"[Mesh]). There is an official list of Medical subject headings (MeSH), with which the indexers label an article regarding its content. Applying MeSH terms can lead to a more efficient research, as the resulting articles will be necessarily about the suggested topic instead of just mentioning it (Baumann, 2016).*

In recent years, NMR spectroscopy has been clearly contributing to the boom of multivariate statistical data analysis approach in order to mine the extensive information of foods chemical profile. Along with mass spectrometry, these are the pair of modern analytical techniques being the workforce and what defines modern food metabolomics (Ralli et al., 2018).

The most widely used technologies in plant metabolomic studies are NMR, gas chromatography-mass spectrometry (GC-MS) and liquid chromatography-mass spectrometry (LC-MS). The equipment selection is based on its sensitivity, selectivity, accuracy and speed, related with the class of compounds, their concentration and physio-chemical properties, and the available amount of biological sample under study (Ghatak et al., 2018). Table 1 summarizes the attributes and limitations of NMR and MS analytical platforms.

Table 1 - Summary of NMR, MS advantages and disadvantages (Amberg et al., 2017; Emwas, 2015; Segers et al., 2019)

| Property             | NMR                                                                     | LC-MS                                                                   | GC-MS                                                              |
|----------------------|-------------------------------------------------------------------------|-------------------------------------------------------------------------|--------------------------------------------------------------------|
| Sample preparation   | Minimal sample preparation (simple extraction)                          | Intermedium. Chromatography and ionization.                             | Extensive. Chromatography, ionization and derivatization required. |
| Sample analysis      | Fast                                                                    | Longer                                                                  |                                                                    |
| Sample recovery      | Non-destructive. Several analysis can be carried on the same sample     | Destructive. However it needs a smaller amount of sample                |                                                                    |
| Sensitivity          | Poor                                                                    | High. Metabolites with nanomolar concentrations can be readily detected |                                                                    |
| Reproducibility      | Fundamental advantage                                                   | Lower                                                                   |                                                                    |
| Resolution           | Lower                                                                   | Higher                                                                  |                                                                    |
| Quantification       | Inherent. Signal intensity proportional to the metabolite concentration | Internal standard needed. Ionization efficiency is a determining factor |                                                                    |
| Identification       | Easy                                                                    | Databases must be improved                                              | Easy (spectral libraries)                                          |
| Range of metabolites | Polar and nonpolar                                                      | Polar (HILIC)                                                           | Volatile polar and nonpolar                                        |
|                      |                                                                         | Non polar (RPLC)                                                        |                                                                    |

HILIC - hydrophilic interaction liquid chromatography; RPLC - reversed phase liquid chromatography.

Indeed, NMR has a lower sensitivity than MS hindering the possibility to detect metabolites present in very small quantities. Moreover, MS also possesses a higher resolution than NMR. On the other hand, MS only detects compounds that readily ionize or are derivatized, which can be challenging sometimes. The incoherent signals caused by ion-suppression and matrix effects are another drawback. MS usually has a time consuming previous chromatographic separation that can induce biological variations in the metabolome. The advantages of using NMR for metabolomics include a relatively high-throughput, non-destructive data acquisition, minimal sample handling, simple methods for metabolite quantitation, and redundant spectral information to improve the accuracy of metabolite identification (Markley et al., 2017; Segers et al., 2019).

Due to its capability of accessing a broad range of metabolites in a single spectrum aligned with the possibility to elucidate unknown compounds, NMR is commonly applied in untargeted fingerprinting metabolomic studies. Conversely, MS methods are usually applied in targeted studies as a result of their powerful sensitivity, and since the aiming metabolites are already known, it is possible to adjust sample preparation in order to minimize interferences (Emwas, 2015).

In fact, it is noted that the coupling of analytical techniques has been of growing interest as the number of studies combining NMR and MS is increasing. A more elaborated coverage of the metabolome is possible by combining the two platforms as the identification of unknown metabolites will be enhanced as well as the accuracy on identifying known ones since the assessment will be based on multiple and confirmatory evidence (Marshall & Powers, 2017).

Effectively, the low sensitivity and resolution featured by NMR alongside its limited accessibility account for the preferred choice of MS to carry metabolomic studies in the past. However, recent technological improvements are reinforcing the key role of NMR in untargeted metabolomic profiling in broad subjects relative to plant sciences. These include advances in sensitivity due to the application of higher magnetic fields, better probes (cryoprobes, micro and in-glow probes) and hyperpolarization. In order to yield more accurate data by solving overlapping peaks, multi-pulse, multi-nuclear and multi-dimensional NMR are promising solutions. Lately, to challenge accessibility focus has been given to the development of benchtop NMR hardware (Giraudeau, 2020).

## 1.6 NMR principles and features

Magnetic properties of the atomic nuclei are the physical basis that engage NMR spectroscopy. The theory behind NMR and the dynamics of nuclear spin systems are thoroughly understood. The interaction between the nuclear magnetic moment and an external magnetic field leads to the nucleus' energy transition to another stationary state, which is detected and recorded as a spectral line. The magnetic energy of the nucleus is confined to certain values and molecules containing atoms with non-zero magnetic moment provide spectral information.

Three principal aspects turn NMR into a powerful broadly applied tool to quantify and determine the structure of unknown compounds or to confirm their presence in samples with known composition. The protons of a given molecule give rise to groups of separated resonance signals, and this spacing is called chemical-shift. This fact provides information regarding the different chemical environments surrounding each nucleus. In addition, the area under each signal is linear to the number of protons that originated the signal. Its integration directly gives

information about the number of nuclei present in the sample. Also, some protons form split signals, such as doublets, triplets and multiplets, due to the spin-spin coupling, a magnetic interaction between nuclei. The magnitude of this interaction (J-coupling constant) gives detail about their positional relation as it depends on the type and number of bonds separating them (Gunther, 1973).

A spectrometer is composed of several elements and the highlighted ones are the magnet and the probe. The magnet produces a stable, homogeneous and intense magnetic field to polarize the spins. The higher the magnetic field, the better resolution and sensitivity as the resonances will be better separated and the greater the thermodynamic preference for nuclear spins to align with the magnetic field. However, this equipment is very expensive and also requires costly maintenance, so usually 600 MHz spectrometers are available in research institutions as these provide a good balance between resolution, sensitivity and cost. The probe is at the centre of the magnetic field, where the sample is inserted. It has coils that bombard radiofrequency signals, tuned at specific frequencies regarding the specific nucleus and magnetic field. The nuclei are excited and emit radiofrequency weak signals conducted by the probe to a console that amplifies, detects and stores them. Besides this, the probe also controls the sample temperature and is perhaps the most delicate part of the spectrometer (Emwas, 2015; Mukhopadhyay, 2007).

Basically, the sample is placed in the centre of a static magnetic field and is irradiated with pulse sequences (high frequency). The samples' response is then detected by an induced current. The amplitude of the NMR response is linear to the concentration of metabolites, providing a relatively easy quantification of compounds for metabolic profiling (Amberg et al., 2017).

Most applications of NMR for metabolomic studies rely on one-dimensional (1D) NMR experiments based on the  $^1\text{H}$  nucleus which offers the highest detection sensitivity, Figure 4. This technique enables the identification and quantification of up to 150 metabolites in a biological sample. In addition, metabolomics applied with other nuclei are not so frequent, as these exhibit lower sensitivity along with low natural abundance of the elements (Serkova et al., 2019).

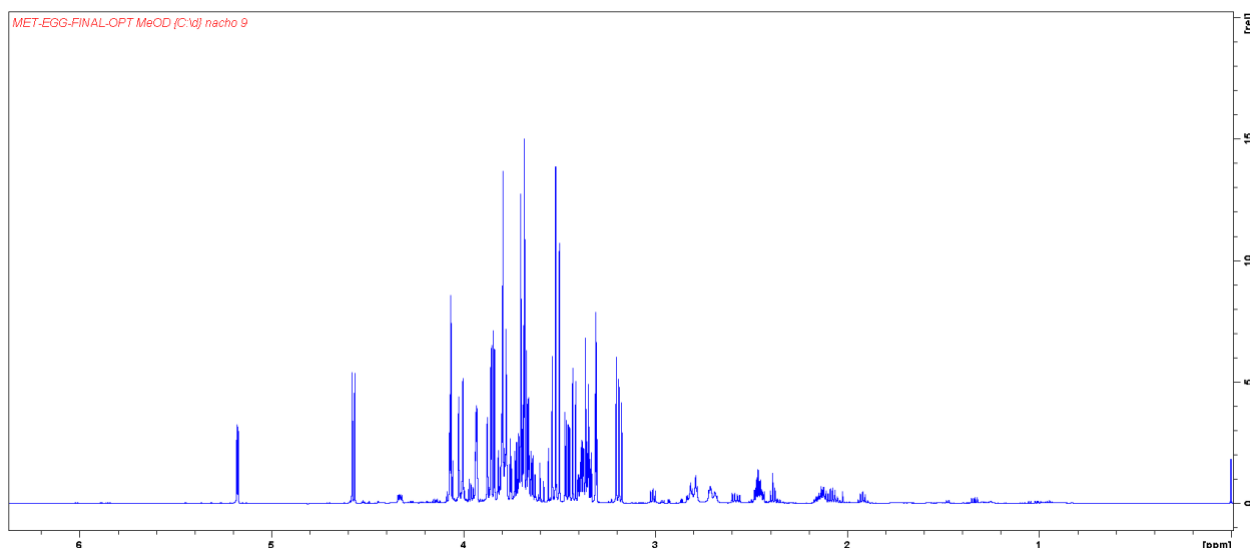


Figure 4 - 1D  $^1\text{H}$ -NMR spectra of a tomato fruit sample. Spectra was obtained in a Bruker Avance III 600 MHz spectrometer and visualized using the Bruker TopSpin (3.6.2) software.

Metabolic samples as food matrices are marked with a big difference in concentrations between water (usually one fraction of the solvent) and the compounds of interest. The water molecule will give rise to a signal in the spectrum that highly contributes to inaccurate metabolite peak areas. Indeed, the technique applied to suppress the water signal is one of the elements that most influences the overall quality of the end spectrum (Giraudeau et al., 2015).

As mentioned before, fruit composition is highly complex because of the vast presence of metabolites with diversified chemical structures, mainly the ones belonging to the secondary metabolism. As a result, fruit extracts spectra obtained with 1D  $^1\text{H}$ -NMR exhibit hundreds of multiplet resonances, and due to the small range dimension of  $^1\text{H}$  (0-10 ppm), it is showcased as an overcrowded spectrum with stacked peaks that hinders the compound's assignment. In order to avoid misleading extrapolations from 1D  $^1\text{H}$ -NMR and enable clear metabolite assignments, two-dimensional (2D) NMR experiments are inevitable followed: homonuclear correlations like  $^1\text{H}$ - $^1\text{H}$  correlation spectroscopy (COSY) and  $^1\text{H}$ - $^1\text{H}$  total correlation spectroscopy (TOCSY); or heteronuclear correlations like  $^1\text{H}$ - $^{13}\text{C}$  heteronuclear single-quantum correlation spectroscopy (HSQC) and  $^1\text{H}$ - $^{13}\text{C}$  heteronuclear multiple-bond correlation spectroscopy (HMBC) (Sobolev et al., 2015).

First,  $^1\text{H}$ - $^1\text{H}$  COSY experiment shows the frequencies of hydrogen along both axes and is used to identify spins directly coupled to each other. When there is a correlation between the signals of the spectrum along both axes, a cross peak is formed, indicating that those nuclei with different chemical shifts are coupled. The  $^1\text{H}$ - $^1\text{H}$  TOCSY experiment is similar as it exhibits the cross peaks of coupled protons. However, the chemical shift of a H is correlated with the

chemical shift of other Hs that are connected in the unbroken coupling chain. As an example, if proton X is coupled with proton Y and proton Y is coupled with Z, COSY spectrum would mainly exhibit the coupling between X and Y, while TOCSY spectrum could display the link of X with Y and Z. Shortly, COSY exposes protons separated up to 2/3 bonds and TOCSY is able to stretch the cross-peaks to protons of the same spin system involving 2/3/4/5 or more bonds. The tuning of the spin-lock time allowed us to see in this specific experiment all these connections.

In addition, heteronuclear experiments detect correlations between different types of nuclei. Just like the homonuclear runs,  $^1\text{H}$ - $^{13}\text{C}$  HSQC spectrum showcases the interaction between a proton and its closest carbon. This method will provide one peak per pair of coupled nuclei and its coordinates represent the chemical shifts of the linked atoms. On the other hand,  $^1\text{H}$ - $^{13}\text{C}$  HMBC technique is able to correlate the chemical shift of a proton with a carbon separated by two or more chemical bonds (Emwas, 2015).

The use of 2D NMR experiments is a key strategy to achieve an unambiguous identification of compounds in mixtures, although it has its own drawbacks. In order to produce the second dimension, 1D spectra are recorded with incremented time evolution periods. As a result, 2D experiments can last up to several hours. The total experiment duration is multiplied by an n-factor which represents the number of points necessary to provide an additional dimension with good resolution, usually several from 160 to 256 (Marchand et al., 2017).

## 1.7 Experimental steps in a metabolomic project

Certainly, a metabolomic study can easily go bust in case it is not thought and planned with caution and consequently properly designed. To provide an accurate response to the proposed biological question it is crucial the construction of an experimental design that guarantees a reliable statistical analysis of the collected data derived from the biological sample (Rodrigues et al., 2019).

First, it is essential a meticulous experimental design. A metabolomic experiment is prone to several sources of variability, either biological or technical, and commonly it is not possible to perform experimental adjustments to avoid their influence. However, whenever possible, biological replication should be taken instead of technical replication as technical variability is usually much smaller (Gibon & Rolin, 2012). Sample size, sample pooling and number of replicates (biological and technical) are key parameters that deeply impact data robustness and suitability for the statistical (multivariate) analysis. In addition, the sample or group size needed to perform the experiment is related to the biological variability of the system under

study compared with the analytical variability of the analytical platform (Barnes et al., 2016). In fact, the larger the number of the samples, the clearer is the information, which translates into an easier identification of patterns or markers for the hypothesis under study (Kim et al., 2011).

After a proper experimental design, the H-NMR metabolomic experimental procedure can be split into four major steps: sample and extract preparation; spectra acquisition and processing; data processing and lately, data interpretation. In Figure 5, the workflow of a plant-based metabolic study is schematized.

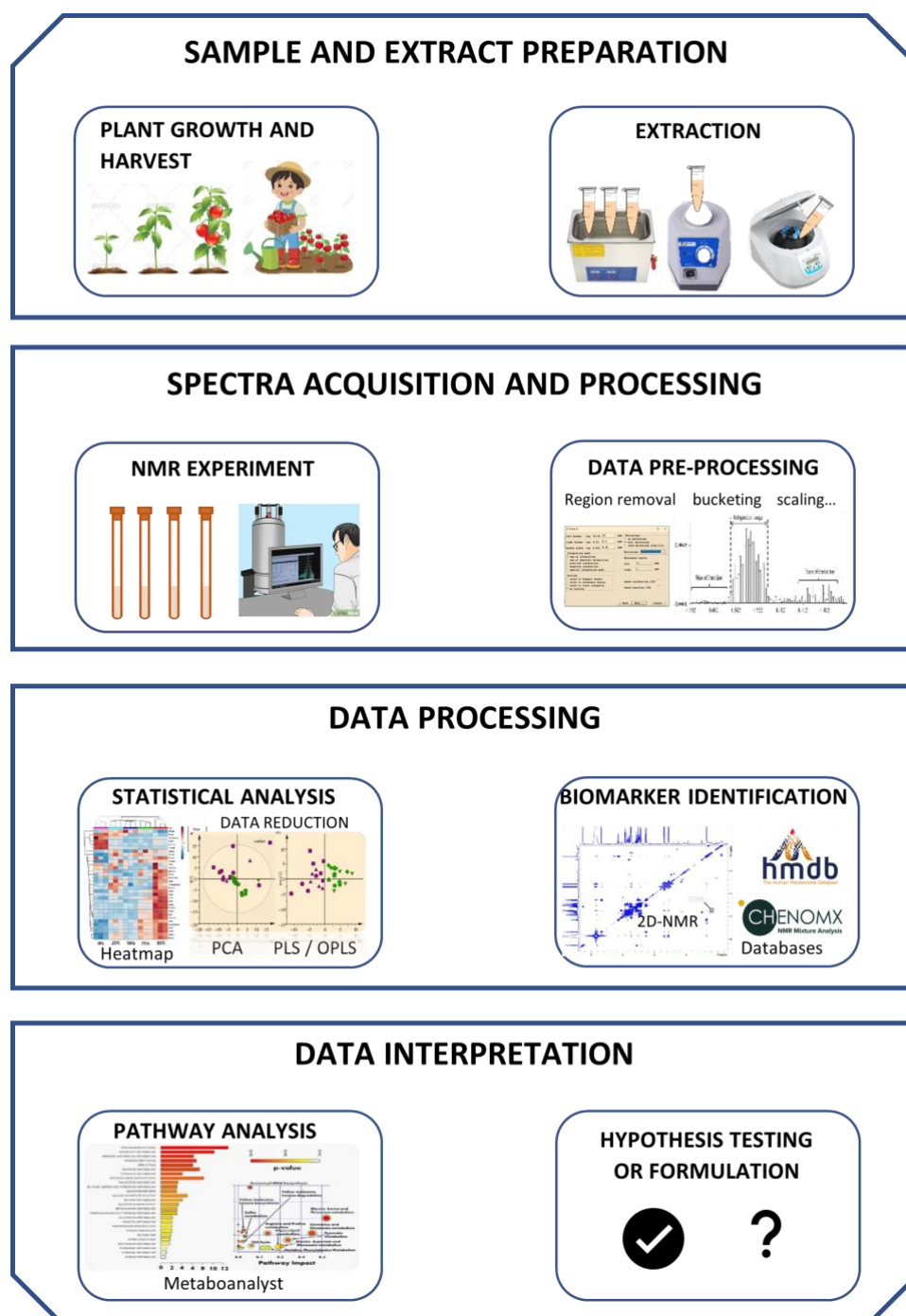


Figure 5 - Scheme of the successive steps carried out on a metabolomic study applied to plant sciences study.

## Sample and extract preparation

The preparation step involves transforming the sample into a solution that can be analysed, starting with the harvest of biological material and all the way to the extraction of its metabolites. Sample harvesting, handling and storage, likewise the quality of the extraction are critical steps which strongly impact the validity of the biological interpretations from the metabolomic results. Sample preparation has to be carried aiming the maximum avoidance of biological irrelevant changes (Kim et al., 2011; Salem et al., 2020). Instantaneously, at harvesting the plant is wounded, cell integrity is lost, and new metabolites appear. The metabolism must be stopped to avoid further biological and physio-chemical changes. Therefore, collection followed by immediate quenching (liquid nitrogen of preferred choice) of the metabolism is strictly necessary (Biais et al., 2011; Rodrigues et al., 2019).

A simple and minimalist sample preparation protocol has its advantages in terms of speed, feasibility, consistency, reproducibility and efficiency. Cryogrinding of the biological sample for plant cell-wall breakage is of particular importance, since powder facilitates the freeze-drying process and consequent extraction. Indeed, water removal by freeze-drying is indispensable as it is the main component of plant tissue and its two protons are detected by NMR (Martins et al., 2018).

Regarding the metabolite extraction protocol, its choice must account for compatibility with the analytical technique and it is another impactful parameter as it will influence metabolite coverage and concentration. In targeted analysis the extraction choice will favour the fraction enrichment of compounds with similar properties while in an untargeted approach it is pretended a solvent that allows a wide coverage of the metabolome. In fact, in the latter, bias results are inevitable as each extraction method has its inherent selectivity. Due to the vast extension of plant metabolome, characterized by compounds with highly distinct chemical nature and solubility ranges, a single extraction procedure will never enable the extraction of the whole range of metabolites (Martins et al., 2018).

For instance, in tomato NMR experiments a methanol-water extraction is suited for detection of polar/semi-polar metabolites such as organic acids, carbohydrates, amino acids, quaternary ammonium compounds and hydroxycinnamic acids, while the presence of chloroform enables the extraction of polar/non-polar metabolites as carotenoids and fatty acids (Abreu & Fernández, 2020).

In fact, methanol-water mixtures have been successfully applied in untargeted tomato metabolomic studies since they render the extraction of a wide range of metabolites (Sauerschnig et al., 2018).

Variations in pH and ionic force between aqueous samples can strongly influence the chemical shifts of metabolites with acidic/basic functional groups such as organic acids and amino acids rendering slightly different spectra, so it is important to use a buffer.

### Spectra acquisition and processing

Regarding NMR acquisition of plant-based samples, that usually exhibit high content in water, it is of main importance the water signal suppression. The pulse-sequence of preferred choice in NMR metabolomics is the  $^1\text{H}$ -nuclear Overhauser effect spectroscopy (NOESY) combined with a continuous-wave presaturation module (Mckay, 2011). On the other hand, for the polar metabolites profiling in samples without water, the standard one-pulse and detection of the FID-based sequence is preferred (Allwood et al., 2011).

In addition, internal reference compounds such as 3-(trimethylsilyl)-2,2',3,3'-tetradeuteriopropionic acid (TSP) and 2,2-dimethyl-2-silapentane-5-sulfonic acid (DDS), or their sodium salts, are added for aqueous media, whereas tetrakis(trimethylsilane)silane (TTMS) is usually employed for organic solvents (Beckonert et al., 2007). As these compounds are added in known concentrations, it is possible to quantify the metabolites present in the sample by peak area comparison.

After spectra acquisition, it is transformed from raw spectra in time domain to spectra in the frequency domain by means of the Fourier Transform. Then, the phase is usually automatedly corrected to obtain an absorption line shape, nevertheless human inspection is necessary to avoid artifacts from automated processing (Deborde et al., 2019). Before the statistical analysis, raw data also undergoes a several step processing to provide related sets of metabolites suited for a thoroughly hypothesis settling. These encompass noise elimination, baseline correction, alignment, data reduction and normalization. It is pretended the removal of differences between spectra that arise from experimental and instrumental methods as well as the achievement of a coincident spectra segmentation to render maximal correlation. In fact, undesired spectral regions (water and solvent signals) are removed, bucketing of regular bin size (0.01, 0.02 and 0.04 ppm) or variable-sized bin and further scaling (mean-centering, autoscaling, pareto scaling among others) to avoid influence of biological insignificant abundant metabolites are performed (Allwood et al., 2011; Deborde et al., 2017).

### Data processing / Multivariate data analysis

Metabolomic analyses provide huge and extensive datasets, due to the hundreds of samples and signals of each spectrum. A strategy usually employed to shorten a large number of analytes to

those that have a bigger response under the studied conditions is a univariate analysis. However, it cannot separate groups based on single-molecule level detail (Bartel et al., 2013). The methods employed to deliver accurate measurement of experiments with the most relevant chemical information come from chemometrics and informatics disciplines which are based on mathematical and statistical strategies to perform pattern recognition and multivariate data analysis (Amberg et al., 2017; Héberger, 2008). The file containing the bucketing/binning data is inserted in software to undergo the multivariate data analysis. These methods go further and capture group similarities or decorrelations with the information of individual molecules. Various multivariate statistical methods are available nowadays, from unsupervised approaches like Principal Component Analysis (PCA) to supervised approaches as Partial Least Squares (PLS) and Orthogonal Projection to Latent Structures (OPLS).

In fact, PCA is the go-to technique for an initial hypothesis and a free, unbiased overview. Without previous metabolite class acknowledgement, this method cluster samples according to their intrinsic similarity or deviance. Sets of metabolites importance, named Principal Components (PC), are displayed orthogonally to each other as a function of the variance that decorrelates them. PCA is of preferred choice when the objective is to obtain group discrimination, clustering and dimension reduction.

On the other hand, PLS supervised method is applied to two matrices, providing the influence of independent variables (Y) on the dependent one (X). X matrix contains the spectra processed data while the Y matrix comprises additional class information, non-spectra related, from the analysed samples. This way, elements responsible for the predefined classification can be identified. Moreover, OPLS is an extension of PLS that through the contribution of variance filters irrelevant data, improving model interpretation by clarifying which are the meaningful variables that give rise or defy the pre-set group class. (Bartel et al., 2013; Dona et al., 2016).

Either unsupervised and supervised methods generate loading plots from where chemical and biological information, such as the identities of metabolites, that justify the apparent biological discrimination, can be deduced. In fact, it is possible through combining 1D NMR quantification with the advantage of precise compound identification derived from 2D NMR. Of indispensable character are public and commercial databases such as Human Metabolome Database (HMDB) and the Chenomx package which contain 2D spectra and structural information, enabling signals comparison to obtain accurate assessments relative to the key contributors for verified class differentiations (Giraudeau, 2020; Kim et al., 2011).

## Data interpretation

A well-executed metabolomic experiment, regardless of its approach, enables gathering accurate, consistent and reproducible data by the quantification and identification of metabolites of a given system. Although it is mainly a portrait of the metabolome, if broadly and frequently applied it can build up huge databases, vast in information, relative to the biological behaviour following distinct treatments during development what turns possible exploiting the biochemistry underlying the expressed phenotype differences. This way, it also leads to the creation of a biological hypothesis, ready to be tested, opening ways to scientific questioning or to the breakthrough of a new paradigm (Hall, 2006; Patti et al., 2012).

Processing, integrating and visualizing modern biological datasets is a formidable task that requires a flexible computational framework to enable a growing variety of genomic, biochemical and phenotypic data types. MetaboAnalyst is an online web tool particularly useful for comprehensive metabolomic data analysis, interpretation, visualization and integration with remaining omics data, which is constantly evolving to meet bioinformatic demand pulled by the metabolomic community, what sets it as the most complete analysis web resource in terms of scope and applications.

Ultimately, the outcome of the implemented data statistical analysis is commonly the precursor for the functional interpretation of metabolomic data, which requires metabolites identities and concentrations. In untargeted approaches, spectra features must be tailored in order to provide a list of compounds suitable for functional analysis while in targeted approaches, data usually requires less processing (Xia & Wishart, 2016). In fact, MetaboAnalyst offers two approaches in functional analysis: Metabolite set enrichment analysis (MSEA) and metabolic pathway analysis (MetPa), with both performing a comparison of the significant compounds identified in uploaded data to the previously settled groups. Moreover, another module is an integrated pathway analysis that correlates genetic experiments with metabolomics, only possible due to the Kyoto Encyclopaedia of Genes and Genomes (KEGG) database. KEGG metabolic pathways enable over-representation and pathway topology analysis that deepen metabolomic interpretation through genetic information. Once again, the construction and improvement of databases is fundamental for successful metabolomic studies as their data interpretation is performed by informatic tools that would be useless without the support from the major databases such as HMDB, KEGG, PubChem, Chemical Entities of Biological Interest (ChEBI) and METLIN (Chong et al., 2019; Xia et al., 2015).

## 1.8 Examples of the application of NMR-based metabolomics in the agricultural field

To conclude and demonstrate the potential and broad application of NMR metabolomics to conduct studies in the agricultural field, some investigations will be pointed. On one hand, it has been applied to characterize and compare the metabolome of fruits from different varieties as performed by Eisenmann et al., (2016) and Girelli et al., (2018) in apples and melons, respectively. More directed to the agricultural practices, Alves Filho et al., (2019) evaluated the impact on wine from two harvest seasons and two vine training systems, helping winemakers improving the product quality. In addition, Zhao et al., (2016) coupled  $^1\text{H}$ -NMR with GC-MS to investigate metabolic changes in cucumber as a result of exposure to copper nanoparticles, which have been used as pesticides. Regarding phytopathogens, Girelli et al., (2019) investigated the metabolic profile of *Xylella fastidiosa* infected and treated adult olive cultivars and confirmed the presence of a disease indicator biomarker.

Relative to the tomato plant, NMR was applied to access if there is a significant metabolic difference in growing tomatoes conventionally or organically (Hohmann et al., 2014; Watanabe et al., 2015). Recently, Alfosea-Simón et al., (2021) verified an increase in plant growth due to the application of a biostimulant with aspartic and glutamic acid in tomato seedlings. Additionally, Falco et al., (2019) compared the metabolic changes in leaves from tomatoes with distinct tolerance to *Tuta absoluta*. Finally, the impact of a shading mesh on the tomato leaves defence compounds and overall fruit quality from different harvests was performed by Abreu et al., (2019).

## 2 General aims

In this study, the impact of covering tomato plants with different coloured plastics (photosensitive) was evaluated in the metabolome of the fruits and the leaves collected from tomato plants grown at Almería's province from 5-weeks of harvests (during the autumn-winter season from 2019-2020) through  $^1\text{H}$ -NMR spectroscopy coupled to multivariate data analysis techniques. Four different covers were tested: (1) single-roof - the common greenhouse layout; (2) blue plastic double-roof; (3) red plastic double-roof and (4) uncoloured standard plastic double-roof.

Apart from the metabolomic analysis, we aimed to examine the effect of the different covers on the tomato fruits' total production and on the fruit nutritional quality and antioxidant capacity (assessed by ORAC method). Additionally, the effect of different covers in the incidence of *Tuta absoluta*, a major tomato pest, was studied and was correlated with the metabolic profiles of the leaves in order to monitor the levels in compounds involved in the plants defence.

On the other hand, the application of different multivariate data analysis techniques to the  $^1\text{H}$ -NMR spectra had three main purposes: (1) assessing the metabolic changes in tomato fruits harvested in five different weeks; (2) understand how the covering scenario or plastic colour influence the fruits' metabolome and its quality; (3) correlate the degree of infestation of *Tuta absoluta* in the plants covered by different plastics with the presence of possible biomarkers in the leaves that may be associated with plants defence mechanisms.

### 3 Materials and methods

#### 3.1 Plant material and greenhouse experiment layout

The crop considered was tomato (*Solanum lycopersicum* L. cv Ventero) developed during the autumn-winter season 2019-2020 in multi-chapel greenhouses. Ventero is a fruit variety of high calibre and long life. Tomato is a multiannual herbaceous plant usually grown as an annual in short cycles (4-5 months) or long cycles (8-9 months) that has a creeping stem that can root, although it is commonly grown in trellises. The planting framework tends to maintain a distance of 80-100 cm between rows and 30-50 cm between plants (Cajamar, 2016). Transplanting was performed on a coconut fiber substrate with a density of 1.0 plants per m<sup>2</sup>. The rows of plants were perpendicular to the ridges of the greenhouse. The experiment was conducted in three multi-chapel greenhouses, U9, U11 and U12, Figure 6 and Figure 7, located in the research station “Catedrático Eduardo Fernández”, which belongs to the Foundation of the University of Almería-Anecoop. It is located in a place called “Los Goterones” in Almería (Spain), 36°51’ North latitude and 2°17’ West Longitude, at 90 m above sea level, Figure 6 and Figure 7.

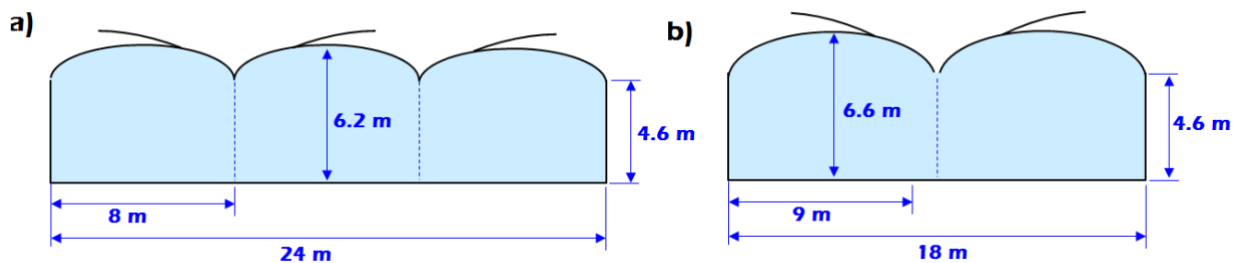


Figure 6 - U9 and U11 greenhouses dimensions (a); U12 greenhouse dimensions.

The greenhouses have a plastic wall and a simple cover. Their surface is 1080 m<sup>2</sup> for the U9 and U11 greenhouses and 810 m<sup>2</sup> for the U12 greenhouse. The main axis is parallel to the North-East / South-West direction (28° with respect to the North direction). The U9, U11 greenhouse has three half-height roof openings and the U12 greenhouse has two half-height roof openings and two side openings. The greenhouses are divided into two sectors, separated by a vertical low-density polyethylene (LD-PE) film, illustrated in Figure 7 with solid black lines.

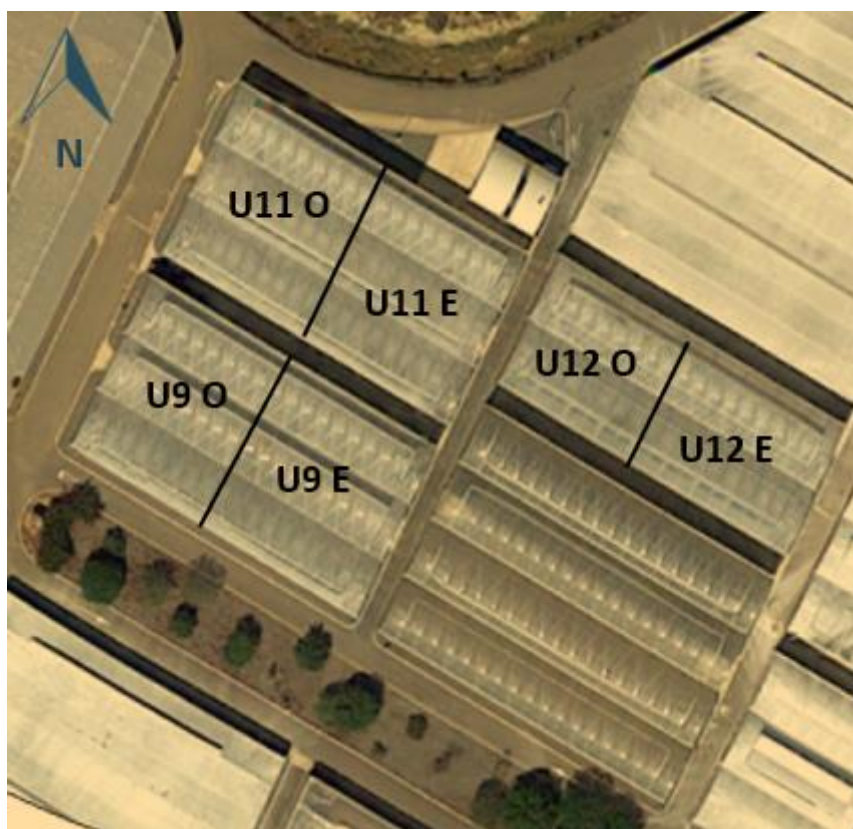


Figure 7 - Greenhouses positioning (Latitude: 36°51' "N; Longitude: 2° 17' W; Altitude: 87 m). N-north; O-west; E-east; S-south.

Experimental plastics based on double roof architectures were installed to analyze its effect. The distribution of materials was as follows:

- 1) In U9 E and U12 E there was a standard uncoloured french Agrypolyane plastic.
- 2) In U9 O there was an experimental photoselective blue plastic.
- 3) In U11 E and U12 O an experimental photoselective red plastic.
- 4) In the U11 O sector there was no plastic cover, remaining as the only one without double roof. We will refer to this sector as control.

Figure 8 shows the particular views on each of these greenhouses.

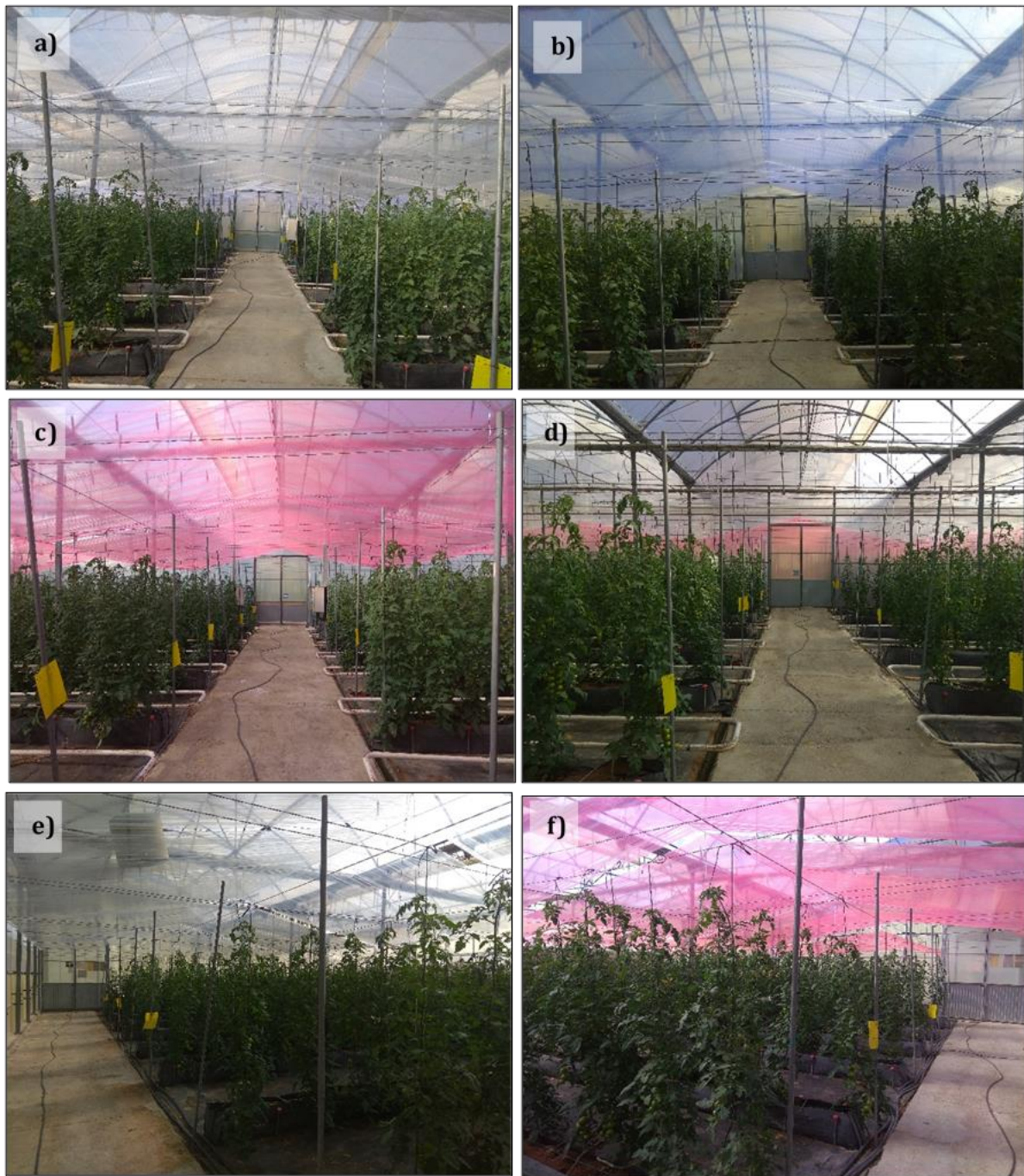


*Figure 8 - Images from tomato culture inside the experimental greenhouses. U9 E with standard Agripolyane double roof (a); U9 O with photo-selective blue plastic double roof (b); U11 E with photo-selective red plastic double roof (c) ; U11 O without double roof (d) ; U12 E with standard Agripolyane double roof (e) and U12 O with photo-selective red plastic double roof (f).*

All the openings of the three greenhouses were fitted with anti-insect nets made of high-density polyethylene. To analyse the possible effect of the treatments on the development of the main greenhouse tomato pest in the Almería region, *Tuta absoluta*, Figure 9a, black colour traps were installed for counting the insects, Figure 9b. This small leafminer burrows inside the tomato leaves, Figure 9c, and in the fruits causing considerable crop losses.



Figure 9 - *Tuta absoluta* butterflies in the insect net inside the side window of the U11E (a), black chromatic trap on the substrate container (b) and galleries in tomato leaves (c) on March 3 2020.

Tomato fruits were harvested successively at 5 consecutive times from January to March: 9/01/2020, 22/01/2020, 06/02/2020, 20/02/2020 and 02/03/2020. Samples for NMR analysis were taken randomly, as twenty fruits from each greenhouse at the corresponding harvest time were pooled together.

Tomato leaves were collected randomly from all the plants of each greenhouse, from the middle part of the plants, in the last day of harvest, 02/03/2020. Samples for NMR analysis were taken from random plants, as 20 leaves from each greenhouse were pooled together.

Results displayed in the Section “Tomato Production Data” were obtained from the company which is developing this greenhouse experiment.

The results are presented with respect to the plastic covering colour and not directed to each greenhouse. In the scenario where the same covering colour was applied in two greenhouses, the average value for each parameter was calculated. This proceeding was performed for all the variables studied.

### 3.2 Sample preparation and extraction

**Tomato fruits.** Twenty tomatoes collected from each greenhouse and respective harvesting stage were pooled together into one mixed sample, washed with water and cut along the transversal direction discarding the half including the peduncle. The remaining halves were introduced in a jar and triturated until obtaining a jelly and consistent product. The previously tagged 30 mL flasks were half filled. After tagging and filling the flasks with each sample and its replicates, these were stored in a freezer at -30 °C. Prior to extraction, tomato samples were freeze-dried, and 35 milligrams were weighted in Eppendorfs in a digital balance. Fruit samples were extracted in 800 mL of a 50:50 (v/v) solvent system based on CD<sub>3</sub>OD and D<sub>2</sub>O-KH<sub>2</sub>PO<sub>4</sub> buffer (pH 6.0) containing the sodium salt of 3-(trimethylsilyl)propionic-2,2,3,3-*d*<sub>4</sub> acid (TSP, 0.01% w/w) and the enzyme inhibitor sodium azide NaN<sub>3</sub> at 90 µM. Extraction had three differentiated steps. First, the mixtures were placed in an ultrasound bath for 15 min at 20 °C. Then they were stirred and vortexed continuously for 5 min at 600 rpm, followed by centrifugation for 5 min (13 500g). Finally, 500 µL of the supernatants were transferred to oven-dried 5 mm NMR tubes with the help of a micropipette. For the experimental conditions carried out in just one greenhouse (i.e photo-selective blue plastic at U9 O and no covering at U11 O), six biological replicates for each sample were performed making up a total of 30 spectrum for each case. For the remaining conditions carried out in two greenhouses (i.e photo-selective red plastic at U11 E, U12 O and the standard uncoloured plastic at U9 E and U12 E), three replicates for each sample were performed making up a total of 15 spectrum for each case. In total, 120 tomato fruit samples were extracted and subsequently investigated by NMR spectroscopy.

**Tomato leaves.** Leaves were immediately frozen with liquid nitrogen, smashed to powder, stored under -30 °C and freeze-dried prior to extraction. An overall of 30 milligrams of freeze-dried leaf samples were extracted with 1 mL of the same solvent system in a 80:20 (v/v) ratio of CD<sub>3</sub>OD and D<sub>2</sub>O-KH<sub>2</sub>PO<sub>4</sub> buffer, respectively. As the tomato fruits, the mixtures were sonicated for 15 min at 20 °C. In this case, vortexing was performed for 10 minutes at 600 rpm and centrifugation was equally carried for 5 min (13 500g). Supernatants were transferred to oven-dried 5 mm NMR tubes. Six biological replicates were performed for the photo-selective blue double roof and for the control, while three replicates were performed for each sample derived from the remaining conditions which were tested in two greenhouses. In total, 24 tomato leaves samples were extracted and subsequently investigated by NMR spectroscopy.

### 3.3 NMR experiments

The tomato (fruits and leaves) extracts from coloured nets with different harvesting days were analysed. In total, there were measured 144 samples by NMR. The spectra were registered at  $293 \pm 0.1$  K on a Bruker Avance III 600 spectrometer with a proton frequency of 600.13 MHz, equipped with an autosampler thermostabilized with 24 positions, using a 5 mm QCI quadruple resonance pulse field gradient cryoprobe. Samples were measured without rotation by using the Bruker 1D noesygppr1d sequence. The acquisition parameters were established in the following manner: dummy scans (DS) - 4, number of scans (NS) - 32, fid size (TD) - 64 K, spectral width (SW) - 20.0 ppm, acquisition time (AQ) - 2.73 s, recycle delay (D1) - 10 s, receiver gain (RG) - 57, fid resolution (FIDRES) - 0.25 Hz and mixing time (D8) - 10 ms. The water signal was suppressed by application of a pre-saturation pulse sequence module. As mentioned previously, the applied 1D-NOESY experiment is recognized as the top-used sequence regarding NMR metabolomic studies (Mckay, 2011).

Phase and baseline were automatically corrected and calibrated to the TSP signal which is automatically located at 0.0 ppm by applying the “sref” command. The spectrometer transmitter was locked to CH<sub>3</sub>OH-*d*<sub>4</sub> (deuterated methanol). Spectra acquisition and processing were performed with TOPSPIN software (version 3.1). Observed multiplicities were labelled as s = singlet; d = doublet; dd = double doublet; t = triplet and m = multiplet.

Moreover, two-dimensional experiments such as <sup>1</sup>H-<sup>1</sup>H homonuclear correlation spectroscopy (COSY), <sup>1</sup>H-<sup>1</sup>H total correlation spectroscopy (TOCSY), <sup>1</sup>H-<sup>13</sup>C heteronuclear single quantum coherence (HSQC) and <sup>1</sup>H-<sup>13</sup>C heteronuclear multiple bonds coherence (HMBC), were performed using standard Bruker sequences. The TOCSY spectrum was obtained with a relaxation delay of 1.5 s, a spin-lock time (D9) of 120 ms, a spectral width of 8403.36 Hz in both dimensions and a receiver gain of 114. Spectra was processed with a sine-bell window function (SSB = 2.0). The HSQC spectrum was obtained with a relaxation delay of 1.0 s, spectral width of 7812.50 Hz in F2 and 37729.66 Hz in F1. A quadratic sine window function (SSB = 2.0) was applied to the HSQC spectrum. The same parameters were applied for the HMBC spectrum except for 7500.00 Hz F2 spectral width. Coupling evolution delays were optimized for <sup>1</sup>J<sub>CH</sub> = 145 Hz and <sup>n</sup>J<sub>CH</sub> = 8 Hz for the heteronuclear experiments.

### 3.4 Oxygen radical absorbance capacity (ORAC) assay

The ORAC method measures the decrease in fluorescence of an organic molecule, fluorescein, due to a structural modification that occurs when it reacts with peroxide radicals (ROO<sup>-</sup>). Peroxide radicals are produced as consequence of the addition of an initiator compound, 2,2'-

Azobis(2-amidinopropane)dihydrochloride (AAPH). If the sample has compounds with antioxidant capacity, these will scavenge peroxide radicals, and the fluorescein fluorescence will keep constant or, at least, keep constant for a longer period.

The ORAC experiment was carried out on a CYTATION microplate reader, from Bio-Tek Instruments, Inc- (Winooski, VT), with 96-well black polystyrene microplates. Fluorescence was measured with an excitation wavelength of 485 nm and an emission filter of 520 nm. The oxygen radical absorbance capacity was performed as described by Bellán et al., 2008, with some modifications. The reaction was carried out with a  $\text{KH}_2\text{PO}_4$  buffer (75 mM, pH 7.4), with a final reaction mixture of 200  $\mu\text{L}$ . Tomato extracts (25  $\mu\text{L}$ , 5:100 diluted) and fluorescein (150  $\mu\text{L}$ , 1  $\mu\text{M}$ ) were placed in the wells of the microplate. Then, 25  $\mu\text{L}$  of AAPH (250 mM) were quickly added with a multichannel pipet. The microplate was immediately placed in the reader, automatically shaken before each reading, and the fluorescence was recorded every 4 min for 72 min. For each assay, calibration solutions were done with Trolox dissolutions, a known antioxidant, in five concentrations (12.5, 25.0, 50.0, 100.0 and 200.0  $\mu\text{M}$ ) plus a buffer solution. The blank was performed with  $\text{KH}_2\text{PO}_4$  buffer solution instead of the antioxidant solution. Every reaction mixture was prepared in triplicate, and at least three independent assays were performed for each sample.

Results were expressed as relative fluorescence in comparison with the initial reading. The area under the fluorescence decay curve (AUC) was calculated according to Bellán et al., 2008, by the following equation:

$$AUC = \left( 0.5 + \left( \sum_{i=1}^{i=18} \frac{f_i}{f_1} \right) \right) * CT \quad (1)$$

Where  $i$  is the number of cycles;  $F$  the unities of fluorescence and  $CT$  the time of each cycle in minutes.

In addition, the difference between the areas below the curves associated to the fluorescence loss in each Trolox concentration and the blank were calculated (AUC net), and these were represented towards Trolox concentrations, yielding a calibration curve. Applying the linear regression ( $y = a + bx$ ), the antioxidant capacity results for each tomato extract were expressed in equivalent  $\mu\text{M}$  of Trolox per gram of lyophilized tomato.

### 3.5 Statistical analysis

The NMR spectra were bucketed with AMIX 3.9.12 (Bruker BioSpin GmbH, Rheinstetten, Germany). Buckets were obtained by integrating intervals of 0.04 ppm and normalizing individual peak intensity in relation to the total intensity registered in the region  $\delta_H$  0.2 to 10.0 ppm. Due to the presence of the residual signals from H<sub>2</sub>O and CH<sub>3</sub>OH, the regions of  $\delta_H$  4.89 to 4.73 and 3.32 to 3.29 ppm, respectively, were excluded from the analysis. Principal component analysis (PCA), partial least square discriminant analysis (PLS-DA) and orthogonal projection to latent structures (OPLS) models were employed. PCA and OPLS were scaled to Pareto, while PLS-DA was scaled to unit variance. These multivariate analyses were carried on SIMCA-P software (v. 14.0, Umetrics, Sweden). To visualize relationships between samples and the different experimental conditions, hierarchical clustering was performed. A heatmap was created using the Statistical analysis tool in MetaboAnalyst 4.0 software, which were based on the Euclidean distance measure and the Ward clustering algorithm. The presented metabolites were selected by analysis of variance (ANOVA) using a significance level of  $p < 0.05$  and post-hoc analysis based on Fisher's least significant difference (LSD). Finally, ORAC data was subjected to statistical analysis by ANOVA. Least significant difference (LSD) test was used to compare means and significance accepted at  $P \leq 0.05$ .

## 4 Results and Discussion

### 4.1 Metabolite profiling on tomato fruits

This work is based on the tomato fruits and leaves metabolic changes induced by the different coloured plastic coverings. First, the metabolic profiling of the tomato fruit was performed through NMR spectroscopy. In general, tomatoes contain 94% water and 6% of dry matter which consists of 55% sugar (primarily glucose and fructose), 20% of alcohol-soluble constituents (proteins, polysaccharides, cellulose, pectin), 12% of organic acids (mainly citric acid), 8% of inorganic substances (minerals) and 5% of other substances, such as amino acids (Herrmann, 1979).

To achieve the most complete and broad picture of the species containing tomato, an optimization of the extracting solvent system was carried. Spectra obtained with extractions performed with methanol 80:20 water (v/v) and methanol 1:1 water (v/v) were compared. In fact, the spectra yielded with the first solvent contained several lipids, due to the increased alcoholic component of the mixture, which allowed the presence of fatty acids in larger quantities that hinder the detection of compounds that have resonances in the same intervals (Annexes A). On the contrary, the experiment engaged with the second, more polar, solvent lead to a more populated spectra with better resolution, enabling the identification of more compounds and with an enhanced representation of the fruit metabolome as pretended. The second solvent system was therefore chosen, and all the NMR experiments described herein were carried out with it.

In general,  $^1\text{H}$  NMR spectra of aqueous extracts of biological matrices are divided in three main regions, according to the groups of compounds present (Botoran et al., 2019; Cazor et al., 2006). The tomato fruit studied in this thesis is not an exception, and these are:

- Region A:  $\delta_{\text{H}}$  0 - 3.21 ppm, the region of hydrogen atoms belonging to amino acids and organic acids;
- Region B:  $\delta_{\text{H}}$  3.21 - 5.5 ppm, typical region for carbohydrates, wherein hydrogens of anomeric carbons are clearly separated from the remaining signals;
- Region C:  $\delta_{\text{H}}$  5.5 - 10 ppm, which reflects the phenolic metabolites and aromatic protons.

In this methanol-water extract, 41 metabolites were unequivocally assigned. These compounds belong to several classes such as amino acids, organic acids, sugars, nucleosides and nucleotides, phenolics (flavonoids and phenylpropanoids), sterols, fatty acids, among others. Figure 10 displays an expansion of these three main regions typical from a tomato fruit  $^1\text{H}$ -NMR spectrum with each identified compound on their respective resonance signals.

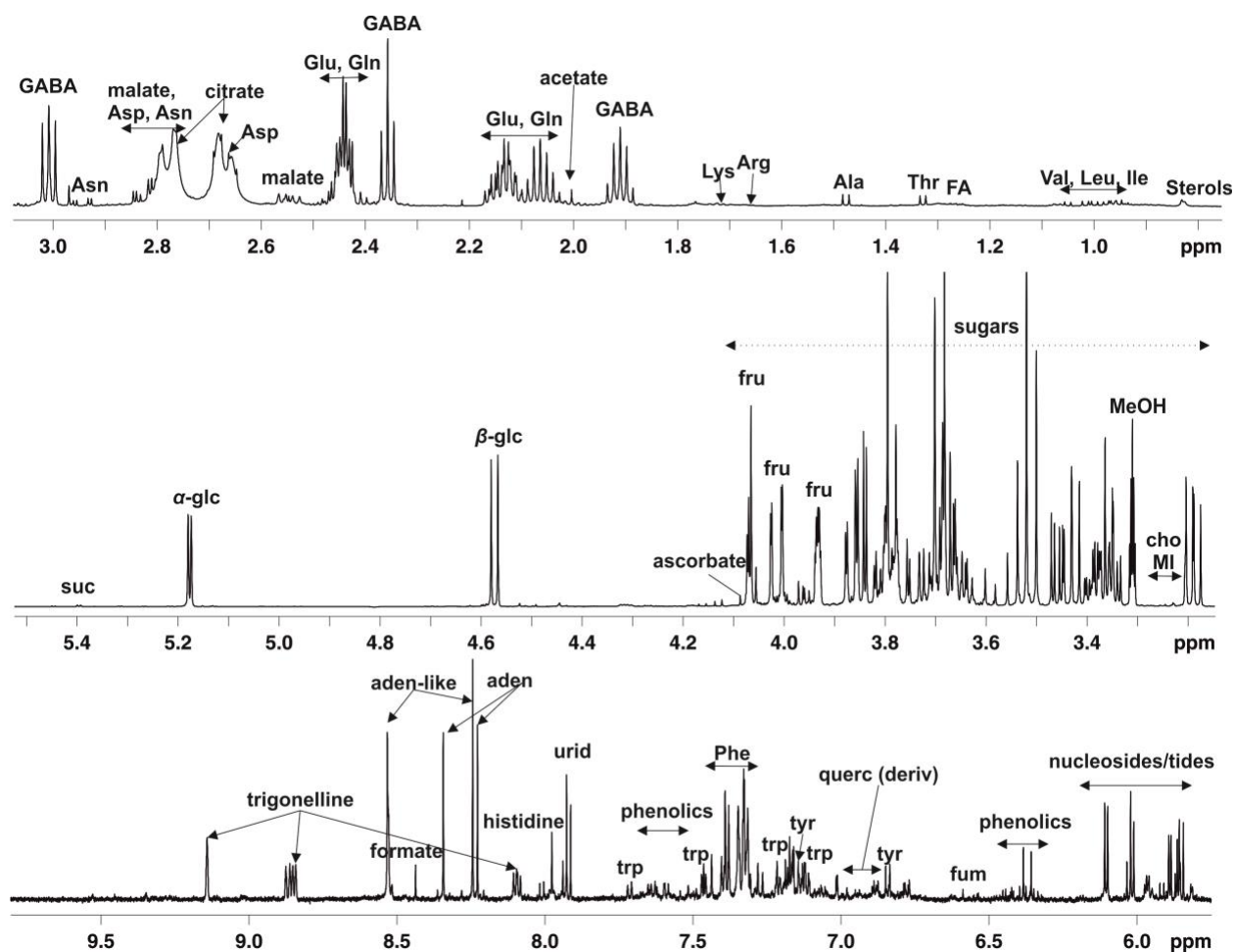


Figure 10 -  $^1\text{H}$  NMR (600 MHz) spectrum of a tomato fruit  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (50:50, v/v) extract. Aden - adenosine, aden-like - adenosine like, Ala - alanine, Arg - arginine, Asn - asparagine, Asp - aspartate, cho - choline, FA - fatty acids, fru - fructose, glc - glucose, Glu - glutamate, Gln - glutamine, Ile - isoleucine, Leu - leucine, Lys - lysine, MeOH - methanol, MI - myo-inositol, Phe - phenylalanine, querc(deriv) - quercetin derivative, suc - sucrose, thr - threonine, trp - tryptophane, tyr - tyrosine, urid - uridine, val - valine. Chemical shifts, coupling constants and multiplicity reported in Table 2.

The first region from  $\delta_{\text{H}}$  0 to 3.21 ppm, starts with sterols and fatty acids. The principal phytosterols found in tomato are  $\beta$ -cholesterol, sitosterol,  $\Delta^5$ -campesterol,  $\Delta^5$ -stigmasterol and  $\Delta^5$ -brassicasterol (signals the presence of an unsaturation in the 5<sup>th</sup> carbon). Their nutritional interest derives from their contribution to reduce cholesterol levels, preventing cardiovascular diseases (Giuffrè, A.M. and Capocasale, 2016; Piironen et al., 2000). Relative to tomatoes' fatty acids content, linoleic acid ( $\omega$ -6) is the main compound accounting for 52% of total fatty acids. Followed by oleic acid, 19%, linolenic acid ( $\omega$ -3) with 4% and lately palmitoleic acid, responsible for 1.82%. Although tomatoes are characterized by a low fatty acids' content, the high proportion (77.6%) of polyunsaturated fatty acids (PUFA) in ripe fruits contributes to their antioxidant capacity (Elbadrawy, 2011). In fact, linoleic ( $\omega$ -6) and linolenic acid ( $\omega$ -3) are the

two essential fatty acids for humans and have to be ingested in our diet. These are very important constituents of all cell membranes, reported to prevent several disease conditions and improve overall nervous system functioning (Das, 2006). Besides these, it is mainly composed by amino acids and organic acids. With respect to amino acids, isoleucine, leucine, valine, threonine, lysine, alanine, arginine, proline, glutamate, glutamine, aspartate, asparagine and  $\gamma$ -aminobutyrate (GABA) were found. The organic acids, acetate, malate and citrate were also found in this region.

The region from  $\delta_H$  3.21 to 5.5 ppm, mainly contains signals from carbohydrates such as fructose, sucrose and glucose. A singlet at  $\delta_H$  3.22 ppm originated from choline was also present. In fact, choline is the precursor of phosphatidylcholine, one of the most important phospholipids in all plant membranes. Relative to human nutrition, it is a vitamin-like compound active in several crucial metabolic routes, being considered an essential nutrient (Zeisel, 2000). Even though it is low concentrated in tomatoes, their high consumption can to some extent compensate (Zeisel et al., 2003). Moreover, *myo*-inositol (MI) and ascorbate (vitamin C), one of the vitamins found in tomatoes and its second strongest antioxidant compound, were also located. MI is one of the nine stereoisomers of inositol, a cyclic sugar. It is often considered as part of the vitamin B complex and it is a biologically important molecule to the human body, playing a structural and functional role in cell membranes, being closely related to second messengers of several metabolic pathways (Milewska et al., 2016). On the other hand, MI biosynthesis pathway has been shown to take part in ascorbic acid accumulation in tomato (Munir et al., 2020; Ntagkas et al., 2019).

In the downfield region from  $\delta_H$  5.5 to 10 ppm, also named aromatic region, signals are weaker, less intense. These were assigned to the aromatic moieties of amino acids such as tyrosine, tryptophan, phenylalanine and histidine. Nucleosides, such as uridine-derived compound and adenosine-derived compounds were identified as were the organic acids, fumarate and formate. Trigonelline was the only characterized alkaloid. Additionally, regarding phenolic compounds, signals belonging to two flavonoids were visualized in the spectra. These probably belong to rutin (quercetin-3-O-rutinoside), and to another quercetin-derived compound (whose characterization was not possible by NMR due to its low concentration in the spectra). Quercetin is a natural flavonoid widely distributed in vegetables and fruits which commonly appears as a glycoside, for example, rutin. Rutin can also be named vitamin P and exhibits therapeutic effects due to its powerful antioxidant and anti-inflammatory properties (Budzynska et al., 2019; Chen et al., 2020). Other studies regarding tomato flavonoids suggested rutin, naringenin and naringenin chalcone as the main flavonoids while kaempferol 3-rutinoside, naringenin 7-glucoside and quercetin-3-O- $\beta$ -glucopyranoside were suggested as the minor flavonoids (Slimestad et al., 2008). These appear in residual quantities in the fruits, explaining why their

resonances were not detected in the spectrum. Relative to phenolic acids, the  $^1\text{H}$ -NMR spectra showed typical signals of the protons of an (*E*)-configured olefinic double bond from three hydroxycinnamic acid derivatives with a coupling constant of about 16 Hz. These probably belong to caffeic and chlorogenic acid, which are the most characteristic phenylpropanoids found in tomatoes, according to Baldina et al., (2016,) who reported that chlorogenic acid and caffeic acid hexoses represented 55% of the average total phenolic content in tomatoes. The third one is most likely *p*-coumaric acid, as high signals of conjugated forms of this acid have been detected by Moco et al., (2007) and tomato fruits have been reported to be rich in *p*-coumaric acid (Kiokias et al., 2020).

A list with every metabolite identified in this tomato fruit extract is provided in Table 2, detailing for each signal its chemical shifts ( $\delta\text{H}$ , ppm), multiplicity and coupling constants (*J*, Hz). Identification and characterization were based in the interpretation of homo- and heteronuclear 1D and 2D NMR experiments ( $^1\text{H}$ -NOESY,  $^1\text{H}$ - $^1\text{H}$  COSY,  $^1\text{H}$ - $^1\text{H}$  TOCSY,  $^1\text{H}$ - $^{13}\text{C}$  HSQC and  $^1\text{H}$ - $^{13}\text{C}$  HMDB) combined with browsing in NMR databases (HMDB, Spectral Reference Libraries from Chenomx software) and literature data (Abreu et al., 2019; C. Deborde et al., 2009; Hohmann et al., 2014; Le Gall et al., 2003).

Table 2 - Peak assignment of metabolites identified in the tomato fruits  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (50:50, v/v) extracts by NMR

| Metabolite         |               | Chemical shifts (ppm), multiplicity and coupling constants (Hz)                           |
|--------------------|---------------|-------------------------------------------------------------------------------------------|
| <b>Amino acids</b> |               |                                                                                           |
| 1                  | Valine        | 1.0 (d, <i>J</i> = 7.2 Hz), 1.05 (d, <i>J</i> = 7.1 Hz)                                   |
| 2                  | Isoleucine    | 0.95 (t, <i>J</i> = 7.4 Hz), 1.01 (d, <i>J</i> = 7.2 Hz)                                  |
| 3                  | Leucine       | 0.98 (d, <i>J</i> = 6.3 Hz), 0.96 (d, <i>J</i> = 6.3 Hz)                                  |
| 4                  | Threonine     | 1.33 (d, <i>J</i> = 6.7 Hz), 4.22 (m)                                                     |
| 5                  | Alanine       | 1.48 (d, <i>J</i> = 7.2 Hz);                                                              |
| 6                  | GABA          | 1.91 (quint, <i>J</i> = 7.3 Hz), 2.38 (t, <i>J</i> = 7.3 Hz), 3.01 (t, <i>J</i> = 7.3 Hz) |
| 7                  | Lysine        | 1.45 (m), 1.72 (m), 1.89 (m), 3.00 (m)                                                    |
| 8                  | Arginine      | 1.68 (m), 1.74 (m), 1.89 (m), 3.24 (m)                                                    |
| 9                  | Proline       | 1.99 (m), 2.05 (m), 2.33 (m), 4.14 (dd)                                                   |
| 10                 | Glutamate     | 2.07 (m), 2.14 (m), 2.45 (m)                                                              |
| 11                 | Glutamine     | 2.12 (m), 2.44 (m)                                                                        |
| 12                 | Aspartate     | 2.68 (dd, <i>J</i> = 17.5, 9.0 Hz), 2.84 (dd, <i>J</i> = 17.5, 3.8 Hz)                    |
| 13                 | Asparagine    | 2.80 (dd, <i>J</i> = 17.0, 8.7 Hz), 2.94 (dd, <i>J</i> = 17.0, 3.9 Hz)                    |
| 14                 | Tyrosine      | 6.84 (d, <i>J</i> = 8.5 Hz), 7.17 (d, <i>J</i> = 8.5 Hz)                                  |
| 15                 | Tryptophan    | 7.13 (m), 7.28 (s), 7.20 (m), 7.46 (d, <i>J</i> = 8.0 Hz), 7.71 (d, <i>J</i> = 8.0 Hz)    |
| 16                 | Phenylalanine | 7.33 (m), 7.39 (m)                                                                        |

17 Histidine 7.37 (d,  $J = 1.2$  Hz), 8.58 (d,  $J = 1.2$  Hz)

### Organic acids

18 Lactate 1.34 (d,  $J = 6.98$  Hz), 4.11 (q,  $J = 6.9$  Hz)  
 19 Acetate 2.02 (s)  
 20 Malate 2.57 (dd,  $J = 16.0, 7.6$  Hz), 2.79 (dd,  $J = 16.0, 5.0$  Hz),  
 4.33 (dd,  $J = 7.6, 5.0$  Hz)  
 21 Citrate 2.69 (d,  $J = 15.7$  Hz), 2.80 (d,  $J = 15.7$  Hz)  
 22 Fumarate 6.60 (s)  
 23 Formate 8.53 (s)

### Sugars

24 Fructose 4.07 (m)  
 25 Glucose 4.57 (d,  $J = 8.3$  Hz), 5.18 (d,  $J = 3.7$  Hz)  
 26 Sucrose 5.45 (d,  $J = 3.7$  Hz)  
 27 Myo-inositol 3.23 (t,  $J = 9.4$  Hz), 3.52 (dd,  $J = 9.9, 2.8$  Hz), 4.05 (t,  $J = 2.8$  Hz)

### Nucleosides/tides

28 Uridine 5.85 (d,  $J = 8.1$  Hz), 5.89 (d,  $J = 4.7$  Hz), 7.92 (d,  $J = 8.1$  Hz)  
 29 Adenosine 6.01 (d,  $J = 6.4$  Hz), 8.23 (s), 8.34 (s)  
 30 Uridine-like 6.03 (d,  $J = 6.1$  Hz), 7.94 (d,  $J = 8.1$  Hz)  
 31 Adenosine-like 6.10 (d,  $J = 5.9$  Hz), 8.24 (s), 8.42 (s)

### Phenylpropanoids and phenolic compounds

32 Cinnamic acid derivative 1 6.29 (d,  $J = 15.8$  Hz), 7.50 (d,  $J = 15.7$  Hz)  
 33 Cinnamic acid derivative 2 6.37 (d,  $J = 15.7$  Hz), 7.57 (d,  $J = 15.8$  Hz)  
 34 Cinnamic acid derivative 3 6.41 (d,  $J = 15.8$  Hz)  
 35 Rutin 6.33 (d,  $J = 1.8$  Hz), 6.55 (d,  $J = 1.8$  Hz), 6.98 (d,  $J = 8.7$  Hz), 7.63 (dd,  $J = 8.7, 2.0$  Hz), 7.67 (d,  $J = 2.0$  Hz)  
 36 Quercetin-like 6.33 (d,  $J = 2.0$  Hz), 6.54 (d,  $J = 2.0$  Hz), 6.97 (d,  $J = 8.4$  Hz), 7.61 (dd,  $J = 8.4, 2.1$  Hz), 7.64 (d,  $J = 2.1$  Hz)

### Others

37 Sterols 0.77-0.83 (s)  
 38 Fatty acids\* 0.87 (-CH<sub>3</sub>, for all FA except n-3), 0.96 (-CH<sub>3</sub>, for n-3),  
 1.24-1.30 (-CH<sub>2</sub>)<sub>n</sub>-, 1.58 (-CH<sub>2</sub>-CH<sub>2</sub>-COOR), 2.04 (-CH<sub>2</sub>-CH=CH-, for UFA), 2.29 (-CH<sub>2</sub>-COOR), 2.74 (=CH-CH<sub>2</sub>-CH=, for PUFA), 5.40 (-CH=CH-, for UFA)  
 39 Choline 3.22 (s)  
 40 Ascorbate 4.64 (d,  $J = 2.2$  Hz)  
 41 Trigonelline 8.09 (dd,  $J = 8.0; 6.1$  Hz), 8.85 (d,  $J = 6.1$  Hz), 8.87 (d,  $J = 8.0$  Hz), 9.14 (s)

---

FA - fatty acids; n-3 -  $\omega$ -3; UFA - unsaturated fatty acids; PUFA - polyunsaturated fatty acids.

Besides vitamin C, tomatoes are also source of vitamin A and E (Cameron & Geana, 2005). In contrast to ascorbic acid (vitamin C), retinol (vitamin A) and tocopherol (vitamin E) are liposoluble, so attempting these phytochemicals detection would require a more apolar solvent system (Martí et al., 2016; Takeoka et al., 2001).

Furthermore, tomatoes contain natural pigments, the liposoluble carotenoids. Lycopene is of particular interest, as it is the fruits' biomolecule with highest antioxidant activity. Indeed, these were not detected because its extraction has to be carried with apolar solvents, such as acetone, chloroform and hexane (Choksi & Joshi, 2007). The extraction and consequent NMR detection of carotenoids, lycopene,  $\beta$ -carotene and phytoene, among others, can be successfully achieved with chloroform extracts, as performed by Masetti et al., (2017) and Tiziani et al., (2006).

## 4.2 Coloured double-roof effect on tomatoes antioxidant activity

Tomato regular consumption has been associated to several health benefits, mainly the reduced risk of cardiovascular disease and certain types of cancer (Burton-Freeman & Reimers, 2011; Cameron & Geana, 2005). Extensive data seeking for the compounds which lead to such health effects has been generated and, in fact, there is epidemiological evidence attributing the total content of phenolics, vitamins and carotenoids for the fruits' protective effect as these compounds exhibit antioxidant properties. Ultimately, the antioxidant activity is used as an indicator of the ability of the product to scavenge free radicals and reduce oxidative stress (Frusciante et al., 2007).

In this work, the antioxidant activity of tomato extracts was measured by ORAC method in 96-wells microtiter plates to understand how different plastic covers on tomato plants influence their fruits' nutritional value. After reading the 96-wells microplates, where fluorescence was measured with excitation and emission wavelengths of 485 and 520 nm, respectively, the fluorescence decay curves along time of Trolox, the reference compound in several concentrations were obtained as shown in Figure 11. A brief description of the analytical processing of data is given below. The detailed proceeding is provided in Annexes B.

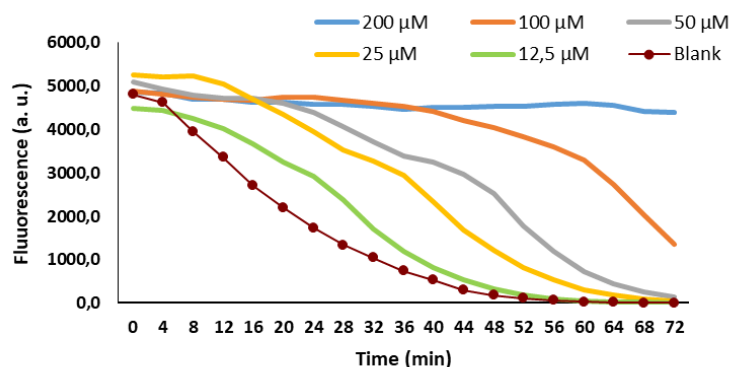


Figure 11 - Trolox fluorescence evolution along time, at different concentrations: 200  $\mu\text{M}$ , 100  $\mu\text{M}$ , 50  $\mu\text{M}$ , 25  $\mu\text{M}$  and 12,5  $\mu\text{M}$ . Fluorescence values were successively measured every 4 minutes in a microplate reader.

Applying the ORAC method, the clean area below the curve ( $\text{AUC}_c$ ) was calculated, by subtracting blank AUC to the sample AUC, thus eliminating the buffer interference. Figure 12 is a visual representation of the proceeding carried to obtain  $\text{AUC}_c$ .

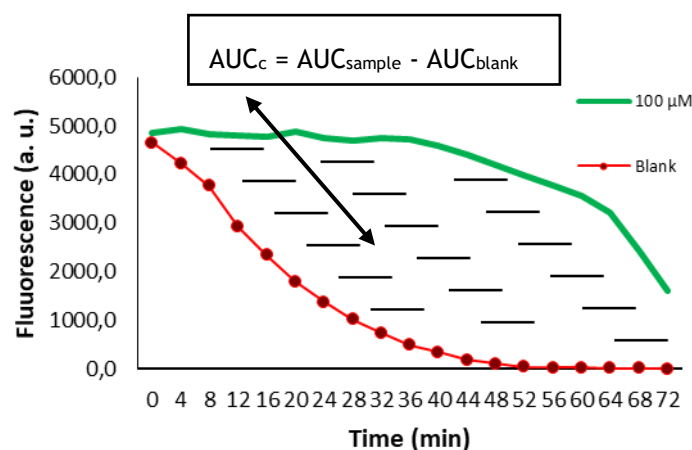


Figure 12 - Visual illustration of the method applied to calculate clean/net AUC of Trolox and every tomato extract sample, through the ORAC method.

The Trolox equivalents of tomato samples were calculated with a linear regression  $y = mx + b$ , where  $y$  is the Trolox  $\text{AUC}_c$  and  $x$  the Trolox concentrations. This calibration curve, as shown in Figure 13, allows the quantification of Trolox equivalent micromoles (TE) per gram of freeze-dried tomato, by introducing the  $\text{AUC}_c$  of each sample in the equation.

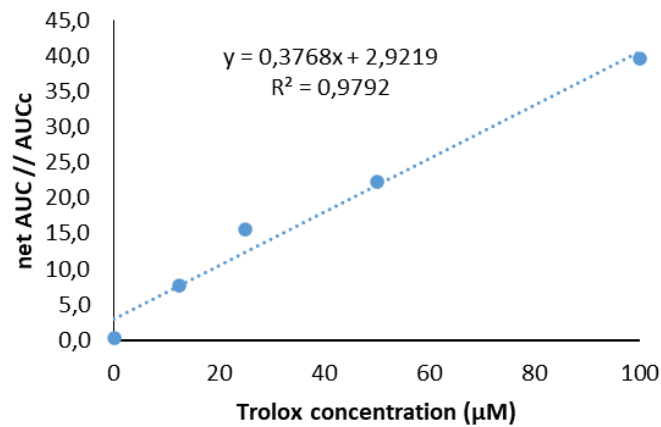


Figure 13 - Example of a Trolox calibration curve used in one of the four independent assays, which enables the quantification of Trolox equivalent micromoles per gram of lyophilized tomato in each sample.

The final results were expressed as equivalent micromoles of Trolox (TE) per gram of freeze-dried tomato, as shown in Figure 14.

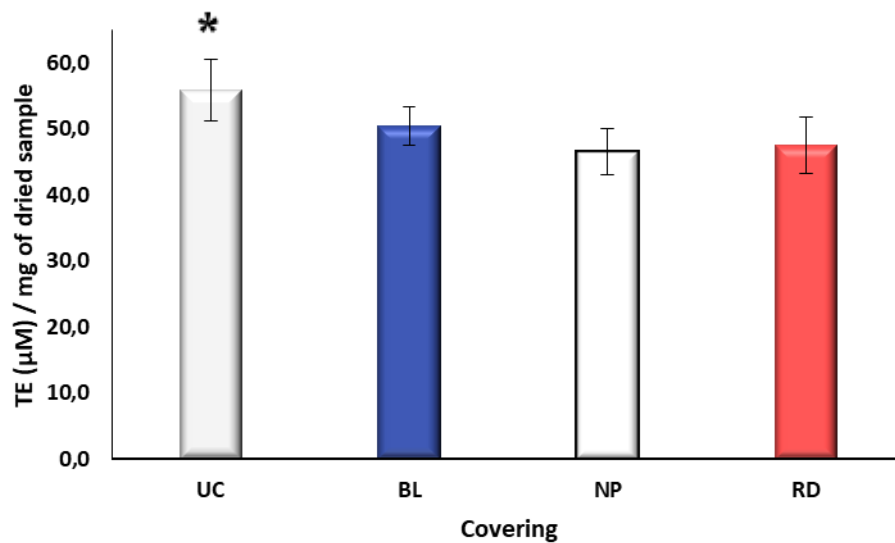


Figure 14 - Antioxidant activity measured by ORAC methods for extracts of tomato fruits from 1st to 5th harvest grown under different coloured plastics. Uncoloured standard plastic double-roof (UC), blue photo-selective plastic double-roof (BL); without double plastic covering (NP) and red photo-selective plastic double-roof (RD). Results are expressed as Trolox equivalent (TE, μM) per gram of dried sample. Bar with \* mean the antioxidant activities is significantly higher compared with the other conditions ( $p < 0.05$ ).

The values obtained were  $55.8 \pm 4.6$ ,  $50.4 \pm 2.8$ ,  $46.5 \pm 3.4$  and  $47.6 \pm 4.2$  TE for tomatoes grown under double-roof with uncoloured plastic covering, double-roof with blue plastic covering, no double-roof and double-roof with red plastic covering, respectively. Indeed, the antioxidant activity of tomatoes covered with the standard, uncoloured plastic is significantly higher ( $p < 0.05$ ) than every other condition under study. The comparison of these results with the metabolic profile of tomatoes from uncoloured plastic will be further analysed below, in order to conclude about the compounds that may be responsible for this increase in the antioxidant activity.

### 4.3 Tomato production data

Besides studying the tomato fruit metabolome, other data regarding the cultivation of tomato plants under double-roofs with specific plastic properties was analysed. The purpose was to detect if the impact of using different plastics is evident on several physio-chemical properties, which dictate the commercial value, and looking for further relations with the metabolic content. The data is presented in the following Tables 3,4, and includes fruit productivity, common tomato quality parameters, chromatic characteristics, photosynthetic parameters and morphological data.

Regarding fruit production, tomato plants' cultivation in greenhouses with uncoloured/coloured double-roof improves both, total and marketable productivity, when compared to the conventional growth in greenhouses with single roof, Table 3. In terms of some parameters usually accessed to measure tomato quality, fruits harvested from greenhouses with red double-roof and single-roof were bigger and heavier, although not statistically significant. Fruit hardness and acidity were very similar for all coverings. However, Brix value, which is positively related with levels of nutritional valuable compounds and fruit flavour, as well as dry matter percentage, that has been an aim of agricultural practices, were both higher for tomato fruits grown under uncoloured and blue double-roofs (Lugasi, 2006; Philouze, 2001), but not statistically significant. Tomato colour is an external characteristic that provides information about the ripening status with huge impact in the consumer's choice, commonly measured by the  $a^*/b^*$  ratio, proportional to maturation (Camelo & Gómez, 2004). In fact, fruits from the blue and uncoloured plastic double-roofs had significantly higher values ( $p < 0.05$ ) of chromaticity.

*Table 3 - Average values for tomato fruit total production, marketable production and quality parameters, obtained for the experiment carried on the growth of Solanum lycopersicum L. cv Ventero, under single-roof and double-roof plastic coverings, developed during the autumn-winter season 2019-2020. Fruit weight - Fw; fruit equator diameter - Fd; fruit hardness - Fh; sucrose fraction - Sf; fruit acidity - pH and dry matter - Dm. Colorimetric coordinates: Lightness - L\* and chromaticity a\*/b\*.*

|                    | Fruit Production           |                                 | Tomato quality |         |            |            |      |        |       |        |
|--------------------|----------------------------|---------------------------------|----------------|---------|------------|------------|------|--------|-------|--------|
|                    | Total (kg/m <sup>2</sup> ) | Marketable (kg/m <sup>2</sup> ) | Fw (g)         | Fd (mm) | Fh (kg/cm) | Sf (°Brix) | pH   | Dm (%) | L*    | a*/b*  |
| <b>Uncoloured</b>  | 5.0                        | 2.9                             | 102.5a         | 57.9a   | 1.9a       | 5.7a       | 4.1a | 7.3a   | 43.7b | 0.84b  |
| <b>Blue</b>        | 4.6                        | 2.9                             | 105.8a         | 57.3a   | 2.0a       | 5.6a       | 4.1a | 7.1a   | 42.5a | 0.88b  |
| <b>Red</b>         | 4.7                        | 3.2                             | 107.7a         | 58.8a   | 2.0a       | 5.4a       | 4.1a | 6.8a   | 43.2a | 0.765a |
| <b>Single-roof</b> | 4.3                        | 2.5                             | 109.3a         | 59.4a   | 2.1a       | 5.4a       | 4.1a | 6.9a   | 42.9a | 0.79a  |

*Different letters in the same column showed statistically significant differences with a confidence degree of 95.0% ( $p < 0.05$ ).*

Regarding leaves photosynthesis, overall leaves from greenhouses with uncoloured plastic double-roof received significantly more ( $p < 0.05$ ) photosynthetic active radiation, mainly, near the central hours of the day, Table 4. Though, leaves from plants grown under blue plastics revealed the highest photosynthetic activity, although not significantly higher than other double-roofs. It might be related with the larger leaf size, also not significantly higher.

*Table 4 - Average values of photosynthesis measurements carried out on the leaves and of the morphological parameters of plants cultivated under single-roof and double-roof with different plastic films. Photosynthetic Activity - PA; Photosynthetic Active Radiation - PAR./ Total stem length - Sl ; stem growth - ΔSg ; length of internodes below last physiologically mature leaf - Li ; stem diameter - Sd ; number of nodes below last physiologically mature leaf - Nn and length of the last leaf at maturity - Lf.*

|                    | Photosynthesis   |                   |                  |                   | Morphology |          |         |         |     |         |
|--------------------|------------------|-------------------|------------------|-------------------|------------|----------|---------|---------|-----|---------|
|                    | PA 10:00-10:30 h | PAR 10:00-10:30 h | PA 11:00-13:00 h | PAR 11:00-13:00 h | Sl (cm)    | ΔSg (cm) | Li (cm) | Sd (cm) | Nn  | Lf (cm) |
| <b>Uncoloured</b>  | 13.5a            | 308.4a            | 14.4ab           | 447.5b            | 457.6a     | 46.5a    | 8.1a    | 11.3a   | 14a | 19.8a   |
| <b>Blue</b>        | 13.2a            | 292.3a            | 16.3b            | 372.8a            | 455.3a     | 46.6a    | 8.9a    | 11.5a   | 15a | 20.5a   |
| <b>Red</b>         | 12.8a            | 308.1a            | 13.1ab           | 410.9a            | 465.3a     | 47.4a    | 7.7a    | 10.9a   | 14a | 19.7a   |
| <b>Single-roof</b> | 10.9a            | 337.4a            | 11.7a            | 424.5a            | 455.2a     | 46.3a    | 8.0a    | 11.1a   | 15a | 18.8a   |

*Different letters in the same column showed statistically significant differences with a confidence degree of 95.0% ( $p < 0.05$ ).*

Although at first sight tomatoes from the blue and uncoloured double-roof seem to have an improved organoleptic and nutritional quality, due to higher values of dry matter and in the Brix degree, it is relevant to point out that these were not validated by the statistical tests. So, these physico-chemical parameters do not evidence a clear improvement in the tomato overall quality by the cultivation under double-roof. In fact, the analysis of the previous data only enables concluding that tomato fruits from blue and uncoloured plastics double-roof have a more appealing look to the consumer, indicated by the chromatic index. Nonetheless, the impact of each covering in the tomato fruit value will be performed by mean of a more detailed metabolome directed analysis. In quantitative terms, every double-roof cover clearly improved the greenhouses' tomato productivity.

In a similar study where tomato plants were grown under coloured photoselective nets, Milenkovi, (2012) also observed an improvement in tomato production for red nets when compared to blue nets. Regarding tomato fruit quality and leaf parameters, Ili et al., (2014) detected higher PAR and leaf size values for tomato plants grown under red plastic. Furthermore, tomatoes grown under blue nets had a higher content of soluble solids and higher maturity and taste indexes, while fruits from red plastics were heavier. All these comparisons are only between blue and red plastic tomatoes.

## 4.4 Multivariate data analysis

The next step was analysing the metabolic modifications occurring in tomatoes along different harvesting weeks and the influence of the plastic covers in the metabolome to get a deeper insight into the impact in fruit quality. Also, the covering effect in the incidence of *Tuta absoluta* was studied through the tomato leaves metabolites. To perform this, the obtained spectra underwent different multivariate statistical analysis techniques.

### 4.4.1 Exploratory analysis (Principal Component Analysis)

First, an exploratory analysis was applied to the 120  $^1\text{H}$ -NMR data of tomato fruit extracts, through the PCA technique to exploit main trends in the  $^1\text{H}$ -NMR data, eliminate outliers and understand which variable induces more discrimination.

The main objective of a PCA exploratory analysis is revealing patterns (grouping or separation) in a set of spectra, that otherwise would not be detectable in the spectroscopic space. The only requirement needed to carry this analysis is a spectra bucket table. A PCA transforms the spectroscopic space into a coordinate space, spanned by Principal Components (PC). The 1<sup>st</sup> PC explains most of the variance, the 2<sup>nd</sup> the second most and successively, as there is a myriad of

PC's. The corresponding 2-dimensional plots, with  $PC_x$  vs  $PC_x$ , are called score plots. To understand, how distinct buckets contributed to the new principal component space there are the loading plots which, importantly, link the statistics provided by the score plots with the metabolite information from the spectra (AMIX, 2006).

As a matter of fact, the goal pretended with the PCA analysis was to identify which variable affects the tomato population the most, its harvesting week, the use of different plastic covers or another unexpected variable that could influence our results. To permit an easier and clear model interpretation each spectrum was labelled according to its characteristic harvesting or its covering.

The PCA scores for the 120  $^1H$  NMR spectra collected from extracts of tomatoes are shown in Figure 15. In fact, the PC's scatter plot which exhibited the better grouping and separation of spectra, relative to their week of harvesting, was PC2 versus PC3, explaining 15% and 10% of the total variance, respectively. Blue dots represent the spectra from the first harvest, and these are clearly separated from the red dots, which represent spectra from the last harvest. Green, light green and yellow dots represent the intermediate harvest, second, third and fourth, respectively, and these are also grouped with their similar and slightly separated from previous and next harvest cluster of spectra. It is observable a successively distribution along the diagonal direction as function of the harvesting time, Figure 15a.

On the other hand, the first PC's scatter plot with spectra indexed by plastic covering shows an aleatory distribution with no detectable patterns, Figure 15c. Discrimination of spectra due to covering type was only observable in the scatter plot between the 6<sup>th</sup> and 7<sup>th</sup> principal components, thus explaining much less total variance between tomato samples. This way, the plastic covering is not the aspect that mostly influences tomato population as the PCA reveals a better discrimination according to the week of harvesting.

In addition, the loading plot, Figure 15b, corresponding to the scatter plots of Figure 15a indicates the  $^1H$  NMR buckets responsible for the different groups of samples. The positioning of objects in a scores plot is directly related with the variables (buckets) in the same direction observed in the loadings plot. For instance, the loadings included in the red square contain metabolites that increase in tomatoes from the last harvest, while the loadings in the blue square increase for tomatoes from the first harvest and these are the major metabolites responsible for the discrimination between spectra.

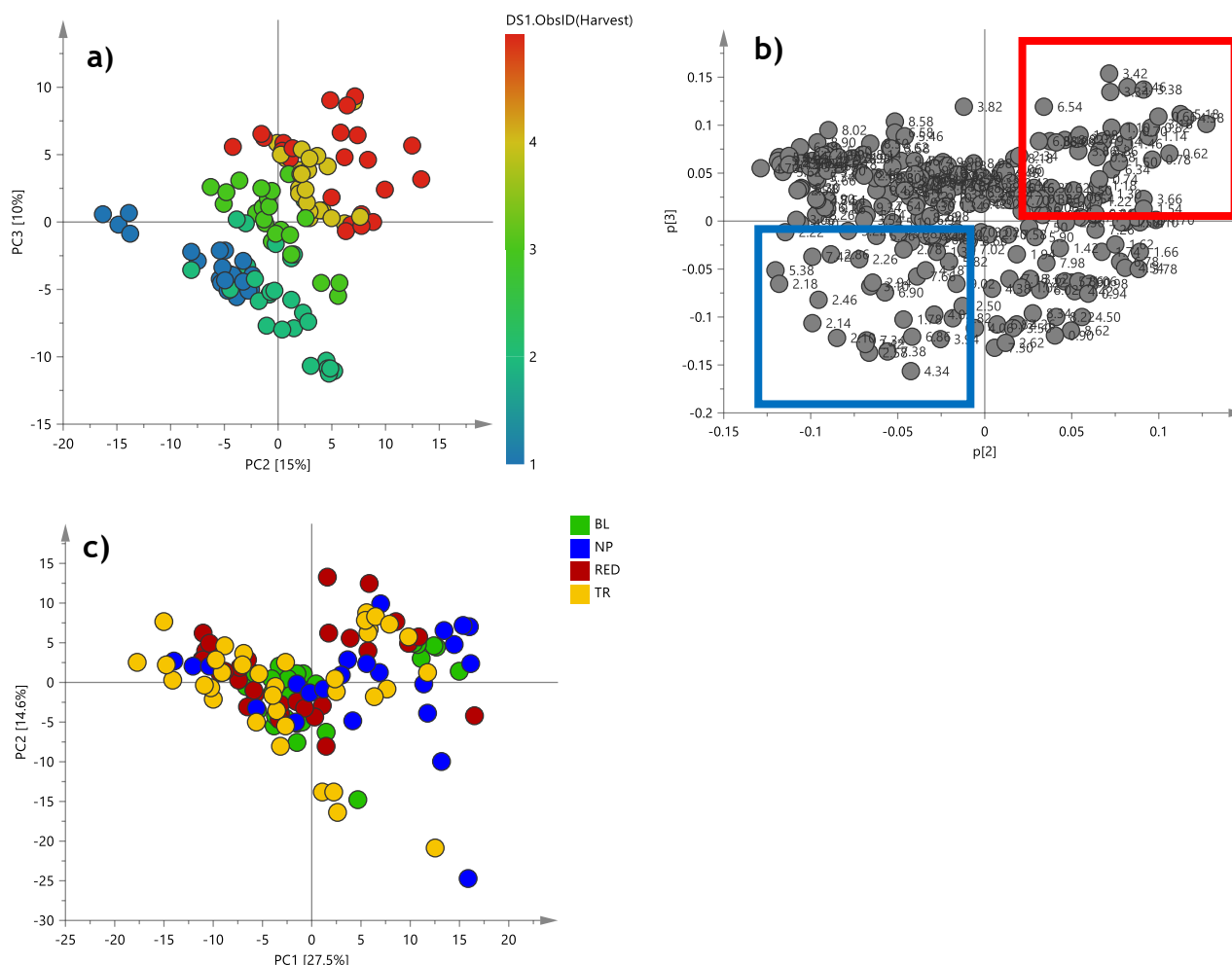


**Figure 15** - (a) PC2/PC3 PCA scores labelled by harvesting time; (b) loadings plots correspondent to the scatter plot in (a). Blue and red squares enclose the buckets which contribute more to the discrimination, for later and first harvests, respectively, and (c) PC1/PC2 PCA scores labelled by plastic covering obtained for  $^1\text{H}$  NMR data of a total of 120 tomato fruit  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (50:50, v/v) extracts covered with different coloured plastics and harvested at five different days (9/01/2020, 22/01/2020, 06/02/2020, 20/02/2020 and 02/03/2020 Almería, Spain). Scaling was done to Univariate. BL- blue plastic double-roof; NP- no double-roof; RED- red plastic double-roof; TR- uncoloured, standard plastic double-roof.

#### 4.4.2 Tomatoes metabolic evolution along the harvests

To improve the discrimination associated with harvesting week and to identify most meaningful, discriminant variables a PLS model was applied to the  $^1\text{H}$  NMR data.

In this supervised approach, the inputs are two tables, X and Y. Table X is the previously mentioned bucket table and table Y, the response matrix, is any type of data regarding a previously known variable under study. If it is a numerical variable, as week of harvesting, the model to be applied is a PLS, in case it is a characteristic or class, the model is named Partial Least Squares Discriminant Analysis (PLS-DA). The aim of a PLS is to find the best correlation

between the two matrixes, X and Y, or how the X variables are influenced by the Y variable. By linear combinations of X and Y, it detects the variance in table X that better fits with Y data (Triba et al., 2015). As it is a highly efficient technique that strongly intensifies group differentiation, it has been consecutively applied in biotechnology, clinical sciences, and importantly, in agri-food studies (Gromski et al., 2015).

The PLS scores plot, Figure 16a, of the first two latent variables exhibited a slight improvement of clustering compared with the PCA scores plot, Figure 15a, as expected. It is noted that spectra are clustered as function of the harvesting week and these clusters are arranged along the first latent variable axis. In fact, tomatoes from the 1<sup>st</sup> and last harvests are clearly distinct in terms of their metabolic profile. Also evident is the tomatoes metabolic evolution along harvesting, as spectra belonging to intermediate harvests are located in the first quadrant for the 4<sup>th</sup> week (yellow dots), around the centre for the 3<sup>rd</sup> week (green dots) and finally in the third quadrant for the 2<sup>nd</sup> week (light green dots). In Figure 16b the loadings plot is shown, what enables a clear observation into the buckets that contribute most to the discrimination of spectra regarding harvesting. Blue and red coloured dots represent the buckets which strongly dictate the separation between spectra based on harvesting time and are the ones with a Variable Importance in Projection (VIP) superior to 1. VIP scores estimate the importance of each variable in the projection used in a PLS model and is often used for variable selection. A variable with a VIP score close to or greater than 1 can be considered important in given model. Variables with VIP scores significantly less than 1 are less important and might be good candidates for exclusion from the model.

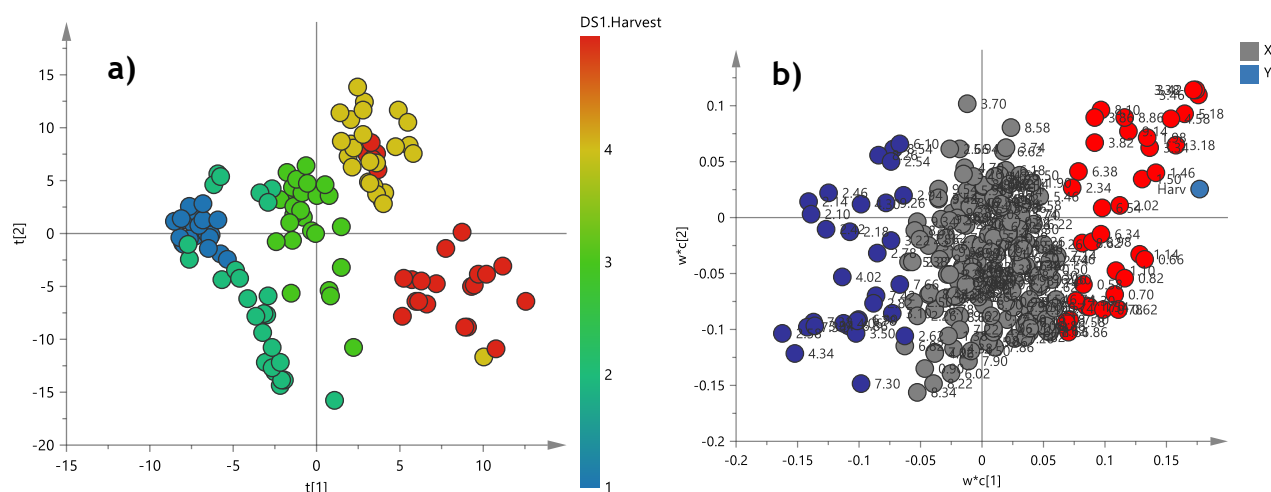


Figure 16 - (a) PLS scores (with  $R^2X=0.592$ ,  $Q^2=0.92$ ), and (b) loadings plot obtained for  $^1H$  NMR data of a total of 120 tomato fruits  $CD_3OD/D_2O$   $KH_2PO_4$  buffer (50:50, v/v) extracts covered with different coloured plastics and harvested at five different days (9/01/2020, 22/01/2020, 06/02/2020, 20/02/2020 and 02/03/2020 Almería, Spain). Scaling was done to Univariate. Blue and red dots in (b) are buckets that scored > 1 in the VIP, for the first and last weeks of harvesting, respectively.

On one hand, by analysing the discriminating loadings for the last weeks of harvest, Table 5, it is possible to state that tomato fruits harvested later showed higher amounts of glucose, lysine, proline, trigonelline, quercetin-like compounds, hydroxycinnamic acid derivatives and fatty acids. On the other hand, the discriminant loadings for the first cuts indicate malic acid, phenylalanine, glutamate, fructose, tyrosine, citric acid, asparagine, *myo*-inositol and adenosine-derivative as the metabolites in higher amounts, Table 5.

*Table 5 - Harvest discriminating metabolites correspondent to the loadings plot in Figure 15b. These metabolites represent buckets that scored > 1 in Variable Importance to Projection.*

| Metabolites that increase in the first harvests | Metabolites that increase in the last harvests |
|-------------------------------------------------|------------------------------------------------|
| Malic acid                                      | Glucose                                        |
| Phenylalanine                                   | Lysine                                         |
| Glutamate                                       | Proline                                        |
| Fructose                                        | Trigonelline                                   |
| Tyrosine                                        | Flavonoids                                     |
| Citric acid                                     | Hydroxycinnamic acid derivates                 |
| Asparagine                                      | Fatty acids                                    |
| Adenosine derivative                            |                                                |
| <i>Myo</i> -inositol                            |                                                |

The analysis of the discriminant loadings showed a decrease on citric and malic acids along the harvesting. Malate and citrate, which account for around 90% of organic acids in ripe fruits, are responsible for the acidity in tomatoes. Organic acids content is highly important because the ratio between sugar/acids, also seen as sweetness/acidity, significantly affects sensorial perception besides indicating a range of quality status upon fruit harvesting (Batista-Silva et al., 2018). In agreement with the discriminant loadings, Anthon & Barrett, (2011) reported an increase in pH and decrease in titratable acidity with the over maturation of tomato fruits, indicating a decrease in acid concentrations. Conversely, with an identical metabolomic approach, Abreu et al., (2019) observed an increase in organic acids and also in the major sugars, glucose and fructose, for the tomatoes harvested later (in a campaign from April to July also at Almería province). Regarding our results, an increase in sugars could be expected since an increase in light intensity occurred (from January to the end of February) and the sugar content is positively correlated with the sunlight that the fruit receives (Turza & Petro-turza, 2009). Nevertheless, only glucose content was increased in tomatoes from last harvest while

fructose showed to decrease. Sugars and acid metabolites are involved in core carbon metabolism as they take part in a variety of key compounds' biosynthetic pathways such as vitamins, phenolics, amino acids and volatile compounds (Batista-Silva et al., 2018; Quinet et al., 2019) and the comprehension of all the metabolic routes and their dynamics can be extremely complex.

The fruits content on flavonoids and hydroxycinnamic acid derivatives, which possess antioxidant properties, showed to increase in the last cuts. The increase in plants' photosynthetic activity leads to the production of sucrose, consequence of higher quantities of sucrose synthase, which is broken down into the hexoses, glucose and fructose, and the accumulation of these sugars can be related to the boost of the fruits' antioxidant defence mechanism, acting as precursors, triggered by the increase in temperature and solar radiation at the end of the crop cycle (Castellano et al., 2007). The increase in phenolic compounds in the last harvest was accompanied by a decrease in their precursor, phenylalanine, which plays a key role in the phenylpropanoid synthesis. We have also observed the increase of some amino acids in tomatoes from the first harvest. Abreu et al., (2019) also reported higher amounts in amino acids for the first harvests when compared to the later ones. In here, glutamate was one of the discriminant amino acids. Interestingly, sugars and organic acids ratio are commonly the taste dictating compounds but high quantities of glutamate are related with the characteristic tomato "umami" taste (María et al., 2010).

Myo-inositol, besides participating in ascorbic acid biosynthesis is also an important precursor for the biosynthesis of several cell wall polysaccharides which are crucial for the fruit proper development. This compound increased in the fruits from the first harvest. Also, an increase in FA was noticed in the last harvest. As mentioned before, these included omega-3 and -6 fatty acids, which contribute to their antioxidant capacity and nutritional value.

Overall, this statistical analysis of 120 <sup>1</sup>H NMR tomato fruit extracts collected in five different weeks indicates the fruits' nutritional value increased along harvesting. Phenolics characterize the metabolome of last harvested tomatoes and these compounds contribute significantly to the fruits' antioxidant activity. Along with the presence of vitamin C and carotenoids, these are the compounds that generate tomatoes' health benefits which, consecutively, drive their worldwide consumption and production.

A table with the full list of discriminating metabolites, with respective bucket and chemical shift in spectra, that scored > 1 in the VIP, is given in Annexes C1.

#### 4.4.3 Tomatoes metabolic profile as a function of the cover colouring

Not all wavelengths of light are beneficial for plant growth and development. It has been well documented that photosynthetic active radiation (PAR, 400-700 nm) directly affects the photosynthesis of plants. The primary photosynthetic pigments in higher plants, chlorophylls *a* and *b*, absorb in the blue (peak near 430 nm) and red regions (peak near 660 nm), while having very little absorption in the green region. For this reason, the greenhouse industry has been applying photoselective nets to increase blue and red light inside greenhouses in order to enhance photosynthesis and productivity (Rajapakse & Shahak, 2007).

Herein, the metabolic changes of tomato fruits produced by plastics of different colours were investigated. Both PLS-DA and Orthogonal Projection to Latent Structures Discriminant Analysis (OPLS-DA) supervised models were applied to the  $^1\text{H}$  NMR data set to assess if there are metabolic differences between tomatoes grown under uncoloured, blue and red plastic double-roof and tomatoes grown under the usual plastic single-roof (control).

The OPLS-DA method is an extension of the PLS regression technique which provides clearer models, easier to interpretate. Basically, OPLS removes the structured-noise from the input data and separates the systematic variation in matrix *X* into two components: variability correlated (linear) and uncorrelated (orthogonal) to *Y* (Vajargah et al., 2012).

As seen in Figure 17b, the OPLS-DA analysis shows a more precise grouping of the spectroscopic data (*X*) according with their distinct treatment or covering condition (*Y*), when compared to the PLS-DA model, Figure 17a. Indeed, the quality parameters also indicate the OPLS model as a better fitting to the data ( $R^2$  closer to 1) and better predictive ability (higher  $Q^2$ ). The closer to 1  $R^2$  is, the better is the data description by the model, while the closer to 1  $Q^2$  is, the better predictability. Although it is complex to give a limit that represents a good predictability as it is highly dependent on the dataset properties, a  $Q^2 > 0.5$  generally is a signal of good predictability (Triba et al., 2015).

In fact, with the OPLS model applied to the  $^1\text{H}$  NMR data of tomato fruit extracts a clear discrimination could be observed among tomatoes grown under different covering conditions, Figure 17b. Fruits from uncoloured double-roof are evidently separated from fruits grown with single-roof, along the *X*-axis, what reveals a profound impact of the double-cover plant farming in the fruits' metabolome. On the other hand, along the *Y*-axis it is explicit the distinction between tomatoes from blue and red plastics double-roof, indicating that the plants' dynamic metabolism is clearly influenced by the covering plastics colour.

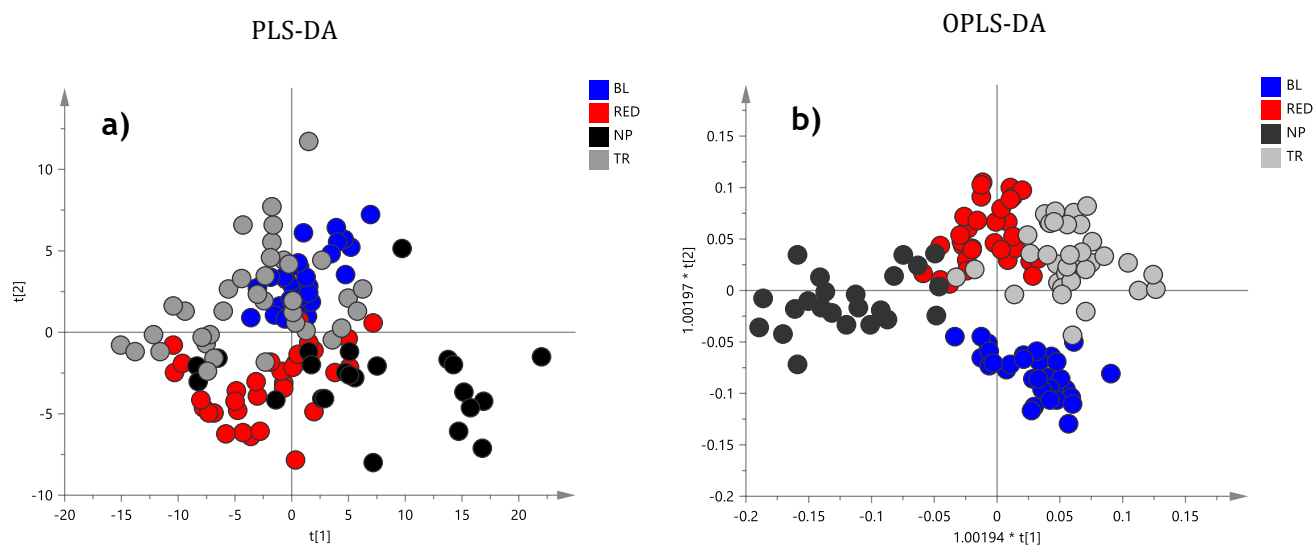


Figure 17 - (a) PLS-DA scores (with  $R^2X=0.828$ ,  $Q^2=0.582$ ), and (b) OPLS-DA (with  $R^2X=0.934$ ,  $Q^2=0.666$ ) obtained for  $^1H$  NMR data of a total of 120 tomato fruits  $CD_3OD/D_2O$   $KH_2PO_4$  buffer (50:50, v/v) extracts covered with different coloured plastics and harvested at five different days (9/01/2020, 22/01/2020, 06/02/2020, 20/02/2020 and 02/03/2020 Almería, Spain). Scaling was done to Univariate for PLS-DA model and Pareto for OPLS-DA model. BL- blue plastic double-roof; NP- no double-roof; RED- red plastic double-roof; TR- uncoloured, standard plastic double-roof.

To obtain an overall picture and compare the tomato fruits' metabolome, induced by each type of plant covering, a hierarchically clustered heatmap, Figure 18, was obtained from a total of 53 discriminating buckets containing significantly altered tomato metabolites. The presented heatmap was built considering the mean values obtained for the five harvests. In Annexes C another heatmap is provided with the values for each harvest week. Clustered heatmaps are a very effective tool for displaying feature variation among groups of samples. Its basic concept is to represent relationships among variables as a colour scale. Actually, it enables a simple and clear visualization of the shifting patterns in metabolites concentrations across samples and/or experimental conditions under study (Xia et al., 2015).

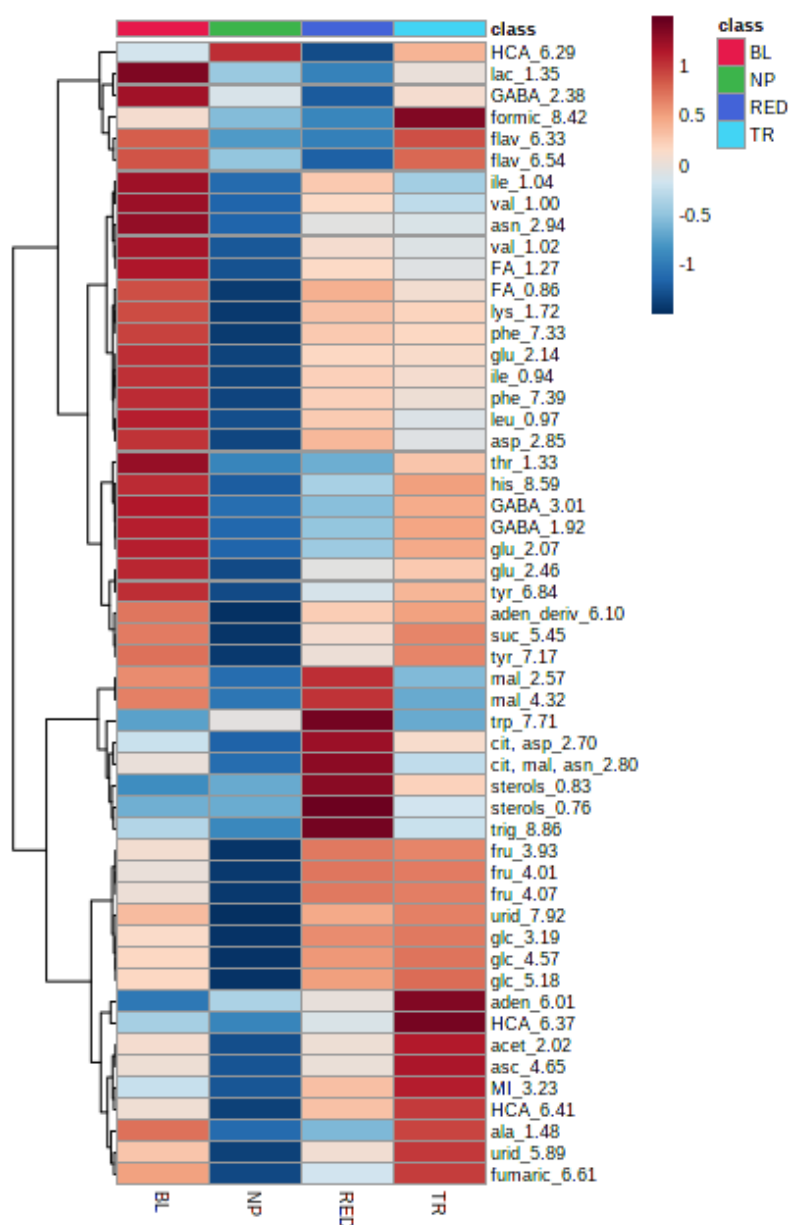


Figure 18- Heatmap showing z-scores of 53 discriminatory buckets from  $^1\text{H}$  NMR data of a total of 120 tomato fruits  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (50:50, v/v) extracts and their assigned metabolites for the four different experimental plastic coverings. The heatmap construction was based on Euclidean distances and Ward clustering. The X axis represents the average of thirty biological replicates for each plastic experiment (BL-blue plastic double-roof; NP-no plastic; RED-red plastic double-roof; TR-uncoloured, standard plastic double-roof) from all harvesting stages. The colour scheme represents the variation on the metabolic content among the four conditions in a scale from dark blue (lowest content) to dark red (highest content). In this case, spectroscopic data underwent variable size bucketing. Abbreviations: acet, acetate; aden, adenosine; ala, alanine; asc, ascorbate; asn, asparagine; asp, aspartate; cit, citric acid; FA, fatty acids; flav, flavonoids; formic, formic acid; fru, fructose; fumaric, fumaric acid; GABA,  $\gamma$ -aminobutyric acid; glc, glucose; HCA, hydroxycinnamic acid derivatives; his, histidine; ile, isoleucine; lac, lactate; leu, leucine; lys, lysine; mal, malic acid; MI, myo-inositol; phe, phenylalanine; pro, proline; sterols; suc, sucrose; thr, threonine; trig, trigonelline; trp, tryptophan; tyr, tyrosine; urid, uridine; val, valine.

By analysing the clustered heatmap, Figure 18, a positive effect from cultivation under double-roof is clearly observed. Tomato fruits cultivated this way present a metabolic content quantitatively much richer than tomatoes grown under the common greenhouse layout.

Tomatoes grown in single-roof only reported the highest quantity of one compound, one hydroxycinnamic acid (loading 6.29), in relation with the remaining fruits grown under double-roof. Comparatively, all the remaining metabolites are present in much smaller quantities suggesting the winter cultivation under double-roof is a better option for tomato plants, enabling the synthesis of fruits with an increased organoleptic and nutritional value.

The fruits belonging to red coloured double-roof showed the highest quantities of malic and citric acid, tryptophane, trigonelline, asparagine and sterols. These tomatoes seem to be the more acidic, consequence of a slower maturation process as observed before in Table 3 where the fruits from red covering scored the lowest in the chromatic index, proportional to maturation, being for the same harvest significantly less ( $p < 0.05$ ) mature than the ones from uncoloured and blue plastics. Moreover, the fruits from red plastic share the highest values in fructose with the ones from uncoloured plastic covering and are second in major sugars total content. Regarding the antioxidant compounds, low quantities were observed for flavonoids, hydroxycinnamic acids and ascorbate.

Tomatoes from blue coloured double-roof revealed, comparatively, the highest amounts on several amino acids (GABA, isoleucine, valine, asparagine, threonine, glutamate, phenylalanine, lysine, leucine and asparagine), flavonoids, and fatty acids. Also, an increase in lactic acid was detected. Overall, these fruits exhibit a very good nutritional value as a source of nutritional essential amino acids and flavonoids, in addition with a probable improvement in fruits' taste due to the abundance of glutamate (Choi et al., 2011; Elbadrawy, 2011).

Tomatoes grown under the uncoloured plastic double-roof presented the highest values of fructose and glucose from the four growth conditions, as well as ascorbate (vitamin C), *myo*-inositol, hydroxycinnamic acids, among others. Also, these have high levels of flavonoids and amino acids (GABA, glutamate, threonine, tyrosine and histidine). Importantly, the high quantities of hydroxycinnamic acid derivatives, flavonoids and vitamin C turn this cultivation practice the best source of antioxidant compounds, which possess several health benefits to the consumers.

As a matter of fact, the investigation of the hierarchical clustered heatmap agrees with the results obtained by the ORAC method, Figure 14. The antioxidant capacity of tomato fruits for each plastic covering obtained by the ORAC method is proportional to the fruits' content in hydroxycinnamic acids, flavonoids and fatty acids exhibited by the heatmap.

While the quality parameters shown in Section 3.3, did not show meaningful differences between the different growing conditions, herein fruits from the double-roof farming seem to have an overall quantitatively and qualitative richer metabolic content. In the end, tomato plants from uncoloured and blue plastic double-roof greenhouses produced tastier fruits with enhanced antioxidant capacity, rendering these covering plastics as the preferred cultivation method in terms of the overall fruit quality.

#### 4.4.4 Plastic cover colouring effect on leaves defence metabolites

In this section, the effect of different coloured plastic coverings on the metabolic content of tomato leaves and its relationship with *Tuta absoluta* incidence was investigated.

To make possible the multivariate data analysis, leaves from each condition, collected in the same day, were extracted with a methanol 80:20 water (v/v) solvent and measured by  $^1\text{H}$  NMR. As the defence compounds synthesized in leaves frequently emit resonance in the downfield aromatic region from  $\delta_{\text{H}}$  5.5 to 10 ppm, compound identification and multivariate analysis were only performed in this region of the spectroscopic data (Zhao et al., 2016).

In fact, NMR-based metabolomics have been widely used to decode the broad range of chemical compounds that might be implicated in the plant defence against biotic stress, mainly the secondary metabolites related with plants resistance to pests and diseases (Verpoorte et al., 2007). Moreover, NMR as a tool to identify the secondary metabolites each plant species produces to counter herbivores and pathogens, has the potential to help in the engineering of resistant cultivars (Leiss et al., 2011).

First, to identify if there was discrimination according to different coverings and analyse eventual discriminant variables a PLS-DA was applied to the  $^1\text{H}$  NMR data of leaves extracts. The PLS-DA scores plot, Figure 19, of the first two latent variables indicates that leaves belonging to plants cultivated in single-roof are clearly metabolically distinct to the ones cultivated under a double roof, either coloured or uncoloured. So, as a clear discrimination was found, the peak assignment for metabolites in the aromatic region was performed to further uncover the loadings responsible for such separation and identify possible biomarkers correlated to a major incidence of *Tuta* infection.

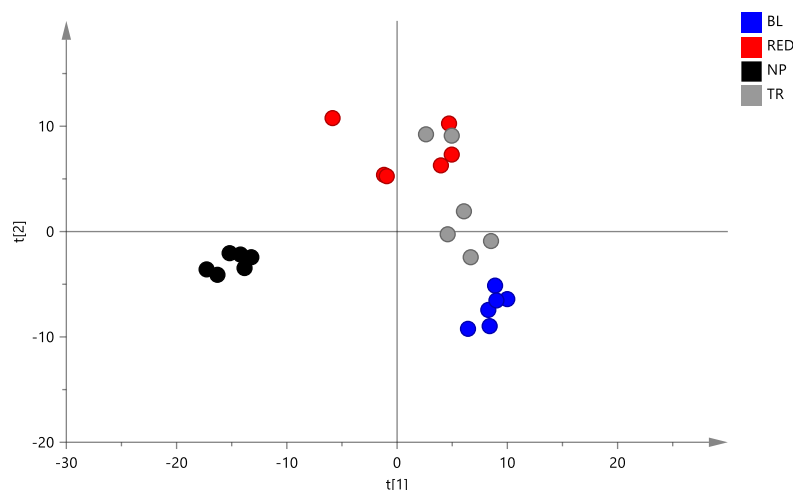


Figure 19 - PLS-DA scores (with  $R^2X = 0.907$ ,  $Q^2 = 0.926$ ) obtained for  $^1\text{H}$  NMR data of a total of 24 tomato leaves  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (80:20, v/v) extracts covered with different coloured plastics from the same harvest (02/03/2020 Almería, Spain). Scaling was done to Univariate. BL- blue plastic double-roof; NP- no double-roof; RED- red plastic double-roof; TR- uncoloured, standard plastic double-roof.

The list with the discriminating metabolites identified in the leaf extract is provided in Table 6, detailing for each signal its chemical shifts ( $\delta_{\text{H}}$ , ppm), multiplicity and coupling constants ( $J$ , Hz). Identification and characterization were based in the interpretation of homo- and heteronuclear 1D and 2D NMR experiments ( $^1\text{H}$ -NOESY,  $^1\text{H}$ - $^1\text{H}$  COSY,  $^1\text{H}$ - $^1\text{H}$  TOCSY,  $^1\text{H}$ - $^{13}\text{C}$  HSQC and  $^1\text{H}$ - $^{13}\text{C}$  HMDB) combined with browsing in NMR databases such as HMDB, Spectral Reference Libraries from Chenomx software and literature data (Abreu et al., 2019; López-gresa et al., 2012, 2010).

Table 6 - Peak assignment of metabolites (limited to the aromatic region) identified in the tomato leaves  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (80:20, v/v) extracts by NMR.

| Metabolite                     | Chemical shifts (ppm), multiplicity and coupling constants (Hz)                         |
|--------------------------------|-----------------------------------------------------------------------------------------|
| <b>Other organic molecules</b> |                                                                                         |
| 1 Tyrosine                     | 6.79 (d, $J = 8.66$ Hz), 7.06 (d, $J = 8.63$ Hz)                                        |
| 2 Tryptophan                   | 7.07 (s), 7.24 (s), 7.3 (m), 7.64 (d, $J = 9.2$ Hz)                                     |
| 3 Phenylalanine                | 7.30 (m), 7.34 (m)                                                                      |
| 4 Formate                      | 8.39 (s)                                                                                |
| 5 Trigonelline                 | 8.08 (dd, $J = 8.4$ ; 6.3 Hz), 8.85 (d, $J = 6.3$ Hz), 8.89 (d, $J = 8.0$ Hz), 9.17 (s) |

**Phenylpropanoids and flavonoids compounds**

|   |                                   |                                                                                                                               |
|---|-----------------------------------|-------------------------------------------------------------------------------------------------------------------------------|
| 6 | Rutin                             | 6.27 (d, $J = 2.1$ Hz), 6.47 (d, $J = 2.1$ Hz), 6.93 (d, $J = 8.54$ Hz), 7.62 (dd, $J = 8.4, 2.2$ Hz), 7.67 (d, $J = 2.1$ Hz) |
| 7 | Hydroxycinnamic acid derivative 1 | 6.41 (d, $J = 15.9$ Hz), 6.79 (d, $J = 8.5$ Hz), 7.0 (d, $J = 8.5$ Hz), 7.10 (dd, $J = 8.5, 2.1$ Hz), 7.57 (d, $J = 16$ Hz)   |
| 8 | Hydroxycinnamic acid derivative 2 | 6.43 (d, $J = 15.9$ Hz), 6.82 (d, $J = 8.2$ Hz), 6.98 (dd, $J = 8.54, 1.9$ Hz), 7.12 (d, $J = 2.1$ Hz)                        |

Then, the contribution plot relative to the leaves grown under the control condition (single plastic roof) was analysed, Figure 20, to examine the metabolic evidence supporting the clear separation from the remaining covering conditions, observed in the PLS-DA scores. Actually, as highlighted in Figure 19, several buckets characteristic from the hydroxycinnamic acids and the flavonoid scored  $>1$  in the VIP, what means these were very important to the model and for the verified discrimination. It is known that these kind of phenolic compounds play a key role in the plants' defence strategy (Dixon et al., 2002).

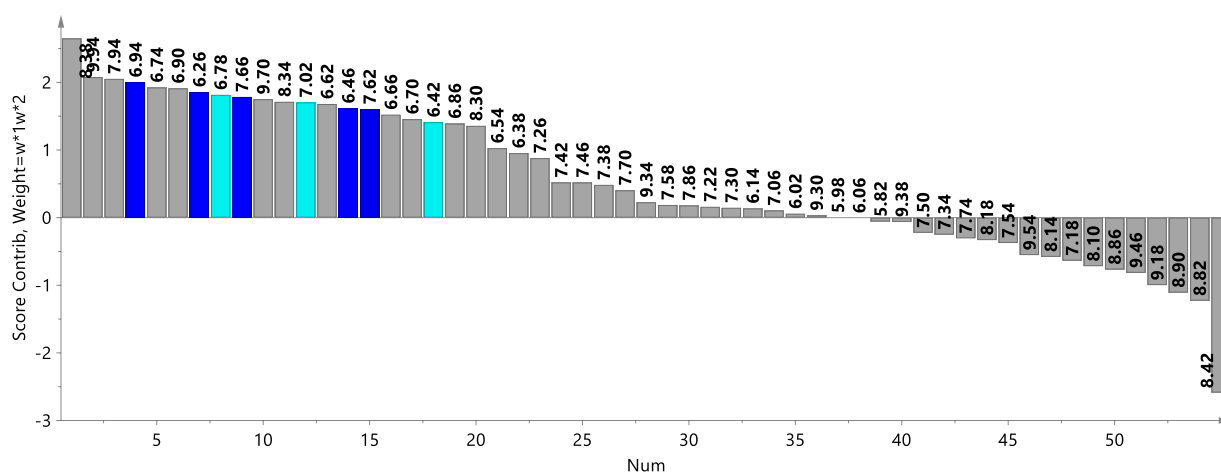


Figure 20 - PLS-DA contribution scores (with  $R^2X = 0.907$ ,  $Q^2 = 0.926$ ) obtained for  $^1H$  NMR data of a total of 24 tomato leaves  $CD_3OD/D_2O$   $KH_2PO_4$  buffer (80:20, v/v) extracts, showing the loadings containing metabolites that increased (upper bars) and decrease (down bars) in leaves from single-roof covering respect to the other plastic-covers. Buckets from the flavonoid and hydroxycinnamic acid derivatives are highlighted, in dark and light blue, respectively. From left to right, in descendent order, these eight coloured buckets scored  $> 1$  in the VIP.

The accurate identification of the hydroxycinnamic acid derivatives was not possible with the NMR experiments carried out. Although, the chemical shifts and coupling constants for the two compounds described in Table 6 match with those from chlorogenic and ferulic acid, which were confirmed by López-gresa et al., (2010) as major hydroxycinnamic acids in tomato leaves.

In fact, upon exposure and damage by herbivores, several plants respond by boosting the phenylpropanoid metabolism, accumulating newly formed phenolics, among them, chlorogenic and ferulic acid (Bernards & Lars, 2008).

Chlorogenic acid has proven to exhibit a broad spectrum anti-herbivore activity, besides having also a negative effect on bacteria, fungi and virus that frequently infect plants (Leiss et al., 2020). This compound is found particularly in the *Solanaceae* family, tomato, potato and eggplant, although present in many other plant species and acts as an anti-nutritive agent, in the case of herbivore insect attacks. It has also been verified that plants with higher levels of chlorogenic acid are less prone to be chosen by herbivores to feed on and as consequence, the metabolic engineering of important *Solanaceae* crops with increased levels of these compounds has been proposed to enhance protection from pathogens and improve crop yields (Kundu & Vadassery, 2019).

Regarding ferulic acid, it has been reported its significant antifungal activity and its relevance in diminishing the incidence of powdery mildew in cucumber. In addition, the compound has shown to act as an oviposition deterrent for *Choristoneura fumiferana* and its high concentration in wheat cultivars was linked with the resistance to *Sitodiplosis mosellana* spread (Hucl et al., 2001; Widarto & Meijden, 2006).

Lastly, the flavonoid rutin is one of the most studied flavonoid glycosides and it has also been related with the improvement of the defence system when facing some stress conditions (Suzuki et al., 2005). Indeed, it has detrimental effects on the development of early stages of larvae from several insects (Simmonds, 2001).

The visual difference between the metabolome that characterizes leaves from plants grown under distinct coverings can be seen in Figure 21. It is represented the phenolic region of the  $^1\text{H}$  NMR spectra from leaves collected in the same day for the four types of covering. On one hand, spectra from double covered leaves presented similar small quantities of compounds involved in plants' defence against pathogens' attack. On the other hand, in the spectrum of leaves grown in the conventional greenhouse layout, a standard plastic single covering, characteristic peaks from those compounds exhibited a considerably higher area and thus a higher quantity.

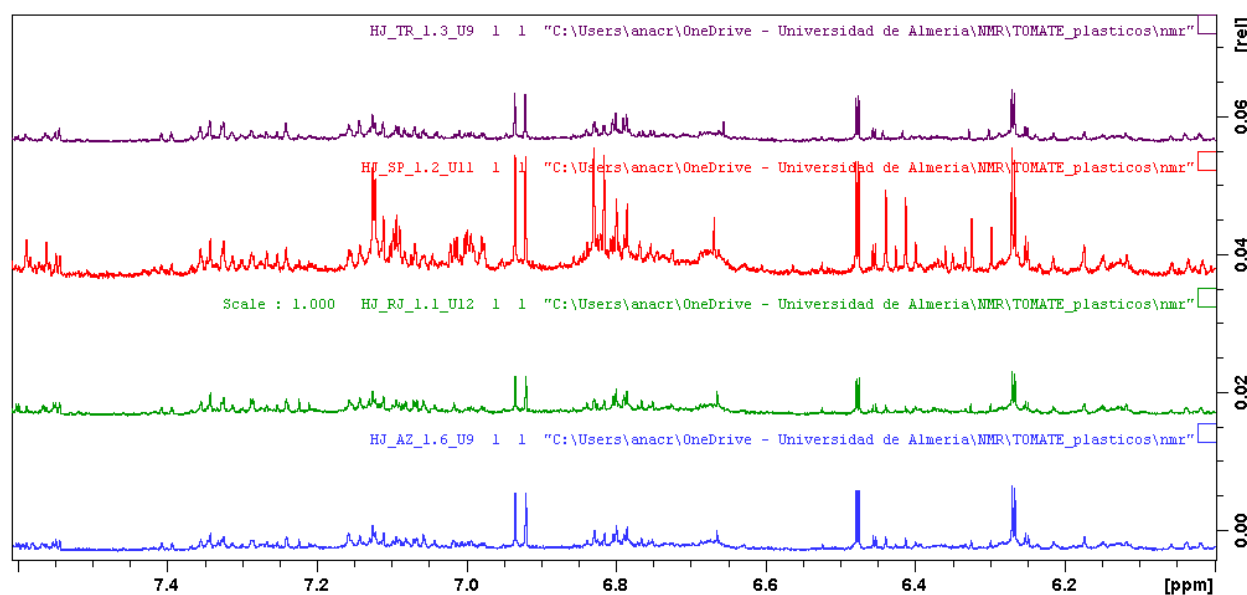


Figure 21 - 1D  $^1\text{H}$ -NMR spectra of tomato leaf samples zoomed in the phenolic region. From top to bottom, displayed spectra are from leaves collected in uncoloured double-roof greenhouses, without double-roof, red double-roof and blue double-roof. Spectra were obtained in a Bruker Advance III 600 MHz spectrometer and visualized in Bruker TopSpin 3.6.2.

As a matter of fact, López-gresa et al., (2010, 2012) investigated the metabolic changes between leaves from healthy plants and leaves from plants infected with *Pseudomonas syringae* pv. tomato, citrus exocortis viroid and tomato mosaic virus, detecting also a higher synthesis of flavonoids and phenylpropanoids for the diseased leaves. In the same line, Falco et al., (2019) studied the interaction between *Tuta absoluta* and tomato leaves, noting once again the triggering of metabolic routes that lead to the production of these organic compounds for plants infested with the pathogen. This renders the leaves metabolic profile as a possible indicator of the plants' health or disease extent condition, as it reveals the end product of changes in metabolomic pathways induced by infections.

As previously said, and according to bibliography, the presence of a higher content of such biologically active compounds in the plants grown under single-roof suggests that these plants were being more attacked by *Tuta absoluta*, than the other plants, as the metabolic pathways implied in the synthesis of defence secondary metabolites were more active. Indeed, this farming condition also revealed the lowest value in marketable production (Table 3). However, these results were not in accordance with the ones provided by the method employed in the insect counting, Figure 22.

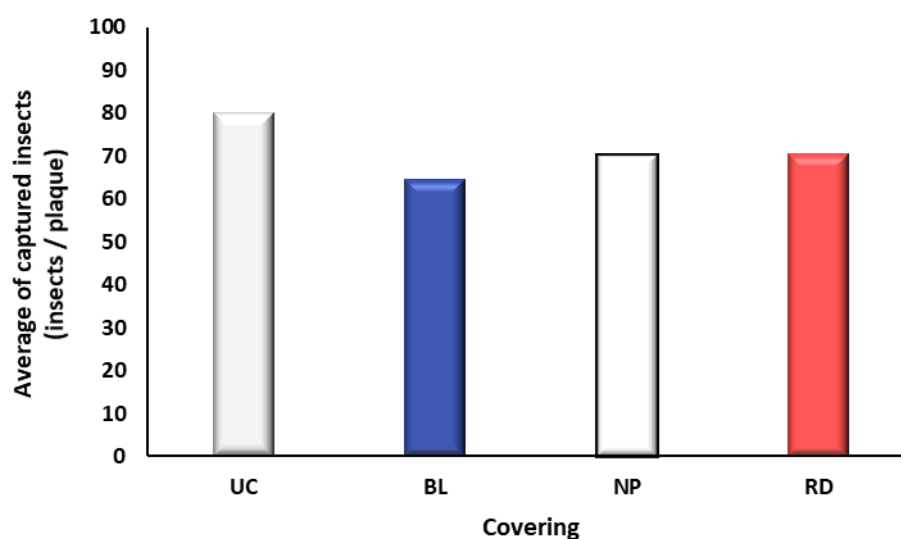


Figure 22 - Average values of *Tuta absoluta* individuals captured in the chromatic trap for all plastic coverings. Uncoloured standard plastic double-roof (UC), blue photo-selective plastic double-roof (BL); without double plastic covering (NP) and red photo-selective plastic double-roof (RD). Insects were counted in 29/01/2020, 05/02/2020 and 12/02/2020.

This method consisted in a black chromatic trap covering the substrate container, as this insect is attracted by the black colour. The number of captured insects was similar for every plastic covering condition, but this result may not be the most accurate. In fact, the metabolomic analysis has a higher scientific rigour enabling the conclusion that tomato plants cultivation under double-roof decreased the incidence of the pest. Also, the increase in plant-defence compounds may be due to the presence of other pests in the plants besides *Tuta absoluta*, which were not considered in this study. A careful analysis of all the pests that may affect the plant and may boost their defence system should be considered in further studies.

## 5 Conclusion and future recommendations

First, in terms of greenhouse productivity, double-roof farming enhanced tomato fruits total and marketable production. The biggest value of marketable tomato fruits per square meter was found for the greenhouse with red plastic covering.

In the ORAC assay, fruits grown under uncoloured double-roof revealed an antioxidant activity significantly higher than the ones grown under the remaining coverings, being followed in order by the tomatoes grown under blue double-roof, red double-roof and single-roof.

The exploratory analysis of the 120 tomato fruits  $^1\text{H}$  NMR spectra by PCA revealed that major metabolic changes in tomato metabolome were caused by the week of harvesting rather than by type of covering. In fact, the PLS model applied to identify the metabolites responsible for such discrimination revealed high values of organic acids and amino acids in tomatoes from the first harvests. On the other hand, high values of glucose, flavonoids and hydroxycinnamic acids were detected for the last cuts. So, tomatoes harvested later are nutritionally richer, being a better source of antioxidant compounds.

Regarding the covering and colour effect, it was possible to observe how an OPLS model is easier to interpret, exhibiting higher quality parameters ( $R^2X$  and  $Q^2$ ) compared to a PLS. The heatmap clearly pointed out that winter double-roof tomato growth is advantageous as the fruits displayed a richer metabolome. For instance, tomatoes from uncoloured double-roof exhibited the highest levels of flavonoids, hydroxycinnamic acid derivatives, vitamin C, major sugars and followed the ones from blue covering in terms of essential amino acids. Fruits covered with blue plastic produced higher amounts of free amino acids and were second in terms of antioxidant compounds. Tomatoes from red plastic covering reported the highest value in organic acids, exhibiting very low content in antioxidants. Finally, the tomatoes from single-roof had the poorest metabolome, having comparatively the lowest amount for almost every metabolite. In fact, the results obtained by the ORAC method are proportional to the levels of antioxidants verified in the heatmap. This way, uncoloured and blue plastic double-roof are the preferred agronomic practices in terms of fruit overall quality as the tomatoes exhibit high contents in essential amino acids with the added value of presenting higher antioxidant properties. On the other hand, as suggested by the low scoring in the chromaticity index (Table 3) tomatoes from red plastics seem to have had a slower maturation due to the higher amounts of organic acids.

Lastly, relative to the covering influence on the incidence of *Tuta absoluta*, the PLS-DA applied to the aromatic region of the tomato leaves revealed a clear separation between leaves spectra from single-roof to the remaining spectra from the three double-roofs. As a matter of fact,

leaves from single-roof produced higher amounts of key phenolic compounds involved in the defensive strategy, indicating that these were more exposed to the herbivore insect attack. Once again, tomato plants' farming under double-roof exhibited better results and the lower incidence of the insect might be, to some extent, responsible for the higher value found for the fruits' marketable production under this condition.

As a final remark, the application of photoselective plastic coverings over the tomato plants improved greenhouse productivity, the phytochemical quality of tomato fruits and provided a better control on a damaging pest. The plastic pigments differentially absorb radiation wavelengths promoting distinct physiological responses, both for the plant and the insect, and the covering also increases the scattering light, thus improving light penetration into the inner parts of plant canopy (Cástilla, 2012; Castilla & Montero, 2008).

This technique represents a non-chemical alternative that could regulate plants' growth and decrease pests' incidence, which can reduce long term production costs, health risks to workers and consumers, and potential environmental pollution.

Studies of this kind have a great impact considering the national and international tomato fruits demand, as well as the importance of the agriculture sector to the city of Almería. The increasing interest and improvement of omic technologies offers new, hopeful opportunities to the agri-food research that can translate into agricultural practices with higher yields and product quality.

Furthermore, to attain a wider coverage of the tomato fruits metabolome, lipophilic extracts could be analysed through NMR. These can be obtained with a more apolar solvent, such as methanol-chloroform ( $\text{CD}_3\text{OD-CDCl}_3$ , 80:20 v/v), allowing the extraction and identification of fatty acids, sterols and carotenoids.

A more complete and precise analysis of the best farming condition can be obtained if data from all possible combinations of covering/greenhouse layout is gathered and underwent further analysis. Herein, the investigation was directed to the type of covering regardless of the respective greenhouse characteristics, and these must also be considered to achieve a thorough analysis.

Further investigations are required to study the influence of the photoselective nets in the tomato plants and in the insect, regarding light quality and consequent physiological responses. Potentially, with the best light manipulation method, the desired balance between productivity and the tomato fruits' organoleptic and nutritional value can be obtained.

## 6 References

- Abreu, A. C., & Fernández, I. (2020). NMR Metabolomics Applied on the Discrimination of Variables Influencing Tomato (*Solanum lycopersicum*). *Molecules*, 25(16). <https://doi.org/10.3390/molecules25163738>
- Abreu, A. C., Marín, P., Aguilera-Sáez, L. M., Tristán, A. I., Peña, A., Oliveira, I., Simões, M., Valera, D., & Fernández, I. (2019). Effect of a Shading Mesh on the Metabolic, Nutritional, and Defense Profiles of Harvested Greenhouse-Grown Organic Tomato Fruits and Leaves Revealed by NMR Metabolomics. *Journal of Agricultural and Food Chemistry*, 67(46), 12972-12985. <https://doi.org/10.1021/acs.jafc.9b05657>
- Agrónoma. (2019). *La producción de tomate en Almería lleva cayendo cuatro campañas consecutivas*. Retrieved February 19, 2021, from <https://sevilla.abc.es/agronoma/noticias/cultivos/tomate/cae-produccion-tomate-almeria-ultimas-campanas/?ref=https%3A%2F%2Fwww.google.es%2F>
- Albida Agriculture . (2018). Retrieved February 19, 2021, from <https://www.albidaagriculture.com/pest-control>
- Alfosea-Simón, M., Simón-Grao, S., Zavala-Gonzalez, E. A., Cámara-Zapata, J. M., Simón, I., Martínez-Nicolás, J. J., Lidón, V., & García-Sánchez, F. (2021). Physiological, Nutritional and Metabolomic Responses of Tomato Plants After the Foliar Application of Amino Acids Aspartic Acid, Glutamic Acid and Alanine. *Frontiers in Plant Science*, 11(January), 1-16. <https://doi.org/10.3389/fpls.2020.581234>
- Allwood, J. W., De Vos, R. C. H., Moing, A., Deborde, C., Erban, A., Kopka, J., Goodacre, R., & Hall, R. D. (2011). Plant metabolomics and its potential for systems biology research: Background concepts, technology, and methodology. In *Methods in Enzymology* (1st ed., Vol. 500). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-385118-5.00016-5>
- Alves Filho, E. G., Silva, L. M. A., Ribeiro, P. R. V., de Brito, E. S., Zocolo, G. J., Souza-Leão, P. C., Marques, A. T. B., Quintela, A. L., Larsen, F. H., & Canuto, K. M. (2019). <sup>1</sup>H NMR and LC-MS-based metabolomic approach for evaluation of the seasonality and viticultural practices in wines from São Francisco River Valley, a Brazilian semi-arid region. *Food Chemistry*, 289(March), 558-567. <https://doi.org/10.1016/j.foodchem.2019.03.103>
- Amberg, A., Riefke, B., Schlotterbeck, G., Ross, A., Senn, H., Dieterle, F., & Keck, M. (2017). NMR and MS methods for metabolomics. *Methods in Molecular Biology*, 1641, 229-258. [https://doi.org/10.1007/978-1-4939-7172-5\\_13](https://doi.org/10.1007/978-1-4939-7172-5_13)
- AMIX. (2006). *Amix-Viewer & Amix Software Manual*. C, 392.

- Anthon, G. E., & Barrett, D. M. (2011). *Changes in pH , acids , sugars and other quality parameters during extended vine holding of ripe processing tomatoes. December 2010*, 1175-1181. <https://doi.org/10.1002/jsfa.4312>
- Baldina, S., Picarella, M. E., Troise, A. D., Pucci, A., Ruggieri, V., Ferracane, R., Barone, A., Fogliano, V., & Mazzucato, A. (2016). Metabolite profiling of Italian tomato landraces with different fruit types. *Frontiers in Plant Science*, 7(MAY2016), 1-13. <https://doi.org/10.3389/fpls.2016.00664>
- Barnes, S., Benton, H. P., Casazza, K., Cooper, S. J., Cui, X., Du, X., Engler, J., Kabarowski, J. H., Li, S., Pathmasiri, W., Prasain, J. K., Renfrow, M. B., & Tiwari, H. K. (2016). Training in metabolomics research. I. Designing the experiment, collecting and extracting samples and generating metabolomics data. *Journal of Mass Spectrometry*, November 2015, 461-475. <https://doi.org/10.1002/jms.3782>
- Bartel, J., Krumsiek, J., & Theis, F. J. (2013). Statistical methods for the analysis of high-throughput metabolomics data. *Computational and Structural Biotechnology Journal*, 4(5), e201301009. <https://doi.org/10.5936/csbj.201301009>
- Batista-Silva, W., Nascimento, V. L., Medeiros, D. B., Nunes-Nesi, A., Ribeiro, D. M., Zsögön, A., & Araújo, W. L. (2018). Modifications in organic acid profiles during fruit development and ripening: Correlation or causation? *Frontiers in Plant Science*, 871(November), 1-20. <https://doi.org/10.3389/fpls.2018.01689>
- Baudoin, W., Nono-Womdim, R., Lutaladio, N., Hodder, A., Castilla, N., Leonardi, C., Pascale, S. De, & Qaryouti, M. (2013). Good Agricultural Practices for greenhouse vegetable crops. In *Good Agricultural Practices for greenhouse vegetable crops*. <https://doi.org/10.1201/b13737-8>
- Baumann, N. (2016). How to use the medical subject headings (MeSH). *International Journal of Clinical Practice*, 70(2), 171-174. <https://doi.org/10.1111/ijcp.12767>
- Beckonert, O., Keun, H. C., Ebbels, T. M. D., Bundy, J., Holmes, E., Lindon, J. C., & Nicholson, J. K. (2007). Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nature Protocols*, 2(11), 2692-2703. <https://doi.org/10.1038/nprot.2007.376>
- Bellán, C. L. U., Os, M. T. M. E., Afrilla, M. P. Z., Ortea, M. I. F., Abaldón, J. A. G., & Elicado, E. N. Ú. (2008). *ORAC-Fluorescein Assay To Determine the Oxygen Radical Absorbance Capacity of Resveratrol Complexed in Cyclodextrins*. 2254-2259.
- Bernards, M. A., & Lars, B. (2008). *Phenylpropanoid Metabolism Induced by Wounding and Insect Herbivory*. 189-211.

- Biais, B., Bernillon, S., Deborde, C., Cabasson, C., Rolin, D., Tadmor, Y., ... Moing, A. (2011). Precautions for Harvest, Sampling, Storage and Transport of Crop Plant Metabolomics Samples. *Plant Metabolomics*.
- Botoran, O. R., Ionete, R. E., Miricioiu, M. G., Costinel, D., Radu, G. L., & Popescu, R. (2019). Amino acid profile of fruits as potential fingerprints of varietal origin. *Molecules*, 24(24), 1-11. <https://doi.org/10.3390/molecules24244500>
- Brandt, S., Pék, Z., Barna, É., Lugasi, A., & Helyes, L. (2006). Lycopene content and colour of ripening tomatoes as affected by environmental conditions. *Journal of the Science of Food and Agriculture*, 86(4), 568-572. <https://doi.org/10.1002/jsfa.2390>
- Budzynska, B., Faggio, C., Kruk-Slomka, M., Samec, D., Nabavi, SF., Sureda, A., Devi, KP., Nabavi SM. (2019). Rutin as Neuroprotective Agent: From Bench to Bedside. *Curr Med Chem*. 26(27):5152-5164. doi: 10.2174/0929867324666171003114154. PMID: 28971760.
- Burton-Freeman, B., & Reimers, K. (2011). Tomato Consumption and Health: Emerging Benefits. *American Journal of Lifestyle Medicine*, 5(2), 182-191. <https://doi.org/10.1177/1559827610387488>
- CADPR. (2018). *Cartografía de invernaderos en Almería , Granada y Malaga. Consejería de Agricultura, Pesca y Desarrollo Rural (CADPR): Andalucía, Spain, 2018*.
- Cajamar. (2016). *Greenhouse agriculture in Almería A comprehensive techno-economic analysis*.
- Camelo, A. F. L., & Gómez, P. A. (2004). *Comparison of color indexes for tomato ripening*. 534-537.
- Cameron, G. T., & Geana, M. V. (2005). Symposium : Relative Bioactivity of Functional Foods and Related Dietary Supplements Functional Foods: Delivering Information to the Oncology Nurse 1 , 2. *The Journal of Nutrition*, 1(February), 1253-1255.
- Castellano, R., Rosales, M. A., Ruiz, J. M., & Romero, L. (2007). *Sucrolytic activities in cherry tomato fruits in relation to temperature and solar radiation*. 113, 244-249. <https://doi.org/10.1016/j.scienta.2007.03.015>
- Cástilla, N. (2012). *Greenhouse Technology and Management* (2nd ed.).
- Castilla, N., & Hernandez, J. (2007). Greenhouse technological packages for high-quality crop production. *Acta Horticulturae*, 761, 285-297. <https://doi.org/10.17660/ActaHortic.2007.761.38>
- Castilla, N., Hernández, J., & Abou-Hadid, A. F. (2004). Strategic crop and greenhouse management in mild winter climate areas. *Acta Horticulturae*, 633, 183-196.

<https://doi.org/10.17660/ActaHortic.2004.633.22>

- Castilla, N., & Montero, J. I. (2008). Environmental control and crop production in Mediterranean greenhouses. *Acta Horticulturae*, 797, 25-36. <https://doi.org/10.17660/ActaHortic.2008.797.1>
- Castro, A. J., López-Rodríguez, M. D., Giagnocavo, C., Gimenez, M., Céspedes, L., La Calle, A., Gallardo, M., Pumares, P., Cabello, J., Rodríguez, E., Uclés, D., Parra, S., Casas, J., Rodríguez, F., Fernandez-Prados, J. S., Alba-Patiño, D., Expósito-Granados, M., Murillo-López, B. E., Vasquez, L. M., & Valera, D. L. (2019). Six collective challenges for sustainability of Almería greenhouse horticulture. *International Journal of Environmental Research and Public Health*, 16(21). <https://doi.org/10.3390/ijerph16214097>
- Cazor, A., Deborde, C., Moing, A., Rolin, D., & This, H. (2006). Sucrose, glucose, and fructose extraction in aqueous carrot root extracts prepared at different temperatures by means of direct NMR measurements. *Journal of Agricultural and Food Chemistry*, 54(13), 4681-4686. <https://doi.org/10.1021/jf060144i>
- Chambers, D. C., Carew, A. M., Lukowski, S. W., & Powell, J. E. (2019). Transcriptomics and single-cell RNA-sequencing. *Respirology*, 24(1), 29-36. <https://doi.org/10.1111/resp.13412>
- Chen, W., Wang, S., Wu, Y., Shen, X., Xu, S., Guo, Z., Zhang, R., Xing, D. (2020). The Physiologic Activity and Mechanism of Quercetin-Like Natural Plant Flavonoids. *Curr Pharm Biotechnol.*;21(8):654-658. doi: 10.2174/1389201021666200212093130. PMID: 32048963.
- Choi, S., Kim, H., Kim, H., Lee, I., Kozukue, N., Levin, C. E., & Friedman, M. (2011). *Free Amino Acid and Phenolic Contents and Antioxidative and Cancer Cell-Inhibiting Activities of Extracts of 11 Greenhouse-Grown Tomato Varieties and 13 Tomato-Based Foods*. 12801-12814.
- Choksi, P. M., & Joshi, V. Y. (2007). A review on lycopene - Extraction, purification, stability and applications. *International Journal of Food Properties*, 10(2), 289-298. <https://doi.org/10.1080/10942910601052699>
- Chong, J., Wishart, D. S., & Xia, J. (2019). *Comprehensive and Integrative Metabolomics Data Analysis*. 68, 1-128. <https://doi.org/10.1002/cpbi.86>
- Chylińska, M., Szymańska-Chargot, M., Deryło, K., Tchórzewska, D., & Zdunek, A. (2017). Changing of biochemical parameters and cell wall polysaccharides distribution during physiological development of tomato fruit. *Plant Physiology and Biochemistry*, 119, 328-337. <https://doi.org/10.1016/j.plaphy.2017.09.010>
- Cocco, A., Deliperi, S., & Delrio, G. (2013). Control of *Tuta absoluta* (Meyrick) (Lepidoptera:

- Gelechiidae) in greenhouse tomato crops using the mating disruption technique. *Journal of Applied Entomology*, 137(1-2), 16-28. <https://doi.org/10.1111/j.1439-0418.2012.01735.x>
- Das, U. N. (2006). *Essential fatty acids : biochemistry , physiology and pathology*. 420-439. <https://doi.org/10.1002/biot.200600012>
- Deborde, C., Moing, A., Roch, L., Jacob, D., Rolin, D., & Giraudeau, P. (2017). Plant metabolism as studied by NMR spectroscopy. *Progress in Nuclear Magnetic Resonance Spectroscopy*.
- Deborde, C. J. F. D. J. A. V. N., Richard-Forget, F., Sylvain, L., Cédric, D., Kamar, H., François, M., Moing, A., & Molinié, R. (2019). *Optimizing 1D 1 H-NMR profiling of plant samples for high throughput analysis: extract preparation, standardization, automation and spectra processing*.
- Deborde, C., Maucourt, M., Baldet, P., Bernillon, S., Biais, B., Talon, G., Ferrand, C., Jacob, D., Ferry-Dumazet, H., de Daruvar, A., Rolin, D., & Moing, A. (2009). Proton NMR quantitative profiling for quality assessment of greenhouse-grown tomato fruit. *Metabolomics*, 5(2), 183-198. <https://doi.org/10.1007/s11306-008-0134-2>
- Desneux, N., Luna, M. G., Guillemaud, T., & Urbaneja, A. (2011). The invasive South American tomato pinworm, *Tuta absoluta*, continues to spread in Afro-Eurasia and beyond: The new threat to tomato world production. *Journal of Pest Science*, 84(4), 403-408. <https://doi.org/10.1007/s10340-011-0398-6>
- Desneux, N., Wajnberg, E., Wyckhuys, K. A. G., Burgio, G., Arpaia, S., Narváez-Vasquez, C. A., González-Cabrera, J., Ruescas, D. C., Tabone, E., Frandon, J., Pizzol, J., Poncet, C., Cabello, T., & Urbaneja, A. (2010). Biological invasion of European tomato crops by *Tuta absoluta*: Ecology, geographic expansion and prospects for biological control. *Journal of Pest Science*, 83(3), 197-215. <https://doi.org/10.1007/s10340-010-0321-6>
- Dixon, R. A., Achnine, L., Kota, P., Liu, C.-J., Reddy, M. S. S., & Wang, L. (2002). *The phenylpropanoid pathway and plant defence – a genomics perspective*. 3, 371-390.
- Dixon, R. A., Gang, D. R., Charlton, A. J., Fiehn, O., Kuiper, H. A., Reynolds, T. L., Tjeerdema, R. S., Jeffery, E. H., German, J. B., Ridley, W. P., & Seiber, J. N. (2006). Applications of metabolomics in agriculture. *Journal of Agricultural and Food Chemistry*, 54(24), 8984-8994. <https://doi.org/10.1021/jf061218t>
- Doering, O., & Sorensen, A. (2018). The land that shapes and sustains us. *How to Feed the World*, 46-58. <https://doi.org/10.5822/978-1-61091-885-5>
- Dona, A. C., Kyriakides, M., Scott, F., Shephard, E. A., Varshavi, D., Veselkov, K., & Everett, J. R. (2016). A guide to the identification of metabolites in NMR-based

- metabonomics/metabolomics experiments. *Computational and Structural Biotechnology Journal*, 14, 135-153. <https://doi.org/10.1016/j.csbj.2016.02.005>
- Dorais, M., Ehret, D. L., & Papadopoulos, A. P. (2008). Tomato (*Solanum lycopersicum*) health components: From the seed to the consumer. *Phytochemistry Reviews*, 7(2), 231-250. <https://doi.org/10.1007/s11101-007-9085-x>
- Dra, I. S. R. A., & Ruiz, L. B. (2015). *La era de las ciencias ómicas*.
- Egea, F. J., Torrente, R. G., & Aguilar, A. (2018). An efficient agro-industrial complex in Almería (Spain): Towards an integrated and sustainable bioeconomy model. *New Biotechnology*, 40(16), 103-112. <https://doi.org/10.1016/j.nbt.2017.06.009>
- Eisenmann, P., Ehlers, M., Weinert, C. H., Tzvetkova, P., Silber, M., Rist, M. J., Luy, B., & Muhle-Goll, C. (2016). Untargeted NMR spectroscopic analysis of the metabolic variety of new apple cultivars. *Metabolites*, 6(3). <https://doi.org/10.3390/metabo6030029>
- Elbadrawy, E. (2011). Evaluation of nutritional value and antioxidant activity of tomato peel extracts. *ARABIAN JOURNAL OF CHEMISTRY*. <https://doi.org/10.1016/j.arabjc.2011.11.011>
- Emwas, A. H. M. (2015). The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research. In *Methods in Molecular Biology* (Vol. 1277). [https://doi.org/10.1007/978-1-4939-2377-9\\_13](https://doi.org/10.1007/978-1-4939-2377-9_13)
- Falco, B. De, Ercolano, M., Manzo, D., Incerti, G., Garonna, A. Pietro, & Lanzotti, V. (2019). *Metabolomics approach based on NMR spectroscopy and multivariate data analysis to explore the interaction between the leafminer Tuta absoluta and tomato ( Solanum lycopersicum )*. May, 1-8. <https://doi.org/10.1002/pca.2850>
- Fang, C., Fernie, A. R., & Luo, J. (2019). Exploring the Diversity of Plant Metabolism. *Trends in Plant Science*, 24(1), 83-98. <https://doi.org/10.1016/j.tplants.2018.09.006>
- Fresh Plaza. (2019) “Almería’s tomato acreage expected to fall significantly again.”. Retrieved February 19, 2021, from <https://www.freshplaza.com/article/9143348/almeria-s-tomato-acreage-expected-to-fall-significantly-again/>
- Fresh Plaza. (2020). *Results of the 2019/20 fruit and vegetable campaign in Almería*. Retrieved February 19, 2021, from <https://www.freshplaza.com/article/9238251/results-of-the-2019-20-fruit-and-vegetable-campaign-in-almeria/>
- Frusciante, L., Carli, P., Ercolano, M. R., Pernice, R., Di Matteo, A., Fogliano, V., & Pellegrini, N. (2007). Antioxidant nutritional quality of tomato. *Molecular Nutrition and Food Research*, 51(5), 609-617. <https://doi.org/10.1002/mnfr.200600158>

- Gautier, H., Diakou-Verdin, V., Nard, C., Reich, M., Buret, M., Bourgaud, F., Ssel, J. L., Caris-Veyrat, C., & Nard, M. (2008). How Does Tomato Quality (Sugar, Acid, and Nutritional Quality) Vary with Ripening Stage, Temperature, and Irradiance? *Journal of Agricultural and Food Chemistry*, 56(4), 1241-1250. [http://pubs3.acs.org/acs/journals/doilookup?in\\_doi=10.1021/jf072196t](http://pubs3.acs.org/acs/journals/doilookup?in_doi=10.1021/jf072196t)
- Ghatak, A., Chaturvedi, P., & Weckwerth, W. (2018). Metabolomics in plant stress physiology. *Advances in Biochemical Engineering/Biotechnology*, 164, 187-236. [https://doi.org/10.1007/10\\_2017\\_55](https://doi.org/10.1007/10_2017_55)
- Gibon, Y., & Rolin, D. (2012). Aspects of experimental design for plant metabolomics experiments and guidelines for growth of plant material. *Methods in Molecular Biology*, 860, 13-30. [https://doi.org/10.1007/978-1-61779-594-7\\_2](https://doi.org/10.1007/978-1-61779-594-7_2)
- Giraudeau, P. (2020). NMR-based metabolomics and fluxomics: Developments and future prospects. *Analyst*, 145(7), 2457-2472. <https://doi.org/10.1039/d0an00142b>
- Giraudeau, P., Silvestre, V., & Akoka, S. (2015). Optimizing water suppression for quantitative NMR-based metabolomics: a tutorial review. *Metabolomics*, 11(5), 1041-1055. <https://doi.org/10.1007/s11306-015-0794-7>
- Girelli, C. R., Accogli, R., Del Coco, L., Angilè, F., De Bellis, L., & Fanizzi, F. P. (2018). <sup>1</sup>H-NMR-based metabolomic profiles of different sweet melon (*Cucumis melo* L.) Salento varieties: Analysis and comparison. *Food Research International*, 114(March), 81-89. <https://doi.org/10.1016/j.foodres.2018.07.045>
- Girelli, C. R., Angilè, F., Coco, L. Del, Migoni, D., Zampella, L., Marcelletti, S., Cristella, N., Marangi, P., Scortichini, M., & Fanizzi, F. P. (2019). <sup>1</sup>H-NMR metabolite fingerprinting analysis reveals a disease biomarker and a field treatment response in *xylella fastidiosa* subsp. *Pauca*-infected olive trees. *Plants*, 8(5). <https://doi.org/10.3390/plants8050115>
- Giuffrè, A.M. and Capocasale, M. (2016). Sterol composition of tomato (*Solanum lycopersicum* L.) seed oil : the effect of cultivar. *International Food Research Journal*, 23(1), 116-122.
- Gromski, P. S., Muhamadali, H., Ellis, D. I., Xu, Y., Correa, E., Turner, M. L., & Goodacre, R. (2015). A tutorial review: Metabolomics and partial least squares-discriminant analysis - a marriage of convenience or a shotgun wedding. *Analytica Chimica Acta*, 879, 10-23. <https://doi.org/10.1016/j.aca.2015.02.012>
- Gunther, H. (1973). *NMR spectroscopy: Basic principles, Concepts and Applications in Chemistry* (Third).
- Guttmacher, A. E., & Collins, F. S. (2003). Welcome to the Genomic Era. *New England Journal of Medicine*, 349(10), 996-998. <https://doi.org/10.1056/nejme038132>

- Hall, R. D. (2006). Plant metabolomics: From holistic hope, to hype, to hot topic. *New Phytologist*, 169(3), 453-468. <https://doi.org/10.1111/j.1469-8137.2005.01632.x>
- Héberger, K. (2008). Chemoinformatics-multivariate mathematical-statistical methods for data evaluation. *Medical Applications of Mass Spectrometry*, 141-169. <https://doi.org/10.1016/B978-044451980-1.50009-4>
- Herrmann, K. (1979). Übersicht über die Inhaltsstoffe der Tomaten. *Zeitschrift Für Lebensmittel-Untersuchung Und -Forschung*, 169(3), 179-200. <https://doi.org/10.1007/BF01333321>
- Hertel, T. W. (2011). The global supply and demand for agricultural land in 2050: A perfect storm in the making? *American Journal of Agricultural Economics*, 93(2), 259-275. <https://doi.org/10.1093/ajae/aaq189>
- Hohmann, M., Christoph, N., Wachter, H., & Holzgrabe, U. (2014). 1H NMR profiling as an approach to differentiate conventionally and organically grown tomatoes. *Journal of Agricultural and Food Chemistry*, 62(33), 8530-8540. <https://doi.org/10.1021/jf502113r>
- Hortidaily. (2018). *Biopesticides against tomato leafminer, Tuta absoluta*. (2019). <https://www.hortidaily.com/article/9130675/biopesticides-against-tomato-leafminer-tuta-absoluta/>
- Hucl, P., Sosulski, F. W., Graf, R., Gillott, C., & Pietrzak, L. (2001). *Screening Spring Wheat for Midge Resistance in Relation to Ferulic*. 0-7.
- Ili, Z. S., Milenkovi, L., Šuni, L., & Fallik, E. (2014). *Effect of Colored Shade Nets on Plant Leaf Parameters and Tomato Fruit Quality*. <https://doi.org/10.1002/jsfa.7000>
- Kearney, J. (2010). Food consumption trends and drivers. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1554), 2793-2807. <https://doi.org/10.1098/rstb.2010.0149>
- Kim, H. K., Choi, Y. H., & Verpoorte, R. (2011). NMR-based plant metabolomics: Where do we stand, where do we go? *Trends in Biotechnology*, 29(6), 267-275. <https://doi.org/10.1016/j.tibtech.2011.02.001>
- Kiokias, S., Proestos, C., & Oreopoulou, V. (2020). Phenolic acids of plant origin-a review on their antioxidant activity in vitro (O/W emulsion systems) along with their in vivo health biochemical properties. *Foods*, 9(4). <https://doi.org/10.3390/foods9040534>
- Klee, H. J., & Giovannoni, J. J. (2011). Genetics and control of tomato fruit ripening and quality attributes. *Annual Review of Genetics*, 45, 41-59. <https://doi.org/10.1146/annurev-genet-110410-132507>

- Klee, H. J., & Tieman, D. M. (2013). Genetic challenges of flavor improvement in tomato. *Trends in Genetics*, 29(4), 257-262. <https://doi.org/10.1016/j.tig.2012.12.003>
- Kundu, A., & Vadassery, J. (2019). *Chlorogenic acid-mediated chemical defence of plants against insect herbivores*. 21, 185-189. <https://doi.org/10.1111/plb.12947>
- Lainez, M., González, J. M., Aguilar, A., & Vela, C. (2018). Spanish strategy on bioeconomy: Towards a knowledge based sustainable innovation. *New Biotechnology*, 40, 87-95. <https://doi.org/10.1016/j.nbt.2017.05.006>
- Le Gall, G., Colquhoun, I. J., Davis, A. L., Collins, G. J., & Verhoeven, M. E. (2003). Metabolite profiling of tomato (*Lycopersicon esculentum*) using <sup>1</sup>H NMR spectroscopy as a tool to detect potential unintended effects following a genetic modification. *Journal of Agricultural and Food Chemistry*, 51(9), 2447-2456. <https://doi.org/10.1021/jf0259967>
- Leiss, K. A., Choi, Y. H., Verpoorte, R., & Klinkhamer, P. G. L. (2011). *An overview of NMR-based metabolomics to identify secondary plant compounds involved in host plant resistance*. 205-216. <https://doi.org/10.1007/s11101-010-9175-z>
- Leiss, K. A., Maltese, F., Choi, Y. H., Verpoorte, R., & Klinkhamer, P. G. L. (2020). *Identification of Chlorogenic Acid as a Resistance Factor for Thrips in Chrysanthemum [ C ][ OA ]*. 150(July 2009), 1567-1575. <https://doi.org/10.1104/pp.109.138131>
- Lesk, A. (2017). CHAPTER 1. Introduction to genomics. *Introduction To Genomics*, 823, 3-20. <https://doi.org/10.1007/978-1-60327-216-2>
- López-gresa, M. P., Lisón, P., Kyong, H., Hae, Y., Verpoorte, R., Rodrigo, I., Conejero, V., & María, J. (2012). Metabolic fingerprinting of Tomato Mosaic Virus infected *Solanum lycopersicum*. *Journal of Plant Physiology*, 169(16), 1586-1596. <https://doi.org/10.1016/j.jplph.2012.05.021>
- López-gresa, M. P., Maltese, F., Bellés, M., Conejero, V., Kim, K., & Choi, H. (2010). *Metabolic Response of Tomato Leaves Upon Different Plant - Pathogen Interactions † Robert Verpoorte b Introduction - Plants utilise various defence mechanisms against their potential biotic stressing agents such as viroids , viruses ,. October 2009*, 89-94. <https://doi.org/10.1002/pca.1179>
- Lugasi, A. (2006). *Tomato Fruit Quality and Content Depend on Stage of Maturity*. 41(6), 1400-1401.
- Marchand, J., Martineau, E., Guitton, Y., Dervilly-Pinel, G., & Giraudeau, P. (2017). Multidimensional NMR approaches towards highly resolved, sensitive and high-throughput quantitative metabolomics. *Current Opinion in Biotechnology*, 43, 49-55. <https://doi.org/10.1016/j.copbio.2016.08.004>

- María, E., Pérez, S., José, M., López, F., & Sánchez, I. (2010). Study of the suitability of HRMAS NMR for metabolic profiling of tomatoes : Application to tissue differentiation and fruit ripening. *Food Chemistry*, 122(3), 877-887. <https://doi.org/10.1016/j.foodchem.2010.03.003>
- Markley, J. L., Brüschweiler, R., Edison, A. S., Eghbalnia, H. R., Powers, R., Raftery, D., & Wishart, D. S. (2017). The future of NMR-based metabolomics. *Current Opinion in Biotechnology*, 43, 34-40. <https://doi.org/10.1016/j.copbio.2016.08.001>
- Marshall, D. D., & Powers, R. (2017). Beyond the paradigm: Combining mass spectrometry and nuclear magnetic resonance for metabolomics. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 100, 1-16. <https://doi.org/10.1016/j.pnmrs.2017.01.001>
- Martí, R., Roselló, S., & Cebolla-Cornejo, J. (2016). Tomato as a source of carotenoids and polyphenols targeted to cancer prevention. *Cancers*, 8(6), 1-28. <https://doi.org/10.3390/cancers8060058>
- Martins, M. C. M., Caldana, C., Wolf, L. D., & de Abreu, L. G. F. (2018). The importance of experimental design, quality assurance, and control in plant metabolomics experiments. *Methods in Molecular Biology*, 1778, 3-17. [https://doi.org/10.1007/978-1-4939-7819-9\\_1](https://doi.org/10.1007/978-1-4939-7819-9_1)
- Masetti, O., Ciampa, A., Nisini, L., Sequi, P., & Dell'Abate, M. T. (2017). A multifactorial approach in characterizing geographical origin of Sicilian cherry tomatoes using <sup>1</sup>H-NMR profiling. *Food Research International*, 100, 623-630. <https://doi.org/10.1016/j.foodres.2017.07.059>
- Mazzei, P., Vinale, F., Woo, S. L., Pascale, A., Lorito, M., & Piccolo, A. (2016). Metabolomics by Proton High-Resolution Magic-Angle-Spinning Nuclear Magnetic Resonance of Tomato Plants Treated with Two Secondary Metabolites Isolated from Trichoderma. *Journal of Agricultural and Food Chemistry*, 64(18), 3538-3545. <https://doi.org/10.1021/acs.jafc.6b00801>
- Mckay, R. T. (2011). How the 1D-NOESY suppresses solvent signal in metabonomics NMR spectroscopy: An examination of the pulse sequence components and evolution. *Concepts in Magnetic Resonance Part A: Bridging Education and Research*, 38 A(5), 197-220. <https://doi.org/10.1002/cmr.a.20223>
- Milenkovi, L. (2012). Effects of the modification of light intensity by color shade nets on yield and quality of tomato fruits. *Scientia Horticulturae*, 139, 90-95. <https://doi.org/10.1016/j.scienta.2012.03.009>
- Milewska, E. M., Czyzyk, A., Meczekalski, B., & Genazzani, A. D. (2016). Inositol and human reproduction. From cellular metabolism to clinical use. *Gynecological Endocrinology*,

- 32(9), 690-695. <https://doi.org/10.1080/09513590.2016.1188282>
- Moco, S., Capanoglu, E., Tikunov, Y., Bino, R. J., Boyacioglu, D., Hall, R. D., Vervoort, J., & De Vos, R. C. H. (2007). Tissue specialization at the metabolite level is perceived during the development of tomato fruit. *Journal of Experimental Botany*, 58(15-16), 4131-4146. <https://doi.org/10.1093/jxb/erm271>
- Mukhopadhyay, R. (2007). Liquid NMR probes: Oh so many choices. *Analytical Chemistry*, 79(21), 7959-7963. <https://doi.org/10.1021/ac071984e>
- Munir, S., Mumtaz, M. A., Ahiakpa, J. K., Liu, G., Chen, W., Zhou, G., Zheng, W., Ye, Z., & Zhang, Y. (2020). Genome-wide analysis of Myo-inositol oxygenase gene family in tomato reveals their involvement in ascorbic acid accumulation. *BMC Genomics*, 21(1), 1-15. <https://doi.org/10.1186/s12864-020-6708-8>
- Natalie. Serkova , Denise. Davis, S. J., & Agarwal, R. (2019). *Quantitative NMR-Based Metabolomics on Tissue Biomarkers and Its Translation into In Vivo Magnetic Resonance Spectroscopy*. <https://doi.org/10.1515/9781618110459-026>
- Nicola, S., Tibaldi, G., & Fontana, E. (2009). Tomato production systems and their application to the tropics. *Acta Horticulturae*, 821, 27-33. <https://doi.org/10.17660/actahortic.2009.821.1>
- Nonaka, S., Arai, C., Takayama, M., Matsukura, C., & Ezura, H. (2017). Efficient increase of  $\Gamma$ -aminobutyric acid (GABA) content in tomato fruits by targeted mutagenesis. *Scientific Reports*, 7(1), 1-14. <https://doi.org/10.1038/s41598-017-06400-y>
- Ntagkas, N., Woltering, E., Bouras, S., De Vos, R. C., Dieleman, J. A., Nicole, C. C., Labrie, C., & Marcelis, L. F. (2019). Light-induced vitamin c accumulation in tomato fruits is independent of carbohydrate availability. *Plants*, 8(4), 1-12. <https://doi.org/10.3390/plants8040086>
- ONU. (2019). World population prospects 2019. In *Department of Economic and Social Affairs. World Population Prospects 2019*. (Issue 141). <http://www.ncbi.nlm.nih.gov/pubmed/12283219>
- Pandey, A., & Mann, M. (2000). Proteomics to study genes and genomes. *Nature*, 405(6788), 837-846. <https://doi.org/10.1038/35015709>
- Pardossi, A., Tognoni, F., & Incrocci, L. (2004). *Mediterranean Greenhouse Technology*.
- Patti, G. J., Yanes, O., & Siuzdak, G. (2012). Innovation: Metabolomics: the apogee of the omics trilogy. *Nature Reviews Molecular Cell Biology*, 13(4), 263-269. <https://doi.org/10.1038/nrm3314>

- Philouze, L. J. (2001). *Genetic analysis of organoleptic quality in fresh market tomato . 1 . Mapping QTLs for physical and chemical traits*. 259-272.
- Piironen, V., Lindsay, D. G., Miettinen, T. A., Toivo, J., & Lampi, A. (2000). *Review Plant sterols: biosynthesis , biological function and their importance to human nutrition*. 966(February).
- Quinet, M., Angosto, T., Yuste-Lisbona, F. J., Blanchard-Gros, R., Bigot, S., Martinez, J. P., & Lutts, S. (2019). Tomato Fruit Development and Metabolism. *Frontiers in Plant Science*, 10. <https://doi.org/10.3389/fpls.2019.01554>
- Rajapakse, N. C., & Shahak, Y. (2007). *Light-quality manipulation by horticulture industry*. 30, 290-312. <https://doi.org/10.1002/9781119312994.apr0321>
- Ralli, E., Amargianitaki, M., Manolopoulou, E., Misiak, M., Markakis, G., Tachtalidou, S., Kolesnikova, A., Dais, P., & Spyros, A. (2018). NMR spectroscopy protocols for food metabolomics applications. *Methods in Molecular Biology*, 1738, 203-211. [https://doi.org/10.1007/978-1-4939-7643-0\\_14](https://doi.org/10.1007/978-1-4939-7643-0_14)
- Reca, J., Trillo, C., Sánchez, J. A., Martínez, J., & Valera, D. (2018). Optimization model for on-farm irrigation management of Mediterranean greenhouse crops using desalinated and saline water from different sources. *Agricultural Systems*, 166(November 2017), 173-183. <https://doi.org/10.1016/j.agsy.2018.02.004>
- Rodrigues, A. M., Ribeiro-Barros, A. I., & António, C. (2019). Experimental design and sample preparation in forest tree metabolomics. *Metabolites*, 9(12), 5-7. <https://doi.org/10.3390/metabo9120285>
- Salem, M. A., De Souza, L. P., Serag, A., Fernie, A. R., Farag, M. A., Ezzat, S. M., & Alseekh, S. (2020). Metabolomics in the context of plant natural products research: From sample preparation to metabolite analysis. *Metabolites*, 10(1), 1-30. <https://doi.org/10.3390/metabo10010037>
- Sauerschnig, C., Doppler, M., Bueschl, C., & Schuhmacher, R. (2018). Methanol generates numerous artifacts during sample extraction and storage of extracts in metabolomics research. *Metabolites*, 8(1), 1-19. <https://doi.org/10.3390/metabo8010001>
- Schrimpe-Rutledge, A. C., Codreanu, S. G., Sherrod, S. D., & McLean, J. A. (2016). Untargeted Metabolomics Strategies—Challenges and Emerging Directions. *Journal of the American Society for Mass Spectrometry*, 27(12), 1897-1905. <https://doi.org/10.1007/s13361-016-1469-y>
- Segers, K., Declerck, S., Mangelings, D., Heyden, Y. Vander, & Eeckhaut, A. Van. (2019). Analytical techniques for metabolomic studies: A review. *Bioanalysis*, 11(24), 2297-2318.

<https://doi.org/10.4155/bio-2019-0014>

- Shimeles, T., Do, S. P., Mu, H. S., & Cheon, S. J. (2017). Review on factors affecting the quality and antioxidant properties of tomatoes. *African Journal of Biotechnology*, 16(32), 1678-1687. <https://doi.org/10.5897/ajb2017.16054>
- Shinozaki, Y., Nicolas, P., Fernandez-Pozo, N., Ma, Q., Evanich, D. J., Shi, Y., Xu, Y., Zheng, Y., Snyder, S. I., Martin, L. B. B., Ruiz-May, E., Thannhauser, T. W., Chen, K., Domozych, D. S., Catalá, C., Fei, Z., Mueller, L. A., Giovannoni, J. J., & Rose, J. K. C. (2018). High-resolution spatiotemporal transcriptome mapping of tomato fruit development and ripening. *Nature Communications*, 9(1). <https://doi.org/10.1038/s41467-017-02782-9>
- Simmonds, M. S. J. (2001). *Importance of flavonoids in insect ± plant interactions : feeding and oviposition*. 56, 245-252.
- Singh, A. K., Dhanapal, S., & Yadav, B. S. (2020). The dynamic responses of plant physiology and metabolism during environmental stress progression. *Molecular Biology Reports*, 47(2), 1459-1470. <https://doi.org/10.1007/s11033-019-05198-4>
- Slimestad, R., Fossen, T., & Verheul, M. J. (2008). The flavonoids of tomatoes. *Journal of Agricultural and Food Chemistry*, 56(7), 2436-2441. <https://doi.org/10.1021/jf073434n>
- Sobolev, A. P., Mannina, L., Proietti, N., Carradori, S., Daglia, M., Giusti, A. M., Antiochia, R., & Capitani, D. (2015). Untargeted NMR-based methodology in the study of fruit metabolites. *Molecules*, 20(3), 4088-4108. <https://doi.org/10.3390/molecules20034088>
- Suzuki, T., Honda, Y., & Mukasa, Y. (2005). *Effects of UV-B radiation , cold and desiccation stress on rutin concentration and rutin glucosidase activity in tartary buckwheat ( Fagopyrum tataricum ) leaves*. 168, 1303-1307. <https://doi.org/10.1016/j.plantsci.2005.01.007>
- Takeoka, G. R., Dao, L., Flessa, S., Gillespie, D. M., Jewell, W. T., Huebner, B., Bertow, D., & Ebeler, S. E. (2001). Processing effects on lycopene content and antioxidant activity of tomatoes. *Journal of Agricultural and Food Chemistry*, 49(8), 3713-3717. <https://doi.org/10.1021/jf0102721>
- Tieman, D., Bliss, P., McIntyre, L. M., Blandon-Ubeda, A., Bies, D., Odabasi, A. Z., Rodríguez, G. R., Van Der Knaap, E., Taylor, M. G., Goulet, C., Mageroy, M. H., Snyder, D. J., Colquhoun, T., Moskowitz, H., Clark, D. G., Sims, C., Bartoshuk, L., & Klee, H. J. (2012). The chemical interactions underlying tomato flavor preferences. *Current Biology*, 22(11), 1035-1039. <https://doi.org/10.1016/j.cub.2012.04.016>
- Tiziani, S., Schwartz, S. J., & Vodovotz, Y. (2006). Profiling of carotenoids in tomato juice by one- and two-dimensional NMR. *Journal of Agricultural and Food Chemistry*, 54(16), 6094-

6100. <https://doi.org/10.1021/jf061154m>

- Triba, M. N., Le Moyec, L., Amathieu, R., Goossens, C., Bouchemal, N., Nahon, P., Rutledge, D. N., & Savarin, P. (2015). PLS/OPLS models in metabolomics: The impact of permutation of dataset rows on the K-fold cross-validation quality parameters. *Molecular BioSystems*, 11(1), 13-19. <https://doi.org/10.1039/c4mb00414k>
- Turza, M. P., & Petro-turza, M. (2009). *Flavor of tomato and tomato products FLAVOR OF TOMATO AND TOMATO PRODUCTS*. September 2013, 37-41. <https://doi.org/10.1080/87559128609540802>
- Vajargah, K. F., Sadeghi-Bazargani, H., Mehdizadeh-Esfanjani, R., Savadi-Oskouei, D., & Farhoudi, M. (2012). OPLS statistical model versus linear regression to assess sonographic predictors of stroke prognosis. *Neuropsychiatric Disease and Treatment*, 8, 387-392.
- Valera, D. L., Belmonte, L. J., Molina-Aiz, F. D., López, A., & Camacho, F. (2017). The greenhouses of Almería, Spain: Technological analysis and profitability. *Acta Horticulturae*, 1170, 219-226. <https://doi.org/10.17660/ActaHortic.2017.1170.25>
- Verpoorte, R., Kim, H. K., & Choi, Y. H. (2007). *NMR-based metabolomics at work in phytochemistry*. September 2006, 3-14. <https://doi.org/10.1007/s11101-006-9031-3>
- Wang, S., Alseekh, S., Fernie, A. R., & Luo, J. (2019). The Structure and Function of Major Plant Metabolite Modifications. *Molecular Plant*, 12(7), 899-919. <https://doi.org/10.1016/j.molp.2019.06.001>
- Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics in Western Equatoria State. *Nature Reviews Genetics*, 10(1), 57.
- Watanabe, M., Ohta, Y., Licang, S., Motoyama, N., & Kikuchi, J. (2015). Profiling contents of water-soluble metabolites and mineral nutrients to evaluate the effects of pesticides and organic and chemical fertilizers on tomato fruit quality. *Food Chemistry*, 169, 387-395. <https://doi.org/10.1016/j.foodchem.2014.07.155>
- Widarto, H. T., & Meijden, E. Van Der. (2006). *Metabolomic Differentiation of Brassica rapa Following Herbivory by Different Insect Instars using Two-Dimensional Nuclear Magnetic Resonance Spectroscopy*. 2417-2428. <https://doi.org/10.1007/s10886-006-9152-6>
- Wishart, D. S., Knox, C., Guo, A. C., Eisner, R., Young, N., Gautam, B., Hau, D. D., Psychogios, N., Dong, E., Bouatra, S., Mandal, R., Sinelnikov, I., Xia, J., Jia, L., Cruz, J. A., Lim, E., Sobsey, C. A., Shrivastava, S., Huang, P., ... Forsythe, I. (2009). HMDB: A knowledgebase for the human metabolome. *Nucleic Acids Research*, 37(SUPPL. 1), 603-610. <https://doi.org/10.1093/nar/gkn810>
- Wittwer, S. H., & Castilla, N. (1995). Protected cultivation of horticultural crops worldwide.

- HortTechnology*, 5(1), 6-23. <https://doi.org/10.21273/horttech.5.1.6>
- Xia, J., Sinelnikov, I. V, Han, B., & Wishart, D. S. (2015). *MetaboAnalyst 3 . 0 -- making metabolomics more meaningful*. 1-7. <https://doi.org/10.1093/nar/gkv380>
- Xia, J., & Wishart, D. S. (2016). *Comprehensive Metabolomics Data Analysis*. September, 1-91. <https://doi.org/10.1002/cpbi.11>
- Yin, Y. G., Tominaga, T., Iijima, Y., Aoki, K., Shibata, D., Ashihara, H., Nishimura, S., Ezura, H., & Matsukura, C. (2010). Metabolic alterations in organic acids and  $\gamma$ -aminobutyric acid in developing tomato (*Solanum lycopersicum* L.) fruits. *Plant and Cell Physiology*, 51(8), 1300-1314. <https://doi.org/10.1093/pcp/pcq090>
- Zeisel, S. H. (2000). *Choline : An Essential Nutrient for Humans*. 16, 669-671.
- Zeisel, S. H., Mar, M., Howe, J. C., & Holden, J. M. (2003). *Human Nutrition and Metabolism Concentrations of Choline-Containing Compounds and Betaine in Common Foods 1*. January 2003, 1302-1307.
- Zhao, L., Huang, Y., Hu, J., Zhou, H., Adeleye, A. S., & Keller, A. A. (2016). <sup>1</sup>H NMR and GC-MS Based Metabolomics Reveal Defense and Detoxification Mechanism of Cucumber Plant under Nano-Cu Stress. *Environmental Science and Technology*, 50(4), 2000-2010. <https://doi.org/10.1021/acs.est.5b05011>

## 7 Annexes

### Annex A - Metabolite profiling on tomato fruits

#### *Solvent optimization for metabolites extraction*

In Figure A1, a visual comparison between spectra obtained with different solvents, methanol/phosphate buffer (80:20 v/v) and methanol/phosphate buffer (1:1, v/v), respectively, is provided. In the top image, lipidic compounds are clearly observed due to the higher apolarity of the extracting solvent. These are characterized by wide peaks which prevent the detection of several compounds with resonance on those intervals. Conversely, in the lower image, representing the extraction carried with methanol 1:1 water (v/v), it is presented a cleaner spectrum with more individualized resonances, enabling the identification of more metabolites, a global profiling, as pretended. This way, the second extracting solvent was selected, and every NMR experiment was performed with it.

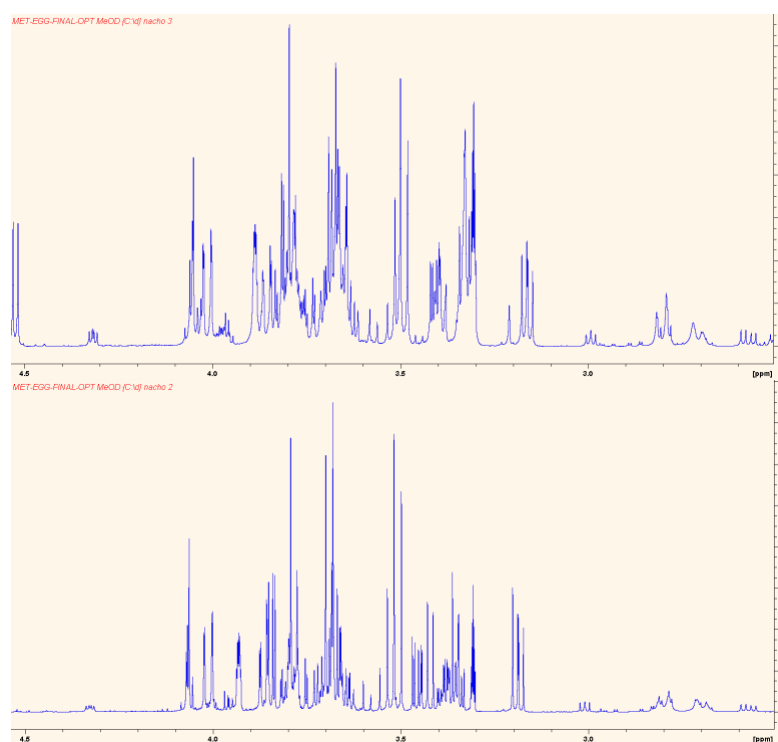


*Figure A1- 1D  $^1\text{H}$ -NMR spectra comparison of tomato fruit samples extracted with different solvents. Top: spectrum obtained with  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (80:20, v/v) fruit extracts. Bottom: spectrum obtained with  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (50:50, v/v) fruit extracts. Spectra were obtained in a Bruker Advance III 600 MHz spectrometer and visualized in Bruker TopSpin 3.6.2.*

## Compound identification by NMR experiment - $\gamma$ -Aminobutyric acid example

$\gamma$ -Aminobutyric acid, also called GABA, is a non-proteogenic amino acid widely distributed in nature that plays a key role as a neurotransmitter, Figure A2. Its presence in the  $^1\text{H}$ -NMR spectra was expected since tomato fruits accumulate higher amounts of this metabolite when compared to the majority of other vegetable crops (Nonaka et al., 2017; Yin et al., 2010).

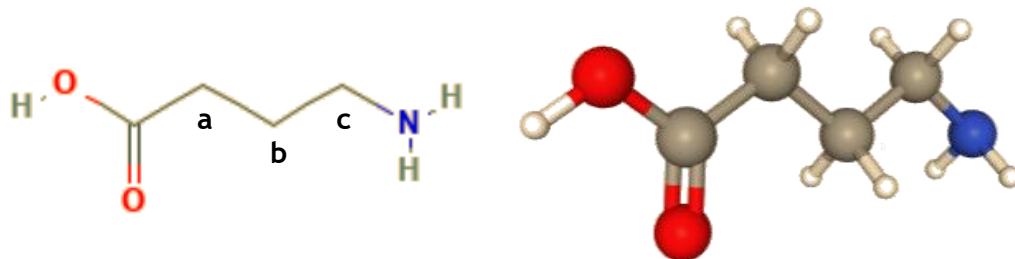


Figure A2 - 2D and 3D chemical structure of  $\gamma$ -Aminobutyric acid. Carbons mentioned in the text are indexed with a, b and c. Images obtained from: National Center for Biotechnology Information (2021). PubChem Compound Summary for CID 119,  $\gamma$ -Aminobutyric acid. Retrieved January 31, 2021 from <https://pubchem.ncbi.nlm.nih.gov/compound/gamma-Aminobutyric-acid>.

The number of signals a compound gives rise to in an NMR spectrum reflects the number of magnetically different protons it has. Relative to the chemical structure of GABA, there are three sets of chemically different protons, each set with a pair of equivalent protons. Carbons a, b and c, each have a pair of equivalent hydrogens which are under a different chemical environment compared to their neighbouring protons.

The chemical environment of a proton provides information regarding the position of its signal in the spectrum in ppm's, the chemical shift. A proton closer to an electronegative atom is under a deshielding effect, as its electron is being moved away from the nucleus, what causes the resonance signal to move downfield (higher ppm's). This way, the  $^1\text{H}$  chemical shifts of GABA in descending order are: a, c and b.

Relative to the number of components, a signal in a  $^1\text{H}$ -NMR spectrum may or may not be split into one or more components. This is called signal multiplicity and gives rise to names such as singlet, doublet, triplet and multiplet. The concept used to figure out the signal multiplicity is the n+1 rule. According to this rule, the signal for the proton under observation will be split into n+1 components depending on protons attached to adjacent carbons, where n is the number of such protons.

Protons from carbon a see two protons in carbon b, so  $2+1=3$ , a triplet is expected.

Protons from carbon b see two protons in carbon a and two protons in carbon c, so  $4+1=5$ , a multiplet with 5 peaks is expected.

Protons from carbon **c** see two protons in carbon **b**, so  $2+1=3$ , another triplet is expected.

With a search into literature data focused in the  $^1\text{H}$ -NMR metabolite profiling of tomato fruits, GABA's three peaks have been frequently assigned in the following regions: 1,89 to 1,93 ppm, 2,32 to 2,42 ppm and 3,00 to 3,03 ppm (C. Deborde et al., 2009; Hohmann et al., 2014; Le Gall et al., 2003). Its location in these regions of the spectrum is also confirmed by searching the compound  $^1\text{H}$ -NMR spectrum in HMDB (Wishart et al., 2009).

Furthermore, relying on CHENOMX software and database it is possible to assign GABA characteristic peaks are located at 3,01 ppm (triplet), 2,38 ppm (triplet) and 1,91 ppm (multiplet), Figure A3. This identification can also be confirmed analyzing the spectra in Bruker TopSpin 3.6.2.

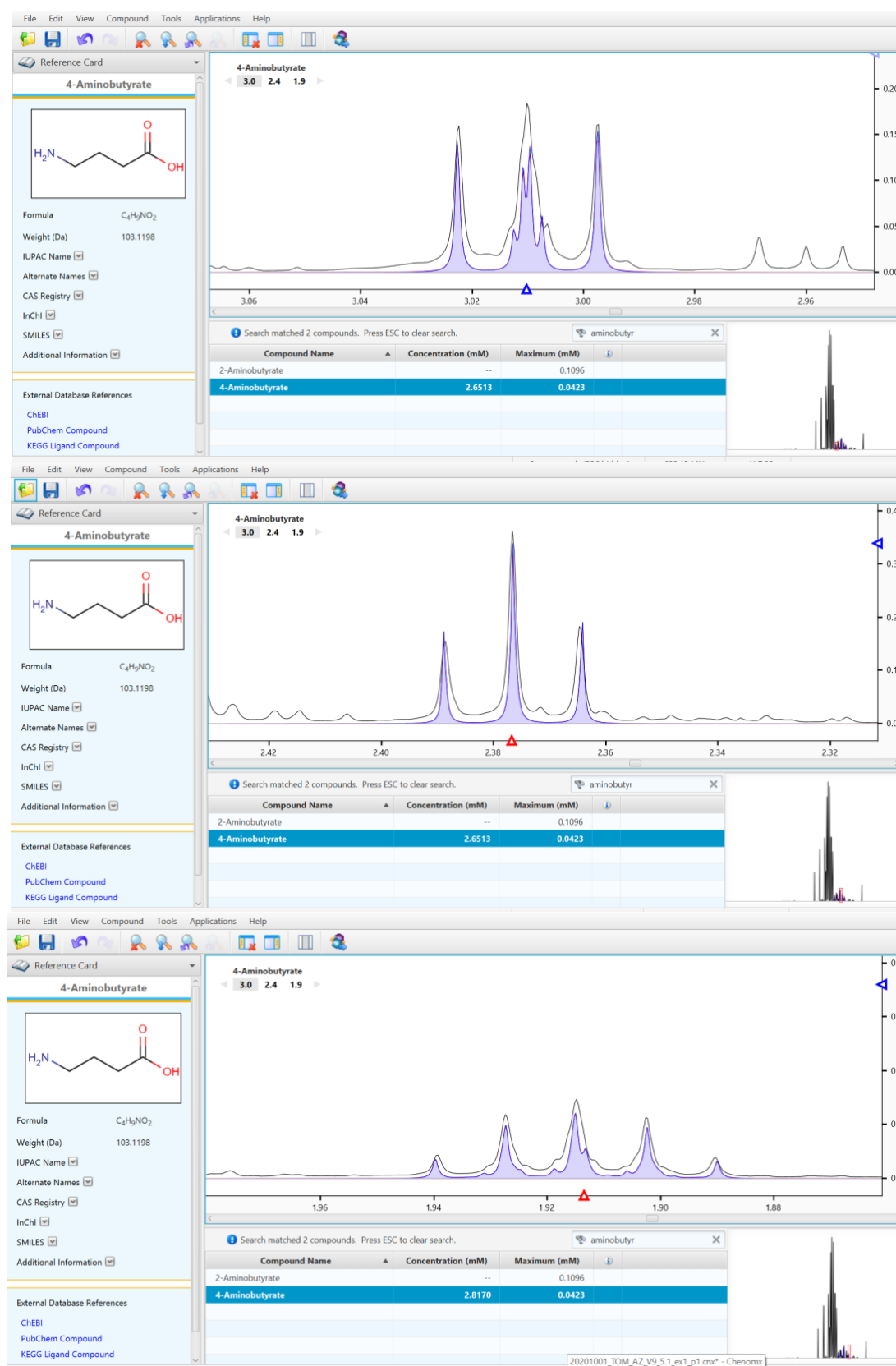


Figure A3 - GABA three characteristic peaks identified in the 1D  $^1\text{H}$ -NMR spectrum of tomato  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (50:50, v/v) fruit extracts. Peak center at 3,01 ppm (t), 2,38 ppm (t) and 1,91 ppm (m) from top to bottom, respectively. Spectrum was analyzed in CHENOMX software.

To confirm compound identification, several 2D NMR experiments were applied.

The  $^1\text{H}$ - $^1\text{H}$  TOCSY experiment allowed the observation of the whole spin system. In this case, the system  $\text{A}_2\text{B}_2\text{C}_2$  is deduced from the 6 cross-peaks circled in red in Figure A4.

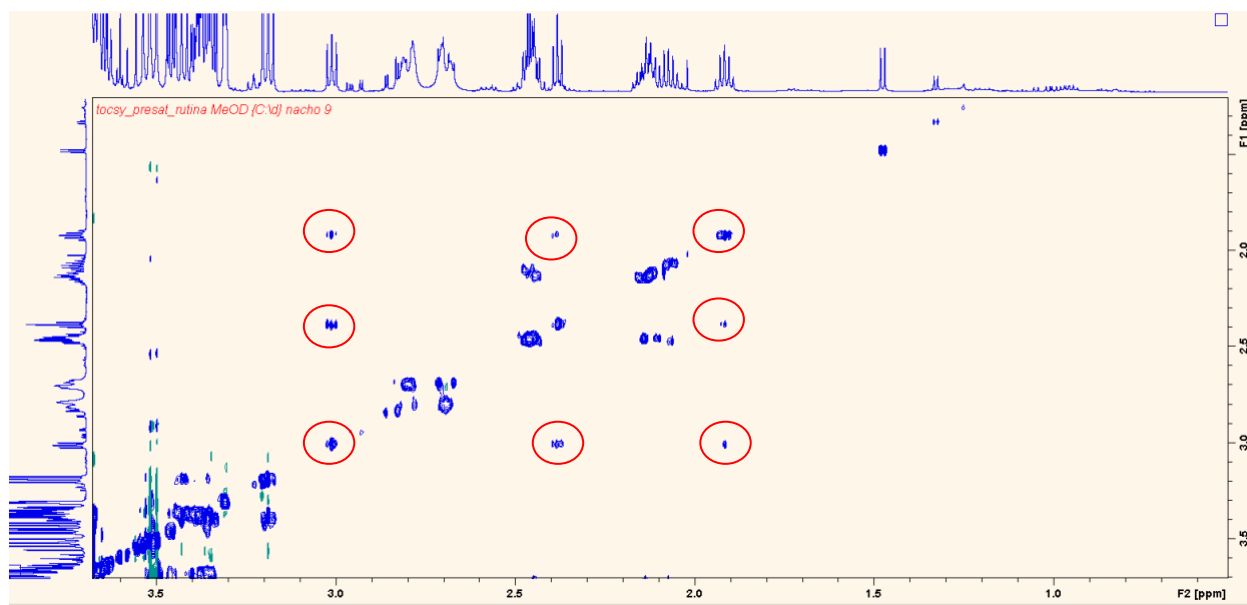


Figure A4 - 2D  $^1\text{H}$ - $^1\text{H}$  TOCSY NMR experiment from tomato  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (50:50, v/v) fruit extracts. TOCSY can provide correlations between all the protons that make up a spin system. Red circles show the correlation between the three type of protons that give rise to the characteristic resonance signals of GABA.

The  $^1\text{H}$ - $^{13}\text{C}$  HSQC experiment allowed the observation of one-bond CH correlations. Figure A5 illustrates how the peaks identified previously in the 2D TOCSY correlate with their corresponding carbon nuclei. By consulting the obtained  $^1\text{H}$  and  $^{13}\text{C}$  NMR chemical shifts in the HMDB database, we confirm the final structure.

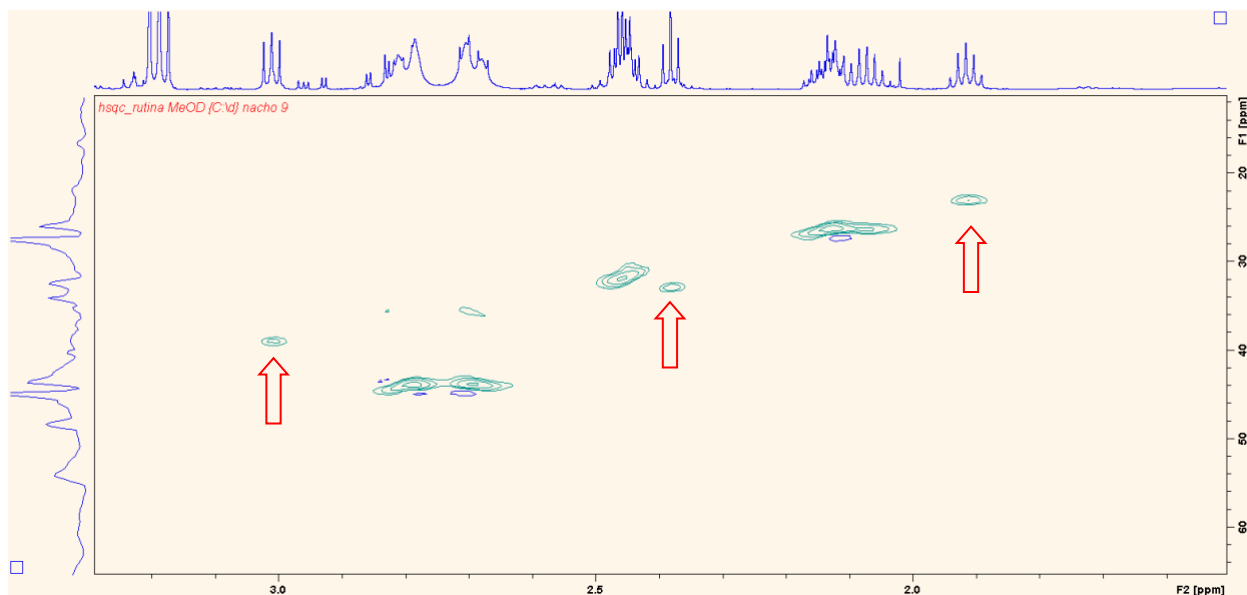


Figure A5 - 2D  $^1\text{H}$ - $^{13}\text{C}$  HSQC NMR experiment from tomato  $\text{CD}_3\text{OD}/\text{D}_2\text{O}$   $\text{KH}_2\text{PO}_4$  buffer (50:50, v/v) fruit extracts.  $^1\text{H}$ - $^{13}\text{C}$  Heteronuclear Single Quantum Coherence Spectroscopy (HSQC) shows which hydrogens are directly attached to which carbon atoms. The  $^1\text{H}$  spectrum is shown on the horizontal axis and the  $^{13}\text{C}$  spectrum is shown on the vertical axis. The HSQC spectrum is most valuable when protons have already been assigned. Red arrows indicate the carbon signal attached to each proton that gives resonance in GABA. With a look at HMDB databases it can be confirmed these atoms belong to the molecule.

## Annex B - Oxygen radical absorbance capacity (ORAC) assay

### *Results processing*

The oxygen radical absorbance capacity was performed as described by Bellán et al., 2008, with some modifications.

After reading the 96-well black polystyrene microplates in a CYTATION microplate reader, from Bio-Tek Instruments, the fluorescence values data was extracted to an excel sheet.

First, the fluorescence values of the Trolox calibration solutions and the blank ( $\text{KH}_2\text{PO}_4$  buffer solution) were plotted along time, Table B1, yielding the Trolox curve, Figure B1.

*Table B1 - Trolox and Blank fluorescence decay values along time. Fluorescence was measured in cycles of 4 minutes.*

| time (min) | Trolox ( $\mu\text{M}$ ) |        |        |        |        | Blanco |
|------------|--------------------------|--------|--------|--------|--------|--------|
|            | 200                      | 100    | 50     | 25     | 12,5   |        |
| 0          | 4867,7                   | 4898,0 | 5092,3 | 5250,0 | 4478,0 | 4805,7 |
| 4          | 4829,3                   | 4817,7 | 4937,0 | 5223,3 | 4450,7 | 4619,0 |
| 8          | 4704,0                   | 4741,0 | 4798,3 | 5232,7 | 4252,0 | 3955,0 |
| 12         | 4706,0                   | 4703,3 | 4730,0 | 5061,0 | 4018,7 | 3360,7 |
| 16         | 4624,0                   | 4677,7 | 4732,7 | 4696,5 | 3667,7 | 2702,0 |
| 20         | 4633,0                   | 4748,7 | 4617,0 | 4360,0 | 3240,0 | 2203,0 |
| 24         | 4579,7                   | 4737,3 | 4393,3 | 3942,0 | 2923,0 | 1730,3 |
| 28         | 4577,0                   | 4676,0 | 4071,3 | 3531,5 | 2374,3 | 1336,7 |
| 32         | 4539,3                   | 4614,7 | 3727,3 | 3267,0 | 1711,0 | 1046,3 |
| 36         | 4475,0                   | 4537,3 | 3392,3 | 2956,0 | 1199,3 | 741,3  |
| 40         | 4502,7                   | 4418,7 | 3243,3 | 2327,5 | 816,3  | 531,0  |
| 44         | 4515,3                   | 4201,3 | 2974,7 | 1684,0 | 546,7  | 301,7  |
| 48         | 4547,3                   | 4046,0 | 2513,7 | 1208,5 | 325,3  | 183,3  |
| 52         | 4533,3                   | 3843,0 | 1782,7 | 824,0  | 190,3  | 115,3  |
| 56         | 4593,3                   | 3593,3 | 1199,3 | 533,0  | 97,3   | 62,0   |
| 60         | 4597,7                   | 3305,0 | 733,3  | 305,0  | 49,3   | 29,7   |
| 64         | 4561,3                   | 2730,7 | 450,7  | 188,0  | 25,7   | 11,7   |
| 68         | 4419,7                   | 2043,0 | 253,7  | 94,5   | 15,3   | 3,3    |
| 72         | 4398,7                   | 1357,0 | 129,3  | 45,5   | 6,3    | 7,7    |

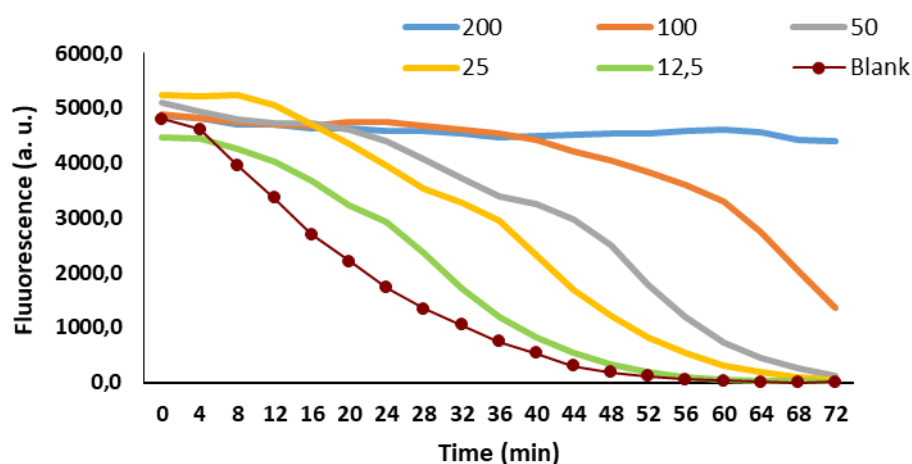


Figure B1 - Trolox fluorescence evolution plotted along time, at different concentrations. Fluorescence values were successively measured every 4 minutes in a microplate reader.

The fluorescence values of each tomato extract sample were also plotted along time, Table 2.

Table B2 - Tomato fruit extracts fluorescence decay values along time. Fluorescence was measured in cycles of 4 minutes.

| TR U9 S1 | TR U9 S2 | TR U9 S3 | TR U9 S4 | TR U9 S5 | Az U9 S1 | Az U9 S2 | Az U9 S3 | Az U9 S4 | Az U9 S5 | SP U11 S1 | SP U11 S2 | SP U11 S3 |
|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|-----------|-----------|-----------|
| 4808,5   | 4684,3   | 4663,3   | 4392,3   | 4687,7   | 4651,7   | 4297,0   | 4391,0   | 4589,5   | 4601,7   | 4280,0    | 4555,0    | 4367,5    |
| 4657,3   | 4759,7   | 4736,7   | 4555,3   | 4651,3   | 4697,0   | 4368,5   | 4438,7   | 4580,7   | 4613,7   | 4328,0    | 4443,0    | 4405,5    |
| 4617,3   | 4709,3   | 4563,7   | 4481,7   | 4536,7   | 4480,5   | 4362,7   | 4443,7   | 4497,7   | 4563,3   | 4424,3    | 4384,3    | 4029,0    |
| 4559,3   | 4549,7   | 4415,7   | 4296,3   | 4465,7   | 4432,7   | 4367,0   | 4345,3   | 4415,7   | 4496,3   | 4306,3    | 4302,3    | 4056,3    |
| 4361,3   | 4494,3   | 4321,0   | 4208,0   | 4328,3   | 4296,7   | 4106,3   | 4327,5   | 4348,0   | 4415,3   | 4076,7    | 4091,3    | 3951,5    |
| 4254,3   | 4335,3   | 4009,3   | 4021,7   | 4169,3   | 4175,0   | 3944,0   | 4057,0   | 4083,7   | 4255,3   | 3846,3    | 3841,3    | 3555,7    |
| 3983,0   | 4028,0   | 3769,0   | 3840,3   | 3918,7   | 3857,5   | 3554,0   | 3779,0   | 3858,0   | 3897,3   | 3392,5    | 3532,7    | 3274,7    |
| 3558,5   | 3696,7   | 3443,3   | 3556,3   | 3642,3   | 3484,7   | 3407,0   | 3525,7   | 3456,0   | 3630,0   | 3160,7    | 3254,3    | 2993,0    |
| 3378,5   | 3335,3   | 3107,7   | 3295,3   | 3332,5   | 3160,5   | 2976,7   | 3239,0   | 3088,5   | 3223,7   | 2707,0    | 2946,3    | 2563,0    |
| 2864,7   | 2894,7   | 2714,3   | 3093,3   | 2884,5   | 2700,0   | 2697,5   | 2838,3   | 2623,5   | 2822,0   | 2211,0    | 2617,0    | 2389,5    |
| 2436,3   | 2503,7   | 2370,7   | 2801,0   | 2526,5   | 2252,3   | 2208,7   | 2458,0   | 2175,0   | 2396,3   | 1815,0    | 2254,3    | 2038,0    |
| 1986,7   | 2065,7   | 1891,0   | 2485,0   | 2099,0   | 1836,3   | 1998,0   | 2097,3   | 1794,5   | 2044,5   | 1407,5    | 1944,0    | 1551,0    |
| 1573,0   | 1586,5   | 1510,5   | 2165,7   | 1714,0   | 1494,7   | 1665,5   | 1743,3   | 1265,0   | 1593,3   | 1078,0    | 1630,3    | 1447,0    |
| 1230,3   | 1324,0   | 1201,5   | 1884,7   | 1334,5   | 1153,7   | 1375,5   | 1429,7   | 1060,0   | 1262,0   | 790,5     | 1309,3    | 1149,5    |
| 901,0    | 1000,7   | 1017,3   | 1598,3   | 1056,0   | 914,5    | 1039,0   | 1116,0   | 580,0    | 964,7    | 523,5     | 1053,7    | 730,0     |
| 622,7    | 682,5    | 635,5    | 1285,3   | 661,7    | 629,3    | 744,0    | 861,7    | 444,7    | 677,0    | 330,5     | 802,3     | 686,0     |
| 423,3    | 503,3    | 423,0    | 1025,0   | 529,0    | 414,3    | 578,5    | 627,0    | 274,7    | 465,0    | 190,5     | 597,3     | 367,0     |
| 234,3    | 339,0    | 291,0    | 792,7    | 274,3    | 262,0    | 351,3    | 441,0    | 148,0    | 304,7    | 94,5      | 414,3     | 316,0     |
| 121,0    | 204,3    | 167,0    | 563,0    | 139,3    | 161,7    | 199,0    | 275,0    | 34,0     | 202,0    | 41,0      | 272,3     | 213,5     |

| SP U11 S4 | SP U11 S5 | RJ U12 S1 | RJ U12 S2 | RJ U12 S3 | RJ U12 S4 | RJ U12 S5 | RJ U11 S1 | RJ U11 S2 | RJ U11 S3 | RJ U11 S4 | RJ U11 S5 |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 4194,3    | 4322,0    | 5157,0    | 4540,0    | 4376,7    | 4862,5    | 4642,3    | 4718,0    | 4523,7    | 4622,5    | 4715,3    | 4611,3    |
| 4286,3    | 4659,7    | 4658,0    | 4582,0    | 4455,3    | 4762,3    | 4830,5    | 4734,7    | 4624,0    | 4774,0    | 4648,3    | 4623,0    |
| 4286,7    | 4497,3    | 4469,0    | 4344,5    | 4334,7    | 4475,5    | 4647,7    | 4571,0    | 4506,7    | 4612,5    | 4623,3    | 4667,3    |
| 4207,7    | 4293,0    | 4272,5    | 4465,0    | 4235,0    | 4384,0    | 4366,5    | 4435,7    | 4355,3    | 4384,5    | 4459,7    | 4502,3    |
| 4096,3    | 4080,7    | 4124,0    | 4218,7    | 4060,7    | 4137,5    | 4314,3    | 4296,3    | 4070,5    | 4208,5    | 4254,7    | 4314,0    |
| 3875,7    | 3776,7    | 4021,3    | 3997,3    | 3827,3    | 3903,0    | 4043,0    | 4011,0    | 4003,7    | 3914,5    | 4030,0    | 4087,3    |
| 3684,7    | 3582,0    | 3770,0    | 3661,7    | 3554,7    | 3622,0    | 3835,0    | 3784,3    | 3708,7    | 3605,5    | 3659,7    | 3863,3    |
| 3417,7    | 3456,0    | 3389,7    | 3273,3    | 3196,3    | 3184,5    | 3578,3    | 3542,7    | 3442,0    | 3108,5    | 2878,6    | 3585,3    |
| 3088,0    | 2889,7    | 3058,7    | 2914,3    | 2809,0    | 2830,0    | 3228,3    | 3229,0    | 3071,7    | 2674,0    | 2972,7    | 3299,0    |
| 2838,7    | 2453,7    | 2704,7    | 2480,0    | 2494,3    | 2408,0    | 3032,7    | 2838,7    | 2923,0    | 2288,0    | 2636,3    | 2943,7    |
| 2487,3    | 2105,7    | 2257,0    | 2084,0    | 2221,3    | 2108,0    | 2712,0    | 2585,3    | 2579,0    | 1833,0    | 2305,7    | 2496,5    |
| 2207,7    | 1843,3    | 1877,7    | 1779,0    | 1941,7    | 1690,5    | 2391,3    | 2216,0    | 2200,0    | 1492,0    | 1937,3    | 2144,5    |
| 1876,3    | 1486,3    | 1527,7    | 1339,0    | 1598,3    | 1395,5    | 2080,7    | 1892,3    | 1925,5    | 1150,0    | 1574,3    | 1814,0    |
| 1559,3    | 1183,0    | 1214,3    | 902,0     | 1322,3    | 1092,5    | 1775,3    | 1541,7    | 1558,0    | 859,0     | 1273,0    | 1503,5    |
| 1299,0    | 898,3     | 921,3     | 659,0     | 1063,3    | 829,0     | 1465,3    | 1240,3    | 1138,0    | 602,0     | 473,1     | 416,5     |
| 1019,3    | 676,0     | 652,7     | 429,0     | 831,7     | 624,0     | 1217,0    | 952,0     | 883,3     | 401,5     | 734,3     | 899,5     |
| 761,7     | 474,3     | 473,7     | 252,0     | 615,0     | 428,0     | 945,3     | 721,3     | 644,3     | 249,0     | 511,7     | 675,0     |
| 560,0     | 297,0     | 296,3     | 115,5     | 438,7     | 271,5     | 722,0     | 518,7     | 551,0     | 135,5     | 337,0     | 491,0     |
| 378,7     | 172,0     | 160,3     | 50,5      | 289,7     | 156,0     | 511,0     | 349,3     | 218,0     | 86,5      | 211,0     | 301,5     |

Then, the area under each fluorescence decay curve (AUC) was calculated by the equation:

$$AUC = \left( 0.5 + \left( \sum_{i=1}^{i=18} \frac{f_i}{f_1} \right) \right) * CT \quad (i)$$

Where  $f_1$  is the initial fluorescence reading at 0 min and  $f_i$  is the fluorescence reading at time  $i$ . CT represents the time of each cycle, 4min.

The net AUC corresponding to the samples was calculated by subtracting the AUC corresponding to the blank. Figure B2 is a visual explanation of the mentioned proceeding. This was carried for Trolox and tomato extract samples, yielding one net AUC for each triplicate, Table B3 and Table B4.

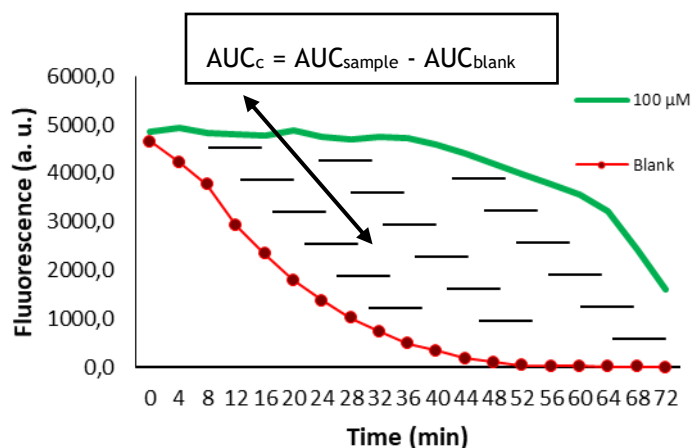


Figure B2 - Visual illustration of the method applied to calculate clean/net AUC of Trolox and every tomato extract sample, through the ORAC method.

Table B4 - Trolox solutions and blank AUC obtained for one independent assay and corresponding clean/net AUC obtained for the Trolox solutions.

|         | Trolox (μM) |      |      |      |      | Blank |
|---------|-------------|------|------|------|------|-------|
|         | 200         | 100  | 50   | 25   | 12,5 |       |
| AUC     | 69,7        | 60,6 | 43,4 | 36,7 | 28,7 | 21,1  |
| AUC net | 48,6        | 39,5 | 22,3 | 15,6 | 7,6  | 0,0   |

After calculating the net AUC (AUC<sub>c</sub>) for each sample, the regression of net AUC of Trolox in different concentrations was performed, Figure B3.

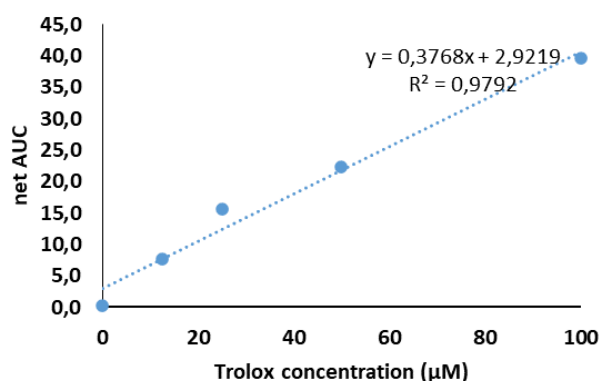


Figure B5 - Example of a Trolox calibration curve used in one of the four independent assays, which enables the quantification of Trolox equivalent micromoles per gram of lyophilized tomato in each sample.

As every sample net AUC is known, solving the regression equation in order to x allows the quantification of Trolox equivalent micromoles (TE) per gram of lyophilized tomato, Table 4.

Table B6 - AUC, clean/net AUC and Trolox equivalent (μM/ g of dried sample) obtained for every tomato fruit extract in one independent assay.

|         | TR U9 S1 | TR U9 S2 | TR U9 S3 | TR U9 S4 | TR U9 S5 | Az U9 S1 | Az U9 S2 | Az U9 S3 | Az U9 S4 | Az U9 S5 | SP U11 S1 | SP U11 S2 | SP U11 S3 | SP U11 S4 |
|---------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|-----------|-----------|-----------|-----------|
| AUC     | 40,1     | 42,1     | 40,2     | 47,5     | 41,5     | 40,2     | 42,9     | 43,9     | 39,2     | 41,8     | 38,2      | 40,4      | 38,4      | 45,8      |
| AUC net | 19,0     | 21,1     | 19,2     | 26,4     | 20,4     | 19,1     | 21,8     | 22,8     | 18,1     | 20,7     | 17,1      | 19,3      | 17,3      | 24,7      |
| TE (μM) | 42,6     | 48,1     | 43,1     | 62,3     | 46,3     | 42,9     | 50,1     | 52,9     | 40,4     | 47,3     | 37,6      | 43,4      | 38,1      | 57,8      |

|         | SP U11 S5 | RJ U12 S1 | RJ U12 S2 | RJ U12 S3 | RJ U12 S4 | RJ U12 S5 | RJ U11 S1 | RJ U11 S2 | RJ U11 S3 | RJ U11 S4 | RJ U11 S5 |
|---------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| AUC     | 41,6      | 36,0      | 38,6      | 41,6      | 36,8      | 44,8      | 42,2      | 43,0      | 36,9      | 38,9      | 42,4      |
| AUC net | 20,5      | 14,9      | 17,5      | 20,5      | 15,7      | 23,7      | 21,1      | 21,9      | 15,8      | 17,8      | 21,4      |
| TE (μM) | 46,8      | 31,8      | 38,7      | 46,6      | 33,9      | 55,2      | 48,4      | 50,5      | 34,3      | 39,5      | 48,9      |

This method was executed in the four independent assays. In the end, results presented are the mean of TE per gram of lyophilized tomato for each plastic covering, regardless of the greenhouse location.

## Annex C - Multivariate Data Analysis

### Tomatoes metabolic evolution along harvests

Table C1 - Full list of harvest discriminating buckets and corresponding metabolites, indicated by the loadings plot relative to the PLS model (Figure 13) applied to the spectra data (X) with harvesting week (Y) as secondary observation. VIP- Variable Importance to Projection.

| INCREASE FOR FIRST HARVESTS |      |                                     | INCREASE FOR LAST HARVESTS |      |                |
|-----------------------------|------|-------------------------------------|----------------------------|------|----------------|
| Loadings                    | VIP  | Metabolites                         | Loadings                   | VIP  | Metabolites    |
| 2.58 (mal)                  | 2.41 | Malic acid                          | 3.46 (glc)                 | 2.62 | Glucose        |
| 4.34 (mal)                  | 2.26 | Malic acid                          | 3.42 (glc)                 | 2.58 | Glucose        |
| 7.34 (phe)                  | 2.12 | Phenylalanine                       | 3.38 (glc)                 | 2.55 | Glucose        |
| 2.14 (glu)                  | 2.11 | Glutamate                           | 5.18 (glc)                 | 2.44 | Glucose        |
| 2.10 (glu)                  | 2.09 | Glutamate                           | 3.18 (glc)                 | 2.33 | Glucose        |
| 3.94 (fru)                  | 2.04 | Fructose                            | 4.58 (glc)                 | 2.28 | Glucose        |
| 7.38 (phe, his)             | 2.04 | Phenylalanine                       | 1.46 (ala, lys)            | 2.12 | Lysine         |
| 2.42 (glu)                  | 1.90 | Glutamate                           | 3.34 (glc)                 | 2.03 | Glucose        |
| 2.46 (glu)                  | 1.87 | Glutamate                           | 1.98 (pro)                 | 2.01 | Proline        |
| 4.06 (fru)                  | 1.72 | Fructose                            | 0.66                       | 2.00 |                |
| 4.02 (fru)                  | 1.71 | Fructose                            | 1.50 (ala, lys)            | 1.95 | Lysine         |
| 6.86 (tyr, HCA))            | 1.63 | Tyrosine                            | 1.14                       | 1.91 | Non identified |
| 2.18 (glu)                  | 1.61 | Glutamate                           | 9.14 (trig)                | 1.82 | Trigoneline    |
| 3.50 (fru)                  | 1.57 | Fructose                            | 8.86 (trig)                | 1.77 | Trigoneline    |
| 1.78                        | 1.53 | -                                   | 0.82 (sterol)              | 1.76 | Sterols        |
| 6.9                         | 1.51 | -                                   | 0.62                       | 1.71 |                |
| 7.0                         | 1.50 | Phenylalanine                       | 2.02 (acet, pro)           | 1.68 | Proline        |
| 4.30 (mal)                  | 1.47 | Malic acid                          | 1.10                       | 1.63 | Non identified |
| 2.82 (cit,asp,asn,mal)      | 1.34 | Malic acid, Citric acid, Asparagine | 0.70                       | 1.62 |                |
| 8.26 (aden-like)            | 1.30 | Adenosine deriv                     | 0.78                       | 1.50 |                |
| 7.42                        | 1.29 | -                                   | 8.10 (trig)                | 1.49 | Trigoneline    |
| 2.78 (cit, mal, asn)        | 1.27 | Malic acid, Citric acid, Asparagine | 6.54 (flav)                | 1.47 | Flavonoids     |
| 9.26                        | 1.18 | -                                   | 6.34 (HCA, flav)           | 1.45 | HCA            |
| 3.22 (MI, cho)              | 1.14 | Myo-inositol                        | 3.82 (sugars)              | 1.38 | Glucose        |
| 2.54 (mal)                  | 1.14 | Malic acid                          | 3.86 (sugars)              | 1.37 | Glucose        |
| 8.54 (form)                 | 1.12 | Formic acid                         | 6.98                       | 1.35 | Flavonoids     |
| 3.10                        | 1.12 |                                     | 1.54                       | 1.32 | Lysine         |
| 6.10 (aden-deriv)           | 1.07 | Adenosine deriv                     | 1.70 (lys, arg)            | 1.30 | Lysine         |
| 7.66                        | 1.04 | -                                   | 0.58                       | 1.24 |                |
| 2.94 (asn)                  | 1.04 | Asparagine                          | 8.82                       | 1.23 | Trigoneline    |
| 2.62                        | 0.98 | Malic acid, Citric acid             | 6.38 (HCA)                 | 1.21 | HCA            |
| 6.82 (tyr)                  | 0.97 | Tyrosine                            | 1.30 (FA)                  | 1.17 | Fatty acids    |

|  |  |  |            |      |             |
|--|--|--|------------|------|-------------|
|  |  |  | 1.86 (lys) | 1.12 | Lysine      |
|  |  |  | 2.34 (pro) | 1.12 | Proline     |
|  |  |  | 7.10       | 1.10 |             |
|  |  |  | 1.58 (FA)  | 1.06 | Fatty acids |

## Tomatoes metabolic profile as a function of the cover colouring

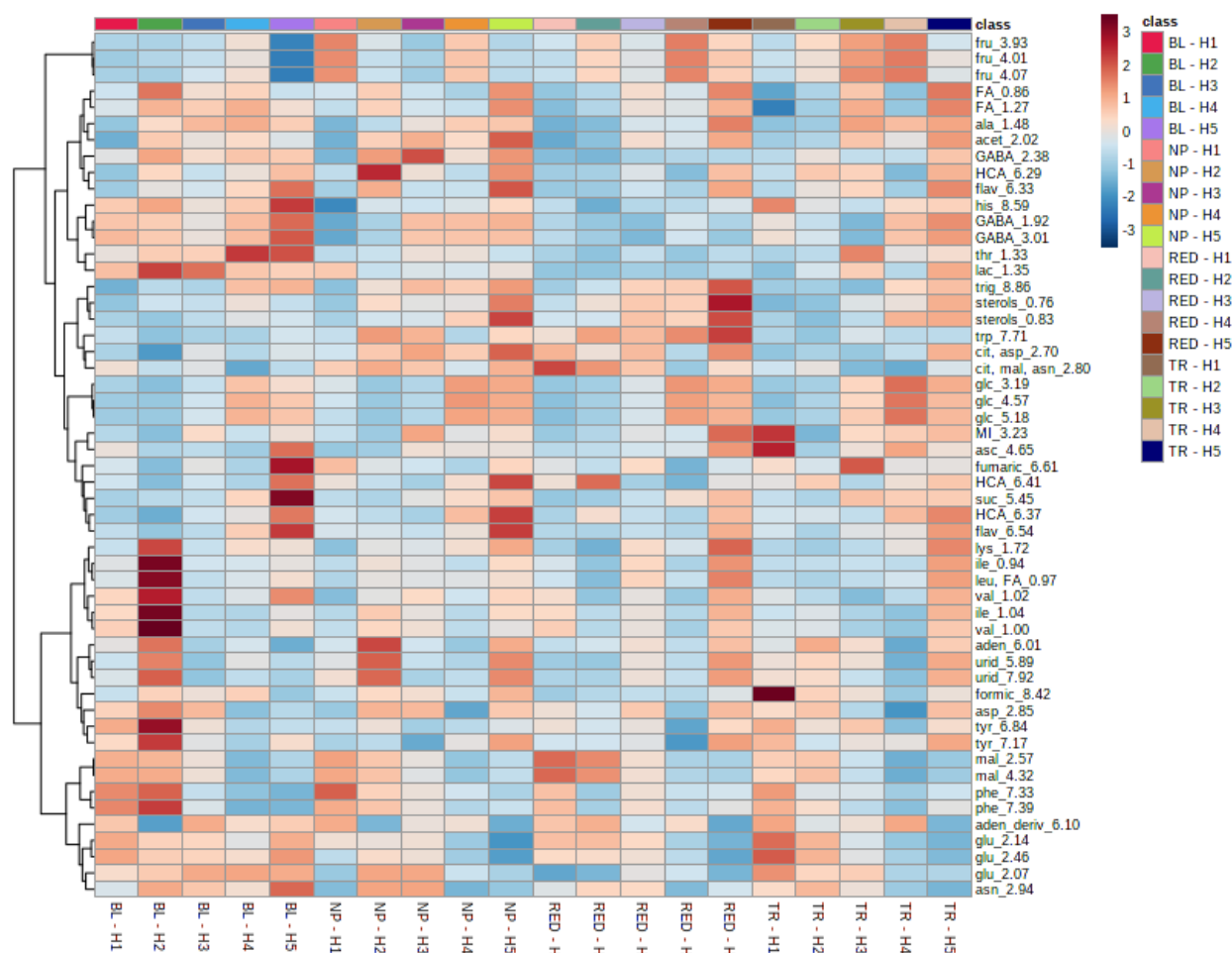


Figure C1 - Heatmap showing z-scores of 53 discriminatory buckets from CD3OD/D2O KH<sub>2</sub>PO<sub>4</sub> buffer (50:50, v/v) extracts of tomato fruits and their assigned metabolites for the four different experimental plastic coverings and each harvesting week. The heatmap construction was based on Euclidean distances and Ward clustering. The X axis represents the average of six biological replicates for each plastic experiment (BL-blue double-roof; NP-no plastic; RED-red double-roof; TR-uncoloured double-roof) from each harvesting stage. The colour scheme represents the variation on the metabolic content among the four conditions in a scale from dark blue (lowest content) to dark red (highest content). In this case, spectroscopic data underwent variable size bucketing. Abbreviations: acet, acetate; aden, adenosine; ala, alanine; asc, ascorbate; asn, asparagine; asp, aspartate; cit, citric acid; FA, fatty acids; flav, flavonoids; formic, formic acid; fru, fructose; fumaric, fumaric acid; GABA,  $\gamma$ -aminobutyric acid; glc, glucose; HCA, hydroxycinnamic acid derivatives; his, histidine; ile, isoleucine; lac, lactate; leu, leucine; lys, lysine; mal, malic acid; MI, myo-inositol; phe, phenylalanine; pro, proline; sterols; suc, sucrose; thr, threonine; trig, trigonelline; trp, tryptophan; tyr, tyrosine; urid, uridine; val, valine.