Plant Cell, Tissue and Organ Culture (PCTOC) (2019) 139:275–294
https://doi.org/10.1007/s11240-019-01679-3

ORIGINAL ARTICLE

Check for updates

# Metabolomic profile of cacao cell suspensions growing in blue light/dark conditions with potential in food biotechnology

Adriana M. Gallego[1,10] · Luisa F. Rojas[2] · Héctor A. Rodriguez[1,3] · Conrado Mora[4] · Lucía Atehortúa[1] · Aura I. Urrea[1] · Mark J. Guiltinan[5,6,7] · Siela N. Maximova[6,7] · Emmanuel Gaquerel[8] · Martha Zuluaga[9] · Natalia Pabón-Mora[10]

Received: 13 April 2019 / Accepted: 21 August 2019 / Published online: 28 August 2019


## Abstract

*Theobroma cacao* is a rich source of flavonoid compounds, which are potent antioxidants. Flavonoids are well-known for their health benefits against cardiovascular diseases, cancer, and improvement of blood pressure. For this reason, cacao mass production has drawn the attention from the functional foods industry. Furthermore, cacao cell suspensions can be used to evaluate complex biosynthetic pathways, such as flavonoids due to the homogeneity of the cell population, the unlimited availability of raw material, the high cell growth and division rates, and the reproducibility of in vitro growth conditions. However, the metabolome of cacao could be affected by exposure to light; especially shorter wavelengths such as blue light trigger targeted flavonoid synthesis. Here, we provide the first report of the metabolomic profile of cacao cell suspensions grown under white-blue and dark conditions. For this, targeted metabolomics was conducted on flavonoids, including bioactive compounds such as catechin, epicatechin and proanthocyanidins (PAs). Moreover, untargeted metabolomics was performed to evaluate the response of the endogenous metabolites exposed to darkness and light. For this, unsupervised and supervised multivariate methods were used. Additionally, a chemical annotation and classification was conducted for the top 50 features obtained from the PLS-DA, in order to identify metabolic pathways that are associated to the light treatments. An increase of glycosylated flavonoids and PAs with higher degrees of polymerization from cells grown under light compared to dark, suggested that light conditions may trigger mechanisms associated with moderate stress. Additionally, lipids, flavonoids, and phytosterols increased after light treatment. The potential of cacao cell suspensions in food biotechnology is discussed, considering that the characterization and quantification of the cacao flavonoid composition are the first steps to evaluate the putative contribution of chocolate to human health.

***Key message*** **Metabolomic profiles of cacao cell suspensions under light and dark conditions suggest that flavonoid modification processes could be involved in defense response under light stress.**

**Keywords** Cacao · Cell suspensions · Metabolomics · Blue LED light · Flavonoids

## Introduction

*Theobroma cacao* seeds are the raw material for the chocolate industry, which is estimated at 90 billion dollars per year (Patras et al. 2014; Anga 2014). Cacao seeds are a rich source of dietary polyphenols, 60% of which are monomeric flavanols (epicatechin and catechin) and proanthocyanidins (PAs) oligomers stored in the cotyledons of cacao seeds (Barnaba et al. 2017). These compounds contribute to the bitter flavor of cocoa products and are responsible for its astringent taste (Magi et al. 2012). Bitterness and astringency are features in high-quality chocolate products and place the study of phenolic compounds at the core of the industry (Elwers et al. 2009). Moreover, cacao flavonoids are strong antioxidants largely due to their *o*-diphenol structure (Elwers et al. 2009). The antioxidant effect of catechins and PAs has been linked to a number of health benefits, including the prevention of chronic and cardiovascular diseases and cancer, as well

Communicated by Francisco de Assis Alves Mourão Filho.

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s11240-019-01679-3) contains supplementary material, which is available to authorized users.

✉ Natalia Pabón-Mora
lucia.pabon@udea.edu.co

Extended author information available on the last page of the article

as potent inflammation inhibition (Sánchez-Rabaneda et al. 2003; Calderón et al. 2009; Gutiérrez-Salmean et al. 2015). Such biological effects occur since they are potent antioxidants and chelators of divalent cations, enzyme inhibitors, and enzyme modulators (Barnaba et al. 2017). Among cacao phenols, PAs have been found to be the most potent antioxidant species (Calderón et al. 2009). Additionally, it has been reported that cocoa contains more phenolic substances than tea and red wine, including a higher antioxidant capacity (Lee et al. 2003). Accordingly, recent epidemiological studies have stressed that phenolic-rich diets including chocolate are associated with beneficial health effects and a longer life expectancy (Ali et al. 2015). Consequently, a health claim on cocoa flavanols based on their properties shown to be beneficial to cardiovascular fitness was granted by the European Food Safety Authority, EFSA (2012).

Several factors affect the quality and quantity of plant polyphenolic compounds. These include genetics (varietal and regional) as well as a large array of environmental variables such as growth conditions, light intensity, humidity, temperature, fertilizer use, wounding and infections, among others (Niemenak et al. 2006; Sampaio et al. 2011; Borges et al. 2013). Light is perhaps one of the most important regulators of growth and development, and it is also key for the biosynthesis of secondary metabolites (Wang et al. 2009; Ghasemzadeh et al. 2010; Fu et al. 2016; Pedroso et al. 2017). Phenolic biosynthesis requires and is enhanced by light (Ghasemzadeh et al. 2010). Wavelengths, mainly blue light, together with light intensity are key factors reported to contribute to the biosynthesis of flavonoids in plants (Manivannan et al. 2015; Zoratti et al. 2014; Pedroso et al. 2017; Li et al. 2018). In cacao, several factors including microclimate and sunlight as well as pod position on the trees have been reported to impact the quantity of polyphenols (Counet et al. 2004; Elwers et al. 2009; Afoakwa et al. 2013; Gu et al. 2013; Krähmer et al. 2015; D'Souza et al. 2017).

Understanding the effects of light on secondary metabolism in field-grown plants can be challenging due to the aforementioned environmental factors influencing flavonoid production. Thus, the use of cell suspension cultures is a powerful platform to evaluate complex biosynthetic pathways and physiological processes at the cellular and molecular levels. This is enabled by the homogeneity in the cell population, large availability of material, high rate of cell growth and reproducibility of cell suspensions (Moscatiello et al. 2013). Cacao polyphenol production has been reported in cell cultures under different conditions and some advances using light conditions to modulate in vitro flavonoid production have been described and patented (Alemanno et al. 2003; Rojas et al. 2008, 2015; Pancaningtyas 2015; Gallego et al. 2016, 2018). This demonstrates the potential of cacao cell cultures as a reliable source of natural products.

Flavonoid metabolic profiles have been well documented in cacao at several time points during seed development (Rigaud et al. 1993; Hammerstone et al. 1999; Bucheli et al. 2001; Stark and Hofmann 2005; Elwers et al. 2009; Pereira-Caro et al. 2013; Wang et al. 2016; D'Souza et al. 2017), in fermented and/or dry seeds (Tomas-Barberán et al. 2007; Patras et al. 2014; Barnaba et al. 2017), as well as in chocolate products such as cocoa powder and cocoa liquor (Sánchez-Rabaneda et al. 2003; Natsume et al. 2000; Wollgast and Anklam 2000; Stark and Hofmann 2006; Calderón et al. 2009; Magi et al. 2012; Ali et al. 2015; D'Souza et al. 2017). Such profiles include other polyphenolic compounds besides flavonoids, for instance, simple phenols, benzoquinones, phenolic acids, acetophenones, phenylacetic acids, hydroxycinnamic acids, *N*-phenylpropenoyl-L-amino acids, phenylpropenes, coumarines, chromones, naphthoquinones, xanthones, stilbenes, anthraquinones, lignans, and lignins (Barnaba et al. 2017). The identification of cacao flavonoids has predominantly been done using a variety of chromatographic techniques for the separation of oligomers, accompanied by independent methods for structural characterization (Rigaud et al. 1993; Hammerstone et al. 1999; Natsume et al. 2000). Today, normal- and reverse-phase high-performance liquid chromatography coupled to mass spectrometry is the most used technique for the identification of flavonoids and their sub-products in cacao seeds at different stages (Counet et al. 2004; Sánchez-Rabaneda et al. 2003; Patras et al. 2014; D'Souza et al. 2017). An accurate characterization of polyphenol composition by LC–MS/MS can contribute to the evaluation of the putative benefits of chocolate to human health (Wollgast and Anklam 2000) and can provide information about biosynthetic responses to variables like light and other abiotic stresses. Furthermore, the characterization of the chemical profile in cacao cell cultures could lead to determining its potential in the design of functional foods.

Despite the fact that cacao has been extensively examined in terms of its chemical composition, very little information and biotechnological approaches are available for cacao cell cultures. Herein, we describe for the first time the metabolic composition of cell cultures derived from cacao seeds, with emphasis on flavonoid content and their potential in biotechnological applications. We have recently reported that light can effectively regulate flavonoid profiles, inducing a faster accumulation of phenolic compounds and shifting of epicatechin/catechin ratios, in particular as a response to switching from white to blue light (Gallego et al. 2018). Here we have analyzed the metabolic profile of cacao cell suspensions grown under white-blue light and dark based on targeted and untargeted metabolomic studies. Finally, our data are discussed in the context of the potential for cacao cell cultures in food biotechnology and human health.

## Materials and methods

### Plant material and growth conditions

Mature, 8-month-old cacao pods of the Trinitario (BIOA accession) were collected at the commercial farm of the Compañía Nacional de Chocolates, during the harvest of August 2016 in the region of San Vicente de Chucurí-Santander, Colombia (06°53′00″N, 73°24′50″W). Cacao cell suspensions were induced following Gallego et al. (2017). The cultures were maintained at 23 ± 2 °C under cool-white fluorescent lamps (12 µmol $m^{-2}$ $s^{-1}$) for 16 h light at 100 rpm. After 12 days, cell suspensions were filtered using a 300 µm mesh and an inoculum (4 g $l^{-1}$ dry weight) was used for establishing cultures. These were sub-cultured every 7 days under dark conditions in order to prepare the suspensions for the light treatments. Media composition was prepared according to Gallego et al. (2017).

### Light treatments

Cell suspension cultures were exposed to two light treatments for 14 days using light-emitting diode (LED) lights. In the first treatment, we evaluated a sequential exposure to 7 days white LED then blue LED (L) for other 7 days. In the second treatment, cells were kept in the dark (D) for 14 days. White and blue LED arrays were designed to give an intensity of 60 µmol $m^{-2}$ $s^{-1}$ under a photoperiod of 16 h light and 8 h dark. Erlenmeyer flasks of 100 ml containing 30 ml of cells were subcultured at day 7 and all samples were collected at day 14 for PAs extraction and further metabolomic processing. Four replicates were collected for each treatment. The same experiment here described was used to obtain transcriptomic data and reported in Gallego et al. (2018).

### Sample preparation

Cell suspension was filtered and washed with ultrapure water. Subsequently, the biomass was lyophilized at − 20 °C. Then 50 mg of dry weight cacao cell suspensions were extracted using 1 ml of extraction solution (70% acetone:29.5% water:0.5% acetic acid), then vortexed for 5 s followed by water bath sonication for 15 min using a benchtop ultrasonic cleaner (Model 2510, Bransonic, Danbury, CT, USA). After sonication, samples were vortexed again and centrifuged at 5000 rpm for 10 min. The samples were re-extracted as before and the supernatants from each extraction were pooled. The degreasing of samples were followed using hexane twice, and then the aqueous phase was filtered through a 0.45 µm polytetrafluoroethylene (PTFE) syringe filter (Millipore, Billerica, MA, USA). This procedure was performed according to Liu et al. (2013). Four biological replicates of each treatment were extracted. To quantify soluble PA levels, the *p*-dimethylamino-cinnamaldehyde (DMACA) method was used. Absorption was measured at 640 nm at 1 min intervals for 20 min to get the maximal readings. The total PA levels were calculated using a standard curve prepared using PAs B2 (Indofine, NJ, USA). For metabolomics analysis, genistein (2.5 µg $ml^{-1}$) was spiked in each sample as an internal standard before HPLC analysis.

### RP-HPLC–MS/MS conditions

Reverse phase HPLC using a Prominence 20 UFLCXR system (Shimadzu, Columbia, MD) was used to perform the metabolomic analysis. 5 µl of extract was injected to a Waters (Milford, MA) BEH $C_{18}$ column (100 mm × 2.1 mm, 1.7 µm particle size). The column was maintained at 55 °C and 20 min aqueous acetonitrile gradient, at a flow rate of 250 µl $min^{-1}$. Solvent A was HPLC grade water with 0.1% formic acid and Solvent B was HPLC grade acetonitrile with 0.1% formic acid. The initial condition was 97% A and 3% B, increasing to 45% B at 10 min and 75% B at 12 min, which was held until 17.5 min before returning to the initial conditions. The eluate was delivered into a 5600 (QTOF) TripleTOF using a Duospray™ ion source (all AB Sciex, Framingham, MA). The capillary voltage was set at 5.5 kV in positive ion mode and 4.5 kV in negative ion mode, with a declustering potential of 80 V. The mass spectrometer was operated in information dependent acquisition mode with a 100 ms survey scan from 100 to 1200 m/z, and up to 20 MS/MS product ion scans (100 ms) per duty cycle using a collision energy of 50 V with a 20 V spread. Analytical standards of catechin, epicatechin, epicatechin gallate, gallocatechin and quercetin were purchased from Sigma (St. Louis, MO, USA). The purity of the standards was 98% and all of these were prepared as stock solutions at 1 mg $l^{-1}$ in methanol. Samples were processed at the metabolomics facility of Penn State University.

### Metabolomics data processing

Accurate precursor masses, MS/MS fragments in negative mode, retention time (RT), and intensities were exported from AB Sciex (Progenesis QI software) to MS-DIAL V. 2.34 (Tsugawa et al. 2015). For data collection, mass range was set from 100 to 1200 m/z, centroid parameter 0.01 Da for MS1 tolerance and 0.05 Da for MS2. For peak detection, linearly weighted smoothing average algorithm was used, setting the smoothing level at 1 scan, minimum peak width at 5 scan, and the minimum peak height at 3000 of amplitudes. For RT peak spotting, the default values were accepted (mass slice of 0.1 m/z and step size of 0.05 m/z).

For peak alignment, filtering, and missing value interpolation we use 0.05 min at RT tolerance and 0.025 Da at MS tolerance.

### Metabolomics data analysis

First at all, the obtained peak areas were normalized to genistein (m/z = 269.0457) internal standard, log transformed and Pareto scaled, missing values imputed, and outliers removed before proceeding to univariate and chemometrics analyses. The input raw data is available in Supplemental Table S1. Statistical analysis was performed using the software MetaboAnalyst 4.0 (Chong et al. 2018). In the flavonoid targeted approach, parameters of metabolite fold change (FC) > 2 and volcano plot with $p$ value < 0.1 and FC > 2 were used for the comparison blue light (L) and dark (D). For the untargeted approach, metabolite FC > 2, and significant changes (FDR adjusted p < 0.05) were calculated using a volcano plot. A partial least squares projection to latent structures discriminant analysis (PLS-DA) were performed between treatments for the complete metabolome. The output consisted of score plots to visualize the contrast between different samples for the two treatments and variable importance in the projection (VIP) loading plots to detect putative biomarkers that explain the cluster separation. All analyses were carried out with four replicates. The peak intensity data were visualized in heatmaps using R. In addition, the web application MetFamily (Treutler et al. 2016) was used to identify light-regulated metabolite families by integrating MS1 abundances with MS/MS spectra. A minimum spectrum intensity of 500 was used. The hierarchical cluster analysis (HCA) was performed using an MS1 abundance threshold of 20,000 counts and a $log_2$-FC (LFC) of two, to identify metabolite families. Relative abundance were used to perform the multivariate statistical analysis.

This was followed by a flavonoid annotation targeted metabolomic approach using a previously in-house designed database of mass spectra (m/z) and RT of known phenolic compounds in cacao beans and chocolate products (Wollgast and Anklam 2000; Natsume et al. 2000; Sánchez-Rabaneda et al. 2003; Stark and Hofmann 2006; Tomas-Barberán et al. 2007; Calderón et al. 2009; Elwers et al. 2009; Magi et al. 2012; Pereira-Caro et al. 2013; Wang et al. 2016; Patras et al. 2014; Ali et al. 2015; Barnaba et al. 2017; D'Souza et al. 2017). The data base was used to identify putative phenolic metabolites in the cacao cell suspensions metabolome. In addition, each metabolite was searched manually on Metlin (https://metlin.scripps.edu/), MassBank (https://massbank.eu/MassBank/index.html), and PubChem (https://pubchem.ncbi.nlm.nih.gov/) databases to confirm their identification. External standards of catechin, epicatechin, quercetin, gallocatechin epigallocatechin and epigallocatechingallate were also run with the same parameters of cacao cell suspensions.

## Results

### Annotation of phenolic compounds in the cacao cell suspensions metabolome

Cacao cell suspensions extracts showed a significant difference (p-value < 0.05) for the total soluble PAs content between treatments with 49.10 mg/ml for L and 29.65 mg/ml for D. Using RP-HPLC–MS/MS, several groups of phenolics previously reported for cacao were identified in the extracts prepared from cacao cell suspension cultures. These included four *N*-phenylpropenoyl-L-amino acids (PPAA) (two aspartic acid-derivatives and two tyrosine-derivatives), a compound belonging to benzoic acids (protocatechuic acid), and a prominent peak of citric acid. Additionally, four sub-groups of flavonoids were also annotated as belonging to flavanone derivatives (naringenin-hexoside, naringin, eriodictyol), flavone and derivatives (luteolin, apigenin-hexosides, luteolin-hexosides), flavanonol (dihydroquercetin, dihydrokaempferol), and flavonol and derivatives (kaempferol, quercetin, quercetin-hexoside). The flavan-3-ol monomers catechin and epicatechin and the well-known oligomeric PAs up to tetramers A-type and B-type were also annotated in the cacao cell suspensions. Finally, two isomers (i) of an unknown peak previously reported for cacao were also detected at *m/z* 329.2336 in 12.01 and 17.48 min RT, respectively. The detailed information of the 63 peaks detected in our cultures in negative mode are shown in Table 1.

The PPAA were annotated with two aspartic acid-derivatives at 278 *m/z* ((+)-*N*-[4′-hydroxy-(E)-cinnamoyl]-L-aspartic acid) and 308 *m/z* ((+)-*N*-[4′-hydroxy-3′-methoxy-(E)-cinnamoyl]-L-aspartic acid), and two tyrosine-derivatives at 358 *m/z* as clovamide (((−)-*N*-[3′,4′-dihydroxy-(E)-cinnamoyl]-3-hydroxy-L-tyrosine) and 326 *m/z* as deoxyclovamide (((−)-*N*-[4′-hydroxy-(E)-cinnamoyl]-L-tyrosine)). All four molecules fragmented in MS2 yielded the characteristics peaks. Citric acid and protocatechuic acid were associated with the peaks at *m/z* 191 and 153, respectively.

In relation to the sub-groups of flavonoids, three peaks related to flavanone derivatives were detected at *m/z* 287, 579, and 433 as eriodictyol, naringenin-7-*O*-neohesperidoside (naringin), and naringenin-hexoside respectively, giving MS/MS peaks characteristic of those compounds. The second sub-group with six flavones were detected at *m/z* 477 (3 luteolin-hexosides), *m/z* 431 (2 apigenin-hexosides), and luteolin (*m/z* 285), which also showed characteristic MS/MS fragments of each molecule. The third sub-group, flavanonols, showed peaks at 303 and 287 associated with dihydroquercetin and dihydrokaempferol (2 isomers). In the last sub-group

**Table 1** RP-HPLC–MS-MS information of phenolic compounds detected in cacao cell suspensions

| Proposed compound | Formula | Ion | RT (min) | Calculated (m/z) | Experimental $[M-H]^-$ (*m/z*) | Δppm | MS/MS (m/z) | References |
|---|---|---|---|---|---|---|---|---|
| Citric acid | $C_6H_8O_7$ | M−H | 1.08 | 191.0197 | 191.0201 | 2.0940 | 57, 87, 111 | D'Souza et al. (2017) |
| Procyanidin A trimer (isomer 1) | $C_{45}H_{36}O_{18}$ | M−H | 1.09 | 863.1829 | 863.1835 | 0.6951 | 287, 575, 711 | Pereira-Caro et al. (2013) and D'Souza et al. (2017) |
| Procyanidin B tetramer (isomer 1) | $C_{60}H_{50}O_{24}$ | M−H | 1.1 | 1153.2619 | 1153.2649 | 2.6013 | 287, 577, 865 | Hammerstone et al. (2009), Pereira-Caro et al. (2013), Patras et al. (2014) and D'Souza et al. (2017) |
| Procyanidin B dimer (isomer 1) | $C_{30}H_{26}O_{12}$ | M−H | 1.11 | 577.1352 | 577.1353 | 0.1733 | 289, 407 | Pereira-Caro et al. (2013), Patras et al. (2014) and Cádiz-Gurrea et al. (2014) |
| Procyanidin B dimer (isomer 2) | $C_{30}H_{26}O_{12}$ | M−H | 1.17 | 577.1352 | 577.1351 | 0.1733 | 289, 407 | Pereira-Caro et al. (2013), Patras et al. (2014) and Cádiz-Gurrea et al. (2014) |
| Procyanidin B dimer (isomer 3) | $C_{30}H_{26}O_{12}$ | M−H | 1.27 | 577.1352 | 577.1338 | 2.4258 | 289, 407 | Cádiz-Gurrea et al. (2014) |
| Procyanidin B trimer (isomer 1) | $C_{45}H_{38}O_{18}$ | M−H | 1.92 | 865.1985 | 865.1999 | 1.6181 | Non fragmented | Pereira-Caro et al. (2013), Patras et al. (2014), Cádiz-Gurrea et al. (2014) and D'Souza et al. (2017) |
| Protocatechuic acid | $C_7H_6O_4$ | M−H | 2.8 | 153.0193 | 153.0194 | 0.6535 | 109 | Barnaba et al. (2017) |
| Procyanidin B trimer (isomer 2) | $C_{45}H_{38}O_{18}$ | M−H | 3.03 | 865.1985 | 865.2019 | 3.9297 | 287, 577 | Pereira-Caro et al. (2013), Patras et al. (2014), Cádiz-Gurrea et al. (2014) and D'Souza et al. (2017) |
| Procyanidin A trimer (isomer 2) | $C_{45}H_{36}O_{18}$ | M−H | 3.05 | 863.1829 | 863.1821 | 0.9268 | Non fragmented | Pereira-Caro et al. (2013) and D'Souza et al. (2017) |
| Procyanidin B dimer (dimer 4) | $C_{30}H_{26}O_{12}$ | M−H | 3.11 | 577.1352 | 577.1373 | 3.6387 | 289, 407 | Cádiz-Gurrea et al. (2014) |
| Procyanidin B trimer (isomer 3) | $C_{45}H_{38}O_{18}$ | M−H | 3.17 | 865.1985 | 865.2020 | 4.0453 | 289, 407, 713 | Pereira-Caro et al. (2013), Patras et al. (2014), Cádiz-Gurrea et al. (2014) and D'Souza et al. (2017) |
| Procyanidin A dimer (isomer 1) | $C_{30}H_{24}O_{12}$ | M−H | 3.18 | 575.1195 | 575.1189 | 1.0433 | Non fragmented | Pereira-Caro et al. (2013) and Cádiz-Gurrea et al. (2014) |
| Procyanidin B (dimer 5) B1 | $C_{30}H_{26}O_{12}$ | M-H | 3.19 | 577.1352 | 577.1364 | 2.0792 | 289, 407 | Cádiz-Gurrea et al. (2014) |
| Procyanidin B dimer (isomer 6) | $C_{30}H_{26}O_{12}$ | M−H | 3.25 | 577.1352 | 577.1349 | 0.5198 | 289, 407 | Pereira-Caro et al. (2013), Patras et al. (2014) and Cádiz-Gurrea et al. (2014) |
| Procyanidin A dimer (isomer 2) | $C_{30}H_{24}O_{12}$ | M−H | 3.33 | 575.1195 | 575.1197 | 0.3478 | 287, 423 | Pereira-Caro et al. (2013) and Cádiz-Gurrea et al. (2014) |
| Procyanidin B dimer (dimer 7) | $C_{30}H_{26}O_{12}$ | M−H | 3.33 | 577.1352 | 577.1343 | 1.5594 | 289, 407 | Cádiz-Gurrea et al. (2014) |
| Procyanidin B dimer (dimer 8) | $C_{30}H_{26}O_{12}$ | M−H | 3.4 | 577.1352 | 577.137 | 3.1189 | 289, 407 | Cádiz-Gurrea et al. (2014) |
| Catechin | $C_{15}H_{14}O_6$ | M−H | 3.77 | 289.0718 | 289.0717 | 0.3459 | 203, 245 | Pereira-Caro et al. (2013) and Cádiz-Gurrea et al. (2014) |
| Procyanidin B trimer (isomer 4) | $C_{45}H_{38}O_{18}$ | M−H | 3.78 | 865.1985 | 865.1982 | 0.3467 | 289, 407, 577, 713 | Tomas-Barberán et al. (2007), Pereira-Caro et al. (2013), Patras et al. (2014), Cádiz-Gurrea et al. (2014) and D'Souza et al. (2017) |

**Table 1** (continued)

| Proposed compound | Formula | Ion | RT (min) | Calculated (m/z) | Experimental $[M-H]^-$ (*m/z*) | Δppm | MS/MS (m/z) | References |
|---|---|---|---|---|---|---|---|---|
| Procyanidin B tetramer (isomer 2) | $C_{60}H_{50}O_{24}$ | M−H | 3.8 | 1153.2619 | 1153.2668 | 4.2488 | 287,407, 575, 865, 983 | Hammerstone et al. (2009), Pereira-Caro et al. (2013), Patras et al. (2014) and D'Souza et al. (2017) |
| Procyanidin B dimer (isomer 9) | $C_{30}H_{26}O_{12}$ | M−H | 3.83 | 577.1352 | 577.1352 | 0.0000 | Non fragmented | Tomas-Barberán et al. (2007), Pereira-Caro et al. (2013), Patras et al. (2014) and Cádiz-Gurrea et al. (2014) |
| Procyanidin B trimer (isomer 5) | $C_{45}H_{38}O_{18}$ | M−H | 3.96 | 865.1985 | 865.1982 | 0.3467 | 287, 407, 577, 695 | Tomas-Barberán et al. (2007), Pereira-Caro et al. (2013), Patras et al. (2014), Cádiz-Gurrea et al. (2014) and D'Souza et al. (2017) |
| Procyanidin B dimer (dimer 10) | $C_{30}H_{26}O_{12}$ | M−H | 4.04 | 577.1352 | 577.1334 | 3.1189 | 289, 407 | Cádiz-Gurrea et al. (2014) |
| Procyanidin B trimer (isomer 6) | $C_{45}H_{38}O_{18}$ | M−H | 4.04 | 865.1985 | 865.2007 | 2.5428 | 287, 407, 577, 695 | Pereira-Caro et al. (2013), Patras et al. (2014), Cádiz-Gurrea et al. (2014) and D'Souza et al. (2017) |
| (+)-*N*-[4′-hydroxy-(E)-cinnamoyl]-L-aspartic acid | $C_{13}H_{13}NO_6$ | M−H | 4.07 | 278.0670 | 278.0675 | 1.7981 | 162, 234 | Cádiz-Gurrea et al. (2014) |
| Procyanidin B (dimer 11) B2 | $C_{30}H_{26}O_{12}$ | M−H | 4.11 | 577.1352 | 577.1347 | 0.8663 | 289, 407 | Cádiz-Gurrea et al. (2014) |
| Procyanidin A tetramer (isomer 1) | $C_{60}H_{48}O_{24}$ | M−H | 4.18 | 1151.2463 | 1151.2467 | 0.3474 | 575, 863 | Cádiz-Gurrea et al. (2014) and D'Souza et al. (2017) |
| Procyanidin B tetramer (isomer 3) | $C_{60}H_{50}O_{24}$ | M−H | 4.18 | 1153.2619 | 1153.2634 | 1.3007 | 287, 407, 575, 865, 983 | Hammerstone et al. (2009), Pereira-Caro et al. (2013), Patras et al. (2014) and D'Souza et al. (2017) |
| Epicatechin | $C_{15}H_{14}O_6$ | M−H | 4.47 | 289.0718 | 289.0726 | 2.7675 | 203, 245 | Pereira-Caro et al. (2013) and Cádiz-Gurrea et al. (2014) |
| Procyanidin A trimer (isomer 3) | $C_{45}H_{36}O_{18}$ | M−H | 4.59 | 863.1829 | 863.184 | 1.2744 | Non fragmented | Pereira-Caro et al. (2013) and D'Souza et al. (2017) |
| Procyanidin B trimer (isomer 7) | $C_{45}H_{38}O_{18}$ | M−H | 4.59 | 865.1985 | 865.2003 | 2.0804 | 289, 407, 577, 695 | Pereira-Caro et al. (2013), Patras et al. (2014), Cádiz-Gurrea et al. (2014) and D'Souza et al. (2017) |
| (+)-*N*-[4′-hydroxy-3′-methoxy-(E)-cinnamoyl]-L-aspartic acid | $C_{14}H_{15}NO_7$ | M−H | 4.62 | 308.0776 | 308.0783 | 2.2722 | 134 | Cádiz-Gurrea et al. (2014) |
| Procyanidin B tetramer (isomer 4) | $C_{60}H_{50}O_{24}$ | M−H | 4.71 | 1153.2619 | 1153.2644 | 2.1678 | 865 | Hammerstone et al. (2009), Pereira-Caro et al. (2013), Patras et al. (2014) and D'Souza et al. (2017) |
| Procyanidin A tetramer (isomer 2) | $C_{60}H_{48}O_{24}$ | M−H | 4.72 | 1151.2463 | 1151.2468 | 0.4343 | Non fragmented | Cádiz-Gurrea et al. (2014) and D'Souza et al. (2017) |
| Procyanidin A tetramer (isomer 3) | $C_{60}H_{48}O_{24}$ | M−H | 4.84 | 1151.2463 | 1151.2461 | 0.1737 | Non fragmented | Cádiz-Gurrea et al. (2014) and D'Souza et al. (2017) |
| Procyanidin A trimer (isomer 4) | $C_{45}H_{36}O_{18}$ | M−H | 4.85 | 863.1829 | 863.1852 | 2.6646 | 287, 449, 575, 711 | Pereira-Caro et al. (2013) and D'Souza et al. (2017) |

**Table 1** (continued)

| Proposed compound | Formula | Ion | RT (min) | Calculated (m/z) | Experimental $[M-H]^-$ (*m/z*) | Δppm | MS/MS (m/z) | References |
|---|---|---|---|---|---|---|---|---|
| Procyanidin A dimer (isomer 3) | $C_{30}H_{24}O_{12}$ | M−H | 4.89 | 575.1195 | 575.1198 | 0.5216 | Non fragmented | Pereira-Caro et al. (2013) and Cádiz-Gurrea et al. (2014) |
| Clovamide ((−)-*N*-[3′,4′-dihydroxy-(E)-cinnamoyl]-3-hydroxy-L-tyrosine) | $C_{18}H_{17}NO_7$ | M−H | 5.02 | 358.0932 | 358.0908 | 6.7022 | 135, 178, 222 | D'Souza et al. (2017) |
| Luteolin-hexoside (isomer 1) | $C_{21}H_{20}O_{11}$ | M−H | 5.38 | 447.0933 | 447.0931 | 0.4473 | 285, 327 | Sánchez-Rabaneda et al. (2003) |
| Procyanidin B dimer (isomer 12) | $C_{30}H_{26}O_{12}$ | M−H | 5.44 | 577.1352 | 577.1349 | 0.5198 | 289, 407 | Pereira-Caro et al. (2013), Patras et al. (2014) and Cádiz-Gurrea et al. (2014) |
| Procyanidin B trimer (isomer 8) | $C_{45}H_{38}O_{18}$ | M−H | 5.47 | 865.1985 | 865.2000 | 1.7337 | 289, 407, 577 | Pereira-Caro et al. (2013), Patras et al. (2014), Cádiz-Gurrea et al. (2014) and D'Souza et al. (2017) |
| Procyanidin B dimer (dimer 13) | $C_{30}H_{26}O_{12}$ | M−H | 5.56 | 577.1352 | 577.1347 | 0.8663 | 289, 407 | Cádiz-Gurrea et al. (2014) |
| Dihydroquercetin | $C_{15}H_{12}O_7$ | M−H | 5.75 | 303.0510 | 303.051 | 0.0000 | 125, 285 | Metlin data base |
| Procyanidin B tetramer (isomer 5) | $C_{60}H_{50}O_{24}$ | M−H | 5.85 | 1153.2619 | 1153.2642 | 1.9943 | 865 | Hammerstone et al. (2009), Pereira-Caro et al. (2013), Patras et al. (2014) and D'Souza et al. (2017) |
| Luteolin-hexoside (isomer 2) | $C_{21}H_{20}O_{11}$ | M−H | 5.89 | 447.0933 | 447.0931 | 0.4473 | 285, 327 | Sánchez-Rabaneda et al. (2003) |
| Procyanidin A dimer (isomer 4) | $C_{30}H_{24}O_{12}$ | M−H | 6.13 | 575.1195 | 575.1191 | 0.6955 | 285, 449 | Pereira-Caro et al. (2013) and Cádiz-Gurrea et al. (2014) |
| Apigenin-hexoside (isomer 1) | $C_{21}H_{20}O_{10}$ | M−H | 6.15 | 431.0984 | 431.0983 | 0.2320 | 269 | Sánchez-Rabaneda et al. (2003) and D'Souza et al. (2017) |
| Deoxyclovamide ((−)-*N*-[4′-hydroxy-(E)-cinnamoyl]-L-tyrosine) | $C_{18}H_{17}NO_5$ | M−H | 6.33 | 326.1045 | 326.1041 | 1.2266 | 301 | Tomas-Barberán et al. (2007) and Cádiz-Gurrea et al. (2014) |
| Naringenin-7-*O*-neohesperidoside (naringin) | $C_{27}H_{32}O_{14}$ | M−H | 6.57 | 579.1719 | 579.1724 | 0.8633 | 271 | Sánchez-Rabaneda et al. (2003) |
| Dihydrokaempferol (isomer 1) | $C_{15}H_{12}O_6$ | M−H | 6.66 | 287.0561 | 287.0565 | 1.3935 | 125, 259 | Metlin database |
| Luteolin-hexoside (isomer 3) | $C_{21}H_{20}O_{11}$ | M−H | 6.84 | 447.0933 | 447.0932 | 0.2237 | 285 | Sánchez-Rabaneda et al. (2003) and D'Souza et al. (2017) |
| Quercetin hexoside | $C_{21}H_{20}O_{12}$ | M−H | 6.91 | 463.0882 | 463.088 | 0.4319 | 301 | Pereira-Caro et al. (2013), Cádiz-Gurrea et al. (2014) and D'Souza et al. (2017) |
| Naringenin-hexoside | $C_{21}H_{20}O_{12}$ | M−H | 7.2 | 433.1140 | 433.1143 | 0.6927 | 271 | Sánchez-Rabaneda et al. (2003) |
| Epicatechin methyl dimer | $C_{32}H_{30}O_{12}$ | M−H | 7.21 | 605.1664 | 605.165 | 2.3134 | 289, 453 | Cádiz-Gurrea et al. (2014) |
| Dihydrokaempferol (isomer 2) | $C_{15}H_{12}O_6$ | M−H | 7.3 | 287.0561 | 287.0565 | 1.3935 | 243, 269 | Metlin database |
| Apigenin-hexoside (isomer 2) | $C_{21}H_{20}O_{10}$ | M−H | 7.45 | 431.0984 | 431.0981 | 0.6959 | 269 | Sánchez-Rabaneda et al. (2003) and D'Souza et al. (2017) |
| Eriodictyol | $C_{15}H_{11}O_6$ | M−H | 7.77 | 287.0561 | 287.0563 | 0.6967 | 125, 151, 135 | Metlin database |
| Luteolin | $C_{15}H_{10}O_6$ | M−H | 7.91 | 285.0405 | 285.0409 | 1.4033 | 133, 199 | Sánchez-Rabaneda et al. (2003) and D'Souza et al. (2017) |
| Quercetin | $C_{15}H_{10}O_7$ | M−H | 7.91 | 301.0354 | 301.0355 | 0.3322 | 151, 211 | Sánchez-Rabaneda et al. (2003) |

**Table 1** (continued)

| Proposed compound | Formula | Ion | RT (min) | Calculated (m/z) | Experimental $[M-H]^-$ (*m/z*) | Δppm | MS/MS (m/z) | References |
|---|---|---|---|---|---|---|---|---|
| Kaempferol | $C_{15}H_{10}O_6$ | M−H | 8.8 | 285.0405 | 285.0405 | 0.0000 | 117, 133 | Sánchez-Rabaneda et al. (2003) and D'Souza et al. (2017) |
| Unknown reported (isomer 1) | $C_{18}H_{34}O_5$ | M−H | 12.01 | 329.2334 | 329.2336 | 0.6075 | 171 | D'Souza et al. (2017) |
| Unknown reported (isomer 2) | $C_{18}H_{34}O_5$ | M−H | 17.48 | 329.2334 | 329.2336 | 0.6075 | 171 | D'Souza et al. (2017) |

belonging to flavonols, three peaks were detected at *m/z* 463, 301, and 285, which were associated with quercetin-hexoside, quercetin, and kaempferol, respectively. Of these, quercetin was also compared with an authentic standard.

The monomers catechin and epicatechin yielded the deprotonated molecule $[M-H]^-$ (*m/z* 289) at 3.77 and 4.47 min, respectively. The MS/MS spectrum showed two peaks of 203 and 245 in both molecules. The masses and RT were compared with catechin and epicatechin external standards. One peak at 605 *m/z* was associated with the epicatechin methyl dimer, which showed 289 and 243 ion products in MS/MS. In relation to the PAs, A-type (4 dimers, 4 trimers, 3 tetramers) and B-type (13 dimers, 8 trimers, and 5 tetramers) were annotated in our cell cultures. A-type dimers showed $[M-H]^-$ ions at *m/z* 575, consistent with catechin and/or epicatechin-based dimers, giving ions under the MS2 spectra at 287, 423, 285, and 449. B-type dimers were the most abundant showing 13 isomers. These yielded ions at *m/z* 577 in MS/MS giving same ion products of 289 and 407. A-type trimers, with the precursor ion *m/z* 863, fragmented into peaks at 287, 575, and 711 peaks. B-type trimers (*m/z* 865) showed ions at 287, 289, 407, 577, 695, and 713. A-type tetramers at *m/z* 1151 $[M-H]^-$ fragmented at 575 and 863 in MS/MS. Lastly, B-type tetramers (*m/z* 1153) fragmented showing similar peaks at 287, 407, 575, 865, and 983. All the known phenolic compounds were distributed within the first 9 min of the runs.

Furthermore, two isomers of an unknown molecule were prominent at 329 *m/z* showing 171 *m/z* fragment in MS/MS in both cases. Based on previous reports for cacao, these were tentatively assigned as the organic acid 9,10,13-trihydroxyoctadec-11-enoic acid (9,10,13-TriHOME). Moreover, MUFAs, PUFAs and, acylglycerols were found increased in the top 50 features from the PLS-DA in the light treatment and classified in lipid chemical class (Supplemental Fig. S1). Thus, the high relative abundance of 9,10,13-TriHOME is in accordance with the relative abundance of lipids. Interestingly, in this profile, we did not detect the hydroxycinnamic acids prevalent in cacao (caffeic, ferulic, *p*-coumaric acids, and chlorogenic acid). Similarly, we did not find peaks associated with methylxanthines (theobromine, 179 m/z and caffeine, 193 m/z) and anthocyanins (cyaniding-3-*O*-arabinoside, 419.1 m/z and cyaniding3-*O*-galactoside, 449.2 m/z) when checking in positive mode. The detailed MS/MS fragmentation pattern for the 63 phenolic compounds is presented in the Supplemental material (File Metabolites spectra).

## Metabolomic analysis of phenolic compounds under blue light and dark conditions

We compared the phenolic compounds in extracts from cacao cell suspensions grown under white-blue LED light

(L) and darkness (D). Samples were taken at the end of the experiment, under blue light and dark, in order to detect differences due to light treatment. A FC comparison between L/D showed 43 peaks with a threshold higher than 2 in L and 4 peaks higher in D (Supplemental Fig. S2). A volcano plot (FC higher than 2 and p-value < 0.1) showed four peaks overexpressed in the blue light and these were annotated as two B-type PAs (isomers of dimer and tetramer, respectively), luteolin, and one of the unknown compounds (9,10,13-TriHOME, isomer 2) (Fig. 1).

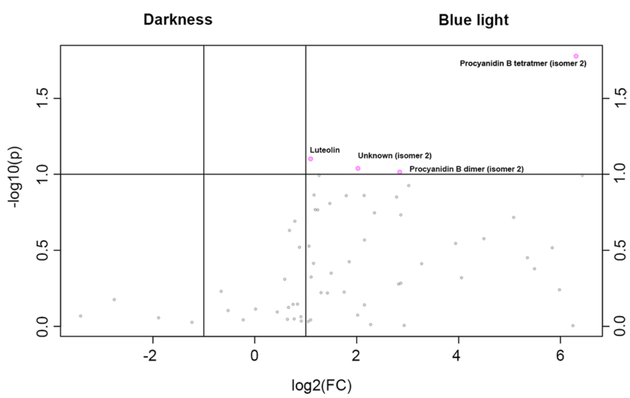


**Fig. 1** Volcano plot for blue light and darkness comparison showing significant phenolic and miscellaneous compounds expressed. Fold change > 2 and p-value < 0.1. (Color figure online)

To gain a better understanding of the 63 phenolic compounds annotated and their relative abundances, we constructed a biochemical map to visualize these compounds (Fig. 2). Citric acid (citrate) and protocatechuic acid were associated with the glycolysis and phenylpropanoid pathways, respectively. The rest of the compounds, except the unknown metabolites, were distributed throughout the flavonoid pathway. The average relative abundance of each metabolite in L and D, evaluated as the metabolite peak area, was calculated using all four biological replicates for each treatment and visualized as a heatmap. The metabolites were mainly concentrated under the blue light, with 51 metabolites out of 63 showing higher peak areas. Interestingly, blue light-induced metabolites included the monomeric forms of most of the oligomeric PAs, all flavanone-hexosides, all flavonols except quercetin, most flavones, and all PPAA. Additionally,

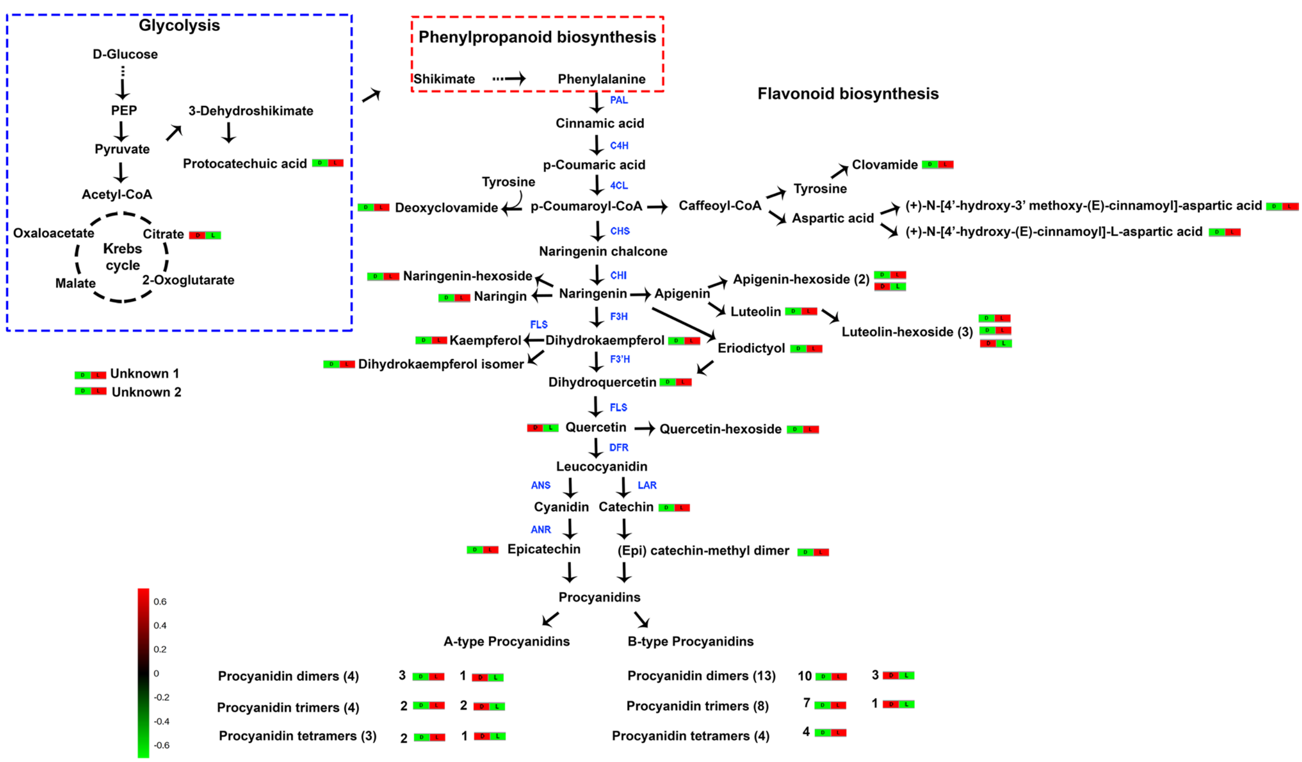


**Fig. 2** Visualization of 63 phenolic and related metabolites found in cacao cell suspensions in a biochemical pathway. Glycolysis (blue dotted square). Phenylpropanoid (red dotted square). Flavonoid (no square). Metabolites in black letters, enzymes in blue letters. Metabolite levels were averaged over four biological replicates after normalization. This is shown in a heatmap for each treatment dark (D) and blue light (L). Metabolite concentrations from low to high are colored from green to red. *PAL* phenylalanine ammonia lyase, *C4H* cinnamate 4-hydroxylase, *4CL* 4-coumarate coenzyme A ligase, *CHS* chalcone synthase, *CHI* chalcone isomerase, *F3H* flavanone 3-hydroxylase, *F3′H* flavanone 3′-hydroxylase, *FLS* flavonol synthase, *DFR* dihydroflavonol-4-reductase, *ANS* anthocyanidin synthase, *ANR* anthocyanidin reductase, *LAR* leucoanthocyanidin reductase. (Color figure online)

light-inhibited metabolites included citric acid, quercetin, two flavones (luteolin-hexoside isomer 3 and apigenin-hesoxide isomer 2), four type A PAs (1 dimer, 2 trimers, 1 tetramer), and four type B PAs (3 dimers, 1 trimer).

We decided to use 43 cacao flavonoid genes from a previous transcriptomic study obtained under the same experiment conducted in this study (Gallego et al. 2018) to perform a Pearson's correlation with the 63 phenolic compounds (metabolomic data) found here, in order to detect putative relationships between metabolites and genes (Supplemental Fig. S3). It is well-known that metabolite accumulation continues after changes in gene expression occur, due to several processes that take place between the expression of a structural gene and the metabolite produced by its encoded enzyme. Examples of such processes include the efficiency of the translation, which may dramatically differ between different genes, the catalytic efficiency of different enzymes varying in orders of magnitude, and finally, the metabolic flux rate, which is not only determined by the concentration of the enzymes but also by a multitude of regulators (positive and negative transcription factors) (Hoppe 2012). Both metabolomic and transcriptomic data were compared under blue light conditions. However, the conclusions derived from this correlation between the flavonoid metabolites and genes studied could require further support, given that this correlation was made for a single time-point. The X axis represents the structural flavonoid genes and the Y axis represents the metabolites (Supplemental Fig. S3). Strong positive correlations (higher than 0.8) with most of the phenolic metabolites were detected for the upregulation of the genes that code for flavonol synthase (*FLS* transcript 2, 3 and 4), two dihydroflavonol-4-reductases (*DFR* transcripts 3 and 4), and one UDP-glucose flavonoid 3-*O*-glucosyltransferase 6 (*UFGT* transcript 7). Interestingly, the genes for the 4-coumarate coenzyme A ligase (*4CL* transcript 4) and leucoanthocyanidin reductase (*LAR* transcript 1), showed negative correlations with most of the metabolites (pcorr > 0.3), yet these correlations were weak. Furthermore, dimers A2 and A3 and quercetin showed positive correlations with several genes. Conversely, only tetramer A3 showed negative correlations with several structural genes (pcorr > 0.5). The annotation and locus name for the transcripts of the flavonoid biosynthetic genes are detailed in the Supplemental Table S2.

Additionally, we plotted Pearson's correlations between pairs of phenolic metabolites (Supplemental Fig. S4). As we expected, the levels of most of the metabolites were highly positively correlated (pcorr > 0.5). Interestingly, we detected that PAs B trimer (i8) had negative correlations with 24 other metabolites, particularly citric acid. Similarly, a weak negative correlation was identified for PAs trimer A (i4).

## Untargeted metabolomic analysis of cacao cell suspensions under blue and dark conditions

An untargeted metabolomics approach was used to determine the impact of light on the metabolome of cacao cell suspensions via HPLC–MS/MS. A partial least squares projection to latent structures discriminant analysis (PLS-DA) was performed to identify class differences between light treatments for all peaks detected. This method aims to maximize the covariance between the independent variables (metabolomics data) and the corresponding dependent Y variable (blue light and darkness) of highly multidimensional data by finding a linear subspace of the explanatory variables. The PLS-DA score plot revealed obvious differences between samples exposed to blue light and darkness (Fig. 3). All samples at L clustered together and were completely separated from samples in D. This indicated clear differences between the sample groups and suggested differences in the metabolic responses under the treatments. The abscissa represents component 1, while the ordinate represents component 2. The PLS-DA revealed that the two highest-ranking principal components accounted for 60.5% of the total variance within the dataset. In particular, more variation within the samples was present in L compared to D, in which samples were tightly clustered (Fig. 4a). In order to investigate the contributors of the separation in the components of the PLS-DA, the metabolic loadings in the VIP were compared for the treatments. Here, 149 and 843 peaks were detected with values > 2.0 and > 1.0, respectively, in the first component (Supplemental Table S3). The top 20 peaks contributing to the separation of the treatments are shown in Fig. 4b. Three peaks previously detected in the volcano plot as PAs B tetramer, luteolin, and 9,10,13-TriHOME were also annotated within the 843 biomarkers also contributing to the separation of the treatments. The t-test in the volcano plot for the complete metabolome (FDR adjusted < 0.05 and FC threshold > 2) showed five differentially expressed peaks under the blue light (Supplemental Fig. S5). The peaks showed *m*/*z* at 503.3363, 297.2449, 283.2657, 465.2069, and 688.3663. A particular characteristic of these differentially expressed metabolites was their elution late in the run, suggesting that they could be non-polar compounds. Furthermore, the PLS-DA VIP features were annotated using public libraries and described above. The top 50 of metabolites were further annotated and classified by chemical class, and a descriptive analysis was also performed using the relative abundance of each class to highlight the differences in terms of chemical groups between light and dark exposure (Supplemental Fig. S1). Differences on relative abundances between treatments for esters, flavonoids, lipids, nucleosides sugars, and tannins were evident between treatments. Particularly, flavonoids and lipids were higher in L compared to D.

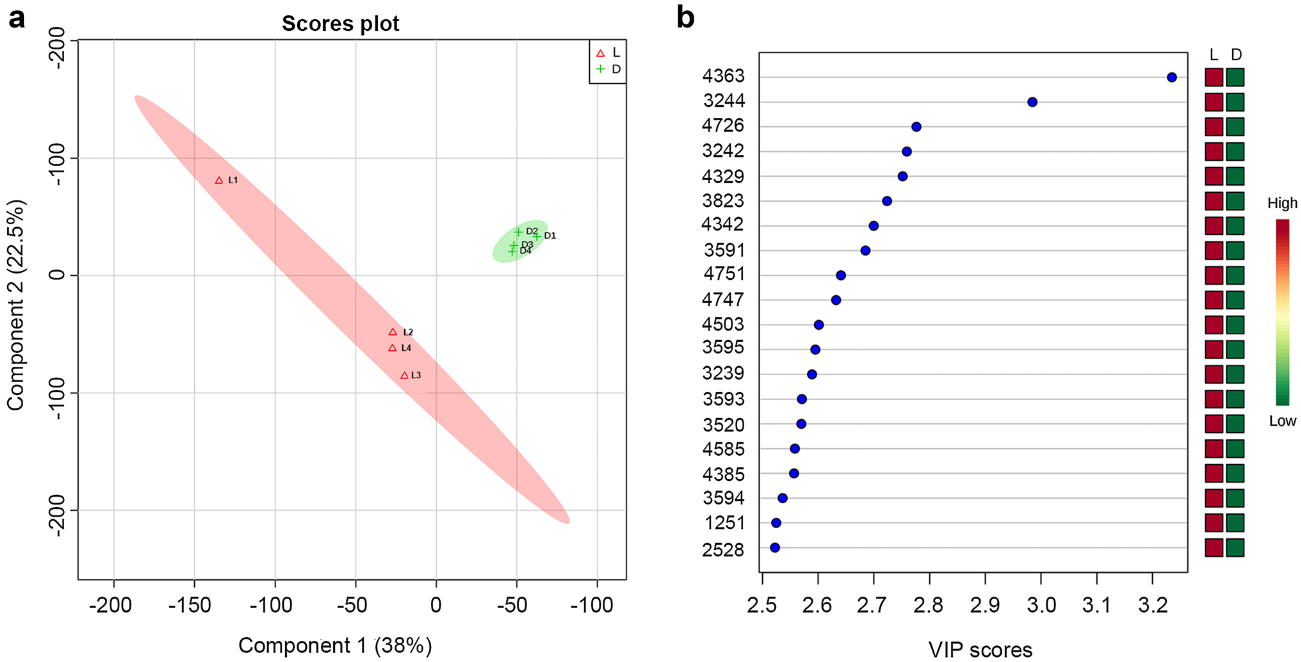


**Fig. 3** PLS-DA score plot and VIP scores of the responses under blue light (L) and dark treatment (D). **a** Samples with different treatments were indicated by different colors. X-axis, Y-axis were labeled with the first principal component and the second principal component explaining 38% and 22.53% of the total variation respectively. **b** VIP loading scores for highest 20 peaks. (Color figure online)

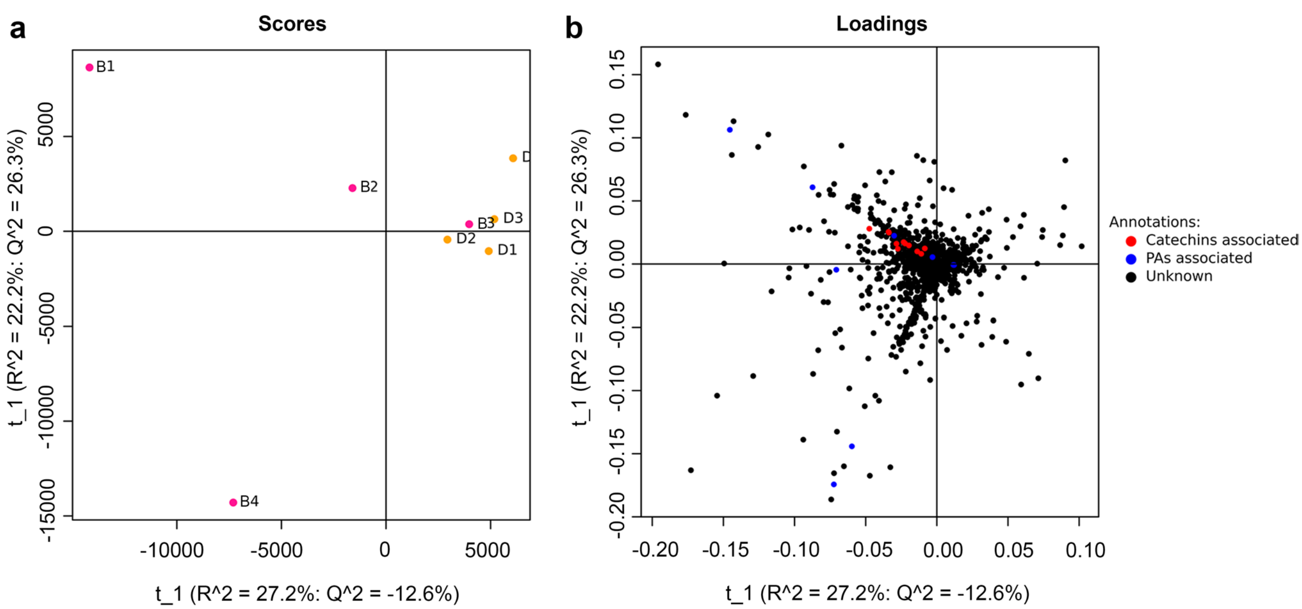


**Fig. 4** PCA of blue light and dark cacao cell suspensions. Comparison of 1097 MS1 features from RP-HPLC–MS/MS data. **a** Scores for eight suspensions under blue (L1–4) and dark (D1–4) and **b** loadings with annotations. Catechin monomers and metabolites associated are represented by red dots, oligomeric forms (dimer up to tetramers) are indicated in blue dots and unknown black dots were not characterized. (Color figure online)

In addition, we aimed to identify blue-light induced metabolites clustering with phenolic compounds. For this, we used the web application MetFamily, which integrates MS1 abundances with MS/MS spectra allowing the characterization of metabolite families. The PCA analysis generated by MetFamily for 1097 MS1 features also detected a clear separation of the treatments. However, the higher variation within L samples affected the $R^2$ and $Q^2$ (Fig. 4a).

The PCA loadings plot showed that MS1 features were more abundant for the blue light. Additionally, structural annotations for metabolites clustering with monomers (red spots) and oligomeric forms (from dimers up to tetramers, blue spots) were also more highly concentrated in L compared to D (Fig. 4b). Interestingly, oligomeric forms showed a higher discriminant power compared to monomers under the L treatment. The structural annotations were identified in signal-clusters through a HCA. The analysis was strengthened using an average of 2000 MS1 abundances and a LFC ≥ 2 to detect abundant light-specific metabolites. The HCA allowed us to identify 349 MS1 features using the corresponding MS/MS spectra obtained from the cacao cell suspensions. The resulting dendrogram indicated a segregation into two main clades with internal spectral similarity (Fig. 5). The first signal-cluster had 92 MS1 features supported by prevalent fragment ions 78.9587 Da (theoretical mass for phosphate, $[PO_3]^-$) and 152.9953 (theoretical mass for glycerolphosphate with a loss of water $[C_3H_6O_5P]^-$). For the second clade, 257 MS1 features were identified with numerous sub-clades usually showing different prevalent fragments per cluster. We detected metabolites clustering with phenolic compounds (colored cluster Fig. 6) with a cluster-discriminant power of 75%. In this sub-clade, seven prevalent ions were present (125.0242, 137.0240, 151.0403, 161.0236, 243.0294, 287.0556, 289.0712). Additionally, this phenolic cluster showed two sub-clusters related with oligomeric forms ("PAs-associated", blue cluster) and metabolites related with monomers ("Catechins-associated", red cluster). In the PAs cluster, we found five peaks at *m/z* of 697.123, 1167.237, 864.193, 864.195 and 864.702. In the catechins cluster, we found a higher number of new peaks (9) at *m/z* of 1137.273 (2 isomers), 849.209, 849.206, 561.140, 273.076, 435.127, 435.093 and 605.165. Altogether, these peaks represented new metabolites clustering with cacao phenolic compounds and appeared to be up-expressed under blue light.

Finally, a heatmap was generated based on the top 50 peaks and subjected to cluster analysis to provide an overview of all samples, highlighting holistic differences in the complex metabolic data (Fig. 6). The analysis provided clear evidence of the effect of blue light in the up-expression of most metabolites, coinciding with the observed pattern shown for the phenolic compounds. The same exercise was done for the first 100 peaks detected, showing the same pattern (Supplemental Fig. S6). Furthermore, chromatograms showed that under light treatment more peaks showed higher abundances compared to dark (Supplemental Fig. S7).

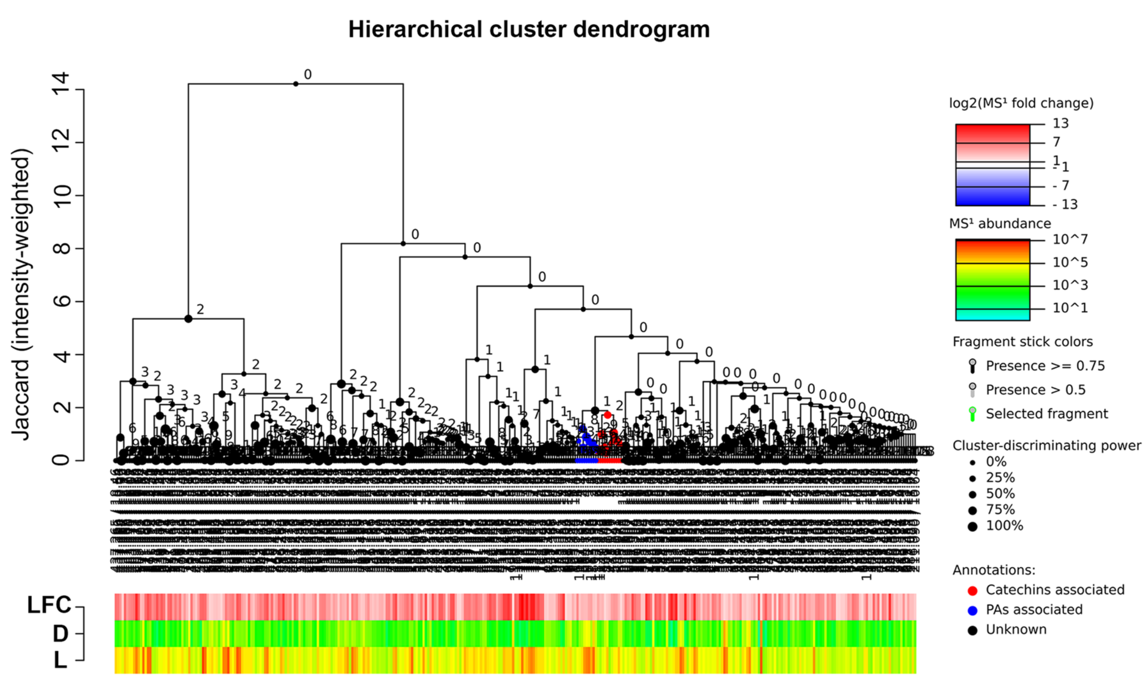


**Fig. 5** Hierarchical cluster analysis of 349 light-specific MS1 features using the corresponding MS/MS spectra obtained from cacao cell suspensions. For comparison of the groups blue light versus darkness, the set of 1097 MS1 features was filtered using an MS1 abundance threshold of 2000 counts and a LFC change of 2. The heatmap below depicts the LFC and the absolute MS1 abundance in blue light (L) and darkness (D) respectively. The 349 filtered MS1 features clearly segregated into two main signal-clusters which in turn further segregated into signal-clusters with different levels of similarity between MS/MS spectra. The analysis was focused in the fifth sub-clade of the second clade with 20 features. It was further split into 2 sub-clusters: PAs associated including dimer up to tetramers with 9 features (blue sub-clade) and catechins associated including monomers (red sub-clade) with 11 features. The rest of the clusters remained uncharacterized in this study. (Color figure online)

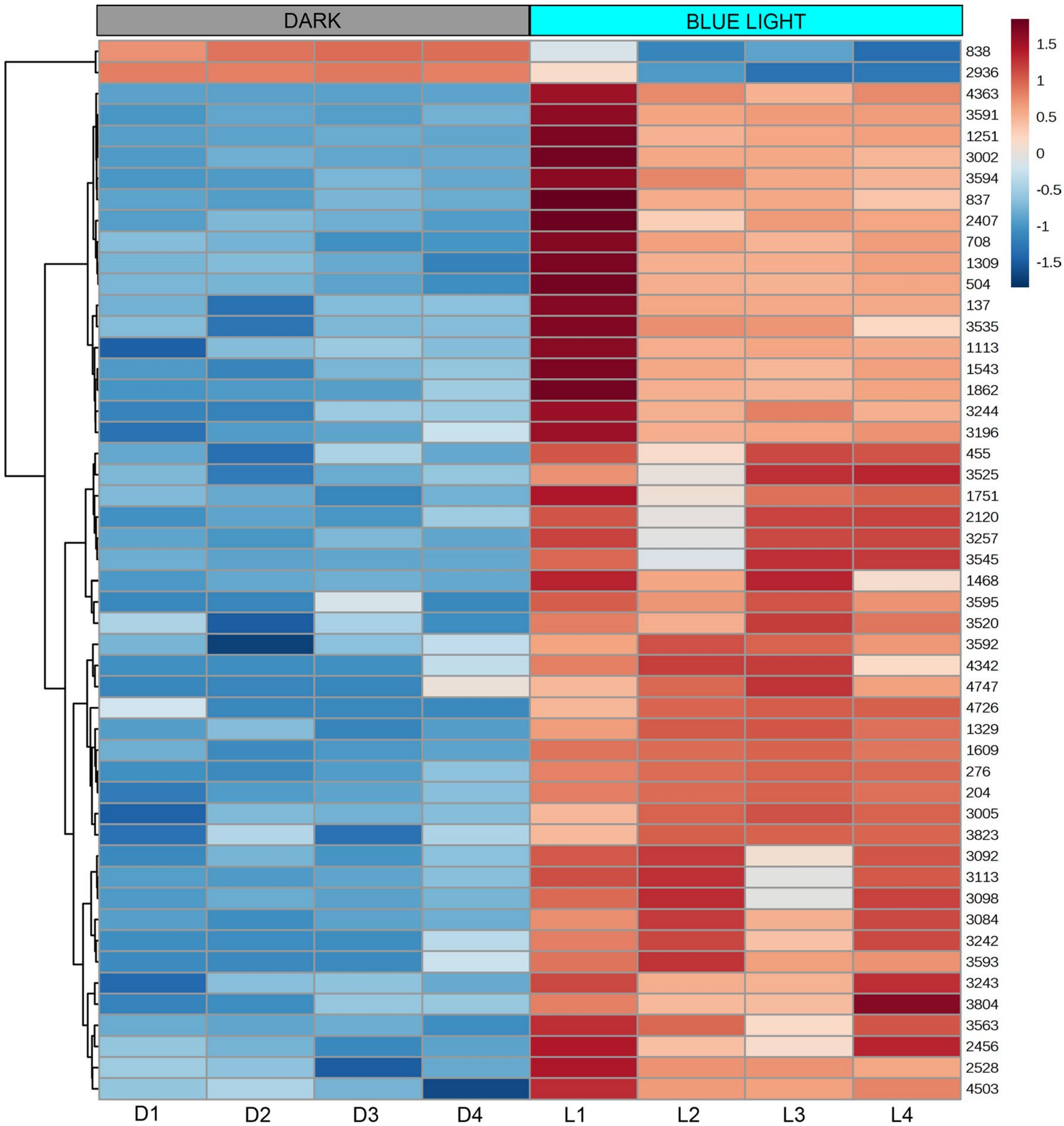


**Fig. 6** Heatmap of 50 highest peak areas detected in the cacao cell suspensions showing differences between dark and blue light. (Color figure online)

## Discussion

### Cacao cell suspensions show a similar phenolic profile to cacao beans and chocolate products

In this study, we provide the first report on the metabolic profiling of cacao cell suspensions. The phenolic compounds found in our cell suspensions were comparable to those previously reported for cacao beans and chocolate products for most of the metabolites (Wollgast and Anklam 2000; Natsume et al. 2000; Sánchez-Rabaneda et al. 2003; Stark and Hofmann 2006; Tomas-Barberán et al. 2007; Calderón et al. 2009; Elwers et al. 2009; Magi et al. 2012; Pereira-Caro et al. 2013; Wang et al. 2016; Patras et al. 2014; Ali et al. 2015; Barnaba et al. 2017; D'Souza et al. 2017). The structures of these metabolites were described according to the registered mass spectra and literature data for cacao. Protocatechuic acid and citric acid have been reported for cacao beans and are produced in the phenylpropanoid and citrate/Krebs pathways, respectively, whereby both routes provide precursors and carbon sources for the flavonoid pathway (Kakkar and Bais 2014; Barnaba et al. 2017; D'Souza et al. 2017). Interestingly, in our cell cultures under blue light/dark treatments, we did not detect alkaloids, anthocyanins or hydroxycinnamic acids commonly found in seeds (Oracz et al. 2015). We hypothesize that the structural enzymes involved in the biosynthesis of alkaloids were not active in this system. This was supported by our transcriptome data, in which we could not find transcripts for the caffeine synthase gene (Tc10v2_g002030) responsible for alkaloid production. Likewise, the expression of seven UFGT transcripts for anthocyanins was very low. On the other hand, the reason for the absence of hydroxycinnamic acids in cell cultures is unknown and needs to be checked with further experimentation.

The cacao cell suspensions yielded four *N*-phenylpropenoyl-L-amino acids (PPAA) as derivatives of aspartic acid and tyrosine, as previously reported for fresh cacao beans and chocolate-derived products (Stark and Hofmann 2005; Patras et al. 2014; D'Souza et al. 2017). The two PPAA from aspartic acid (+)-*N*-[4′-hydroxy-(E)-cinnamoyl]-L-aspartic acid and (+)-*N*-[4′-hydroxy-3′-methoxy-(E)-cinnamoyl]-L-aspartic acid have also been reported in cell suspension cultures of *Arabidopsis thaliana* and *Beta vulgaris*, respectively (Bokern et al. 1991; Mock et al. 1993). Interestingly, tyrosine-derivatived PPAA, clovamide and deoxyclovamide, have been reported not only in cacao seeds and cocoa liquor but also in cacao somatic embryos (Alemanno et al. 2003; Pereira-Caro et al. 2013). Three sub-groups of flavonoids, flavanone, flavonols, and flavones detected in the cacao cell suspensions have been also reported for cacao beans and chocolate products (Sánchez-Rabaneda et al. 2003; Barnaba et al. 2017). As for flavononols, dihydroquercetin has been reported in an in silico model for flavonoid biosynthesis in cacao (Rodriguez and Infante 2011). It is not surprising to detect flavononols in this model, because these are precursors of flavonols and PAs. Also, the conversion of dihydroquercetin into catechin is known to occur in vitro (Dixon et al. 2005).

Monomers of catechin and epicatechin and oligomeric forms of flavan-3ols (PAs) are the most reported compounds in cacao. These have been identified in cacao fresh beans, fermented/dried beans, and chocolate products (Elwers et al. 2009; Magi et al. 2012; Pereira-Caro et al. 2013; Ali et al. 2015; Wang et al. 2016; D'Souza et al. 2017). In this research, we found both monomers and the oligomeric forms up to tetramers of the A and B types. Catechin and epicatechin yielded the same deprotoned form in negative mode (289) and fragmentation pattern 245 (likely from the loss of –$CH_2$–CHOH group from a catechin unit), as previously reported in cacao beans and chocolate products (Pereira-Caro et al. 2013; Cádiz-Gurrea et al. 2014). Additionally, we detected a peak *m*/*z* of 203, corresponding to the ion $[M{-}H{-}CO_2{-}C_2H_2O]^-$ as cleavage of the A-ring of flavan-3-ol (Rockenbach et al. 2012; Zhao et al. 2013).

In relation to the oligomeric forms, most of the PAs showed similar ion products within each category (dimer, trimers and tetramers) corresponding to the A or B types. B-type PAs are more often reported for cacao and chocolate products, specially dimers (Hammerstone et al. 2000). The dimers consist of an extension unit and a terminal unit. The heterocyclic ring of the flavan-3-ol joins fragments through retro-Diels–Alder (RDA) and heterocyclic ring fission mechanisms (Rockenbach et al. 2012). The 577 *m*/*z* $[M{-}H]^-$ ion corresponds to the different PAs dimers. Here, its MS2 fragmentation pattern showed ion products like 289 and 407. The *m*/*z* 289 corresponds to the monomeric forms, catechin and epicatechin, with a heterolytic cleavage of the interflavan linkage in the second flavanyl unit, while *m*/*z* 407 $[M{-}H{-}152{-}18]^-$ results from RDA fragmentation and subsequent elimination of water (Sánchez-Rabaneda et al. 2003; Pereira-Caro et al. 2013; Gu et al. 2013). For trimers (865, $[M{-}H]^-$), besides the three ions previously mentioned, we also found the fragment 287 (heterolytic cleavage of the interflavan linkage in the first flavanyl unit) and 695 *m*/*z*, which correspond to a loss of 170 Da from the precursor ion 865, due to the RDA fission with an additional loss of water (Weinert et al. 2012; Rockenbach et al. 2012). For tetramers (1153, $[M{-}H]^-$), we found all the previous fragments plus a *m*/*z* of 983, which corresponds to a loss of 170 Da (− 152 Da RDA fragmentation of the B-ring of flavan-3-ols and -18 OH) (Zhang et al. 2017). Further, we also detected a 125 *m*/*z* ion in all the oligomers, representing the intact A-ring of the catechin structure (Gu et al. 2013). This ion is also commonly found in green tea catechins extracts (Miketova et al. 2000). Altogether, the results indicate that the phenolic compounds are distributed throughout the flavonoid pathway; thus, the complete flavonoid biosynthetic pathway was recovered in vitro as occurs *in planta*.

## The metabolomic analysis reveals blue light-specific compounds in cacao cell suspensions

Plants contain light-signaling systems and undergo metabolomic perturbation and reprogramming under light stress in order to adapt to environmental changes (Tohge et al. 2018). In this research, cacao cell suspensions were exposed to white-blue LED light (L) and darkness (D), whereby light-specific responses were detected at the level of phytochemical production the end of the experiment under blue light and darkness. A review by Taulavuori et al. (2017) determined that blue LED light (450–490 nm) boosts the phytochemical production in several plant systems. Interestingly, we found that parallel to the higher flavonoids abundance found in L, a high lipid abundance was present in the chemical classification for the top 50 features (Fig. S1). This result could be associated to a higher flow of flavonoids transport traffic from Golgi to vacuole, thus requesting de novo biosynthesis of membrane lipids. Secondary metabolite biosynthesis in land plants is a complex process in which at least two main factors are involved. First, this is a species-specific process, meaning that specific compounds are naturally produced in particular plant species (like catechins and PAs as major phenolics in cacao). Second, environmental factors, in this case light quality, are factors that strongly modify the level of phytochemical production (Taulavuori et al. 2018). Short wavelengths, such as blue light, have been reported to stimulate flavonoid production in vivo and in vitro, mediating a high transcriptional regulation of structural genes and transcription factors as well as microRNAs (Li et al. 2018; Kumar and Dangi 2016). Recently, a new mechanism has been proposed by which the activation of cryptochrome 1 (*CRY1*) through blue light results in a direct enzymatic conversion of molecular oxygen ($O_2$) to reactive oxygen species (ROS) and hydrogen peroxide ($H_2O_2$) in vitro, suggesting that both mechanisms may indeed be of physiological relevance for the cells (Consentino et al. 2015; El-Esawi et al. 2017). Blue wavelength is just above the UV light spectra and it was selected herein as it could boost secondary metabolism without damaging the main cellular structures (DNA, proteins, and lipids), as occurs with UV light.

The correlation between metabolomic with transcriptomic data (Gallego et al. 2018) revealed most of the phenolic compounds showed a positive correlation with the expression levels of the *FLS*, *DFR* and *UFGT* genes. This positive correlation can be explained by the important functions played by these enzymes inside the flavonoid pathway. Both *FLS* and *DFR* represent branching points from the main pathway for anthocyanin and PAs formation respectively, two main metabolic families enriched in cacao. *UFGT* is the enzyme responsible for the glycosidation of flavonoids, thus it makes sense that the high diversity of flavonoids forms found in cacao are associated with a high activity of *UFGT*. On the other hand, the highly positive correlation between abundant polyphenol metabolites is expected as they are co-regulated and belong to the same pathway (Schlotterbeck et al. 2006). Despite the fact that we did not observe significant differences for total polyphenol content (TPC) between blue light and darkness at the last time point in Gallego et al. (2018), the soluble PAs fraction used for the metabolomic analysis showed significant differences between treatments. The differences can be attributed to the specificity of the methods. Folin–Ciocalteu reagent, used in TPC, react with all molecules with OH functional groups, thus is less specific to detect PAs, the major component in cacao. Instead the DMACA reagent used for the soluble PAs extraction is specific to PAs because it binds to meta-oriented dihydroxy- or trihydroxy-substituted benzene rings in the flavonoids (Abeynayake et al. 2011), thus the specificity of the detection is increased, allowing to obtain significant differences between treatments.

By clustering of phenolic compounds using a HCA, B-type tetramers (1167, 697) and B-type hexamers (864.1 and 864.7) were grouped with the PAs group. Interestingly, the tetramers found here were suggested as consisting of one (epi)gallocatechin and three (epi)catechin units (Lin et al. 2016). Furthermore, we annotated a previously unreported light-specific metabolite family for cacao. Masses of this unreported family were associated with propelargonidin derivatives (1137, 849, 561) and phloretin derivatives (435, 273), which clustered with the catechins. Propelargonidin derivatives corresponded to pentameric propelargonidin (Han et al. 2013), whereas phloretin derivatives corresponded with the aglycone (435) and its hexoside (273) (Franceschi et al. 2012). A close relationship between propelargonidin and PAs has been reported for *Uncaria tomentosa* (Rubiaceae), together with a high antioxidant activity described particularly for propelargonidin (Navarro et al. 2017). On the other hand, phloretin has been reported to show antioxidant activity in apple, which could be attributed in part to radical stabilization by 2,6-dihydroxyacetophenone via tautomerization (Lou and Ho 2016). Altogether, these findings suggest that cacao cell suspensions have the enzymatic machinery to produce interesting metabolites besides the well-known catechins, with strong reported antioxidant activities that can be further explored for food or medical purposes.

In the above-annotated light-specific metabolites, we detected two trends; one associated with a higher abundance of hexoside forms and the other is associated with a higher abundance of oligomeric forms. We found five different flavonoid-hexosides overexpressed under blue light. In plants, glycosylation plays important functions in flavonoid production, since it increases solubility in the aqueous cellular milieu, protects the reactive OH-groups from auto-oxidation, and allows the transport of flavonoids from the endoplasmic reticulum to various cellular compartments, as well as their secretion to the apoplast (Zhao and Dixon 2009; Agati and Tattini 2010). Under stress conditions, the diversity of secondary metabolites has been associated with a critical role in the plant defense response (Khan and Mohammad 2011). Thus, glycosylation (addition of sugar moieties to aglycones) is considered the main mechanism for generating structural diversity in secondary metabolites (Rai et al. 2016). Furthermore, glycosylation has been implicated in a radical scavenging activity that improves tolerance toward oxidative and drought stress in *Arabidopsis* (Nakabayashi et al. 2014). Additionally, we found a higher ratio of flavone-hexosides (two luteolin and one apigenin hexoside). This could be explained by the fact that dihydroxy B-ring-substituted forms such as quercetin 3-*O* and luteolin 7-*O*-glycosides are considered light-responsive flavonoids, more than the monohydroxy B-ring-substituted counterparts, like apigenin 7-*O* and kaempferol 3-*O*-glycosides (Brunetti et al. 2013), based on the number of hydroxy groups. Furthermore, and consistent with our results, the ratios of luteolin to apigenin derivatives have been reported to steeply increase in response to UV-B, UV-B + UV-A, or PAR (photosynthetic active radiation, over the 400–700 nm waveband) irradiance (Brunetti et al. 2013). Overall, these results indicate that flavone-hexosides increase in response to the oxidative stress caused by light; however, additional chemical studies should be done to confirm these results.

Oligomeric forms like dimers, trimers and tetramers were differentially expressed under blue light and caused the separation of the treatments in the discriminant analysis, even more than monomers. This is evident in the PCA metabolites loading plot (Fig. 4). However, the effect of light conditions on the polymerization degree has not been explored. Hernández et al. (2009) reported that oxidative processes are necessary to form the PA building blocks, and PA polymerization has been proposed to be an antioxidative mechanism. In wine, for example, catechins are polymerized under oxidative conditions, thereby building polymers (Arapitsas et al. 2012). Several authors have reported that the antioxidant activity of PAs is positively related to their degree of polymerization (i.e., polymer > oligomer > monomer) (Plumb et al. 1998; Spranger et al. 2008; De Sá et al. 2014). This has been hypothesized to be due to the structure of these compounds, which contain a higher number of *ortho*-dihydroxy structures in the B ring that have the highest scavenging activities, as occurs in oligomeric PAs. This hypothesis is supported by the fact that the antioxidant activity of PAs is dictated partially by the oligomer chain length (Gris et al. 2011). The polymerization degree was studied in two *Mesembryanthemum edule* plants coming from contrasting climatic regions (Djerba and Monastir). Plants grown in Djerba, subjected to more difficult environmental conditions as compared to those from Monastir, showed a higher content of polymeric forms. The authors concluded that the gradients in concentrations within-species may reflect different requirements for dealing with abiotic stresses (Falleh et al. 2012). Thus, it is likely that a higher prevalence of oligomeric forms under blue light allow cacao cells to overcome the effects of oxidative stress. In this sense, the occurrence and polymerization degree could act as a marker of environmental adaptation (Falleh et al. 2012). Altogether, our results indicated that, under light conditions, cacao cells respond by increasing the phytochemical machinery to overcome oxidative stress induced by light-specific mechanisms. This results in increased abundances of flavonoid-hexosides and oligomeric forms. Although cacao cell suspensions were exposed from day 7 to the end of the experiment under blue light, it is likely that the early treatment with white light also have contributed to the trends observed in glycosylation and degree of polymerization.

## Potential of cacao cell cultures in food biotechnology

In view of the increasing interest in the use of antioxidants in human health, the choice of an antioxidant for a given application will depend on both the availability of the resource and the chemical profile. It is well-known that pre-industrial and industrial processes in chocolate manufacturing negatively affect the content and antioxidant activity of cacao beans (Di Mattia et al. 2017). This results in reduced flavonoid content in the final chocolate products. Additionally, several studies using cacao phenolic extracts have highlighted the benefits of cocoa intake for human health, particularly those related with cardiovascular events (Latif 2013; De Araujo et al. 2013; Bitzer et al. 2015; Magrone et al. 2017). Moreover, those benefits resulted in a health claim established by the EFSA (2012). Accordingly, the characterization and quantification of the polyphenol composition are among the first steps to evaluate the putative contribution of chocolate to human health (Wollgast and Anklam 2000). In this research, we detected that blue light had a positive effect on phytochemical enrichment; consequently, the cacao cultures proved to be a matrix rich in flavonoids. Although commercially successful plant cell factories are scarce, primarily due to the associated costs, potential legal liabilities, and the limited understanding in the control of secondary metabolism in vitro, some successful cases are available in the market (Ramirez-Estrada et al. 2016). In this sense, flavonoid-rich matrices such as cacao cultures could be used in the design of functional foods providing health benefits to the consumers.

**Acknowledgements** We thank Ruta N (Project Number C007), Compañía Nacional de Chocolates and Universidad de Antioquia for funding. We also thank the Huck Institutes of the Life Sciences, Genomics and Bioinformatics Core Facilities at Penn State University (PSU), especially to Benjamin Knollenberg for providing the internal standard and gave useful comments to the research. Finally, we thank the Members of the Biotecnología and Evo-Devo en Plantas Groups, the Guiltinan-Maximova Lab at PSU, Anderson Lab, Dr. Phill Smith and the Metabolomics Facility at PSU, and Members of Rausch Lab in Heidelberg University (Germany) for providing laboratory and office space for the research.

**Author contributions** AMG conceived the project, performed chemical extractions and made the draft of the manuscript. LFR conceived the project and assisted with chemical analysis. HAR, AMG and EG analyzed the metabolomic data. AIU assisted in the cell culture establishment, experimental standardization and writing. LA established LEDs platforms for in vitro cultures and revised the manuscript. CM wrote the section of food biotechnology. MJG, SNM, designed the experiments, assisted in the data analysis and revised the manuscript. MZ performed the classification of chemical classes, extracted the chromatograms, revised the results related with multivariate analyses. NPM supervised the research, wrote and discussed the results with AMG. All authors revised the article.

**Data availability** All metabolomics data are available in MetaboLights (http://www.ebi.ac.uk/metabolights/).

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References


Abeynayake SW, Panter S, Mouradov A, Spangenberg G (2011) A high-resolution method for the localization of proanthocyanidins in plant tissues. Plant Methods 7:13

Afoakwa EO, Quao J, Takrama J, Budu AS, Saalia FK (2013) Chemical composition and physical quality characteristics of ghanaian cocoa beans as affected by pulp pre-conditioning and fermentation. J Food Sci Technol 50:1097–1105

Agati G, Tattini M (2010) Multiple functional roles of flavonoids in photoprotection. N Phytol 186:786–793

Alemanno L, Ramos T, Gargadenec A, Andary C, Ferriere N (2003) Localization and identification of phenolic compounds in *Theobroma cacao* L. somatic embryogenesis. Ann Bot 92:613–623

Ali F, Ranneh Y, Ismail A, Esa NM (2015) Identification of phenolic compounds in polyphenols-rich extract of Malaysian cocoa powder using the HPLC–UV–ESI–MS/MS and probing their antioxidant properties. J Food Sci Technol 52:2103–2111

Anga J (2014) The world cocoa economy: current status, challenges and prospects. In: Multi-year expert meeting on commodities and development. ICCO

Arapitsas P, Scholz M, Vrhovsek U, Di Blasi S, Biondi Bartolini A, Masuero D, Perenzoni D, Rigo A, Mattivi F (2012) A metabolomic approach to the study of wine micro-oxygenation. PLoS ONE. https://doi.org/10.1371/journal.pone.0037783

Barnaba C, Nardin T, Pierotti A, Malacarne M, Larcher R (2017) Targeted and untargeted characterisation of free and glycosylated simple phenols in cocoa beans using high resolution-tandem mass spectrometry (Q-Orbitrap). J Chromatogr A 1480:41–49

Bitzer Z, Glisan S, Dorenkott M, Goodrich K, Ye Y, Keefe S, Lamber J, Nelison A (2015) Cocoa procyanidins with different degrees of polymerization possess distinct activities in models of colonic inflammation. J Nutr Biochem 179:95–105

Bokern M, Heuer S, Wray V, Witte L, Macek T, Vanek T, Strack D (1991) Ferulic acid conjugates and betacyanins from cell cultures of *Beta vulgaris*. Phytochemistry 30:3261–3265

Borges LL, Alves SF, Sampaio BL, Conceição EC, Bara MTF, Paula JR (2013) Environmental factors affecting the concentration of phenolic compounds in *Myrcia tomentosa* leaves. Braz J Pharmacogn 23:230–238

Brunetti C, Di Ferdinando M, Fini A, Pollastri S, Tattini M (2013) Flavonoids as antioxidants and developmental regulators: relative significance in plants and humans. Int J Mol Sci 14:3540–3555

Bucheli P, Rousseau G, Alvarez M, Laloi M, McCarthy J (2001) Developmental variation of sugars, carboxylic acids, purine alkaloids, fatty acids, and endoproteinase activity during maturation of *Theobroma cacao* L. seeds. J Agric Food Chem 49:5046–5051

Cádiz-Gurrea ML, Lozano-Sanchez J, Contreras-Gámez M, Legeai-Mallet L, Fernández-Arroyo S, Segura-Carretero A (2014) Isolation, comprehensive characterization and antioxidant activities of *Theobroma cacao* extract. J Funct Foods 10:485–498

Calderón AI, Wright BJ, Hurst WJ, Van Breemen RB (2009) Screening antioxidants using LC–MS: case study with cocoa. J Agric Food Chem 57:5693–5699

Chong J, Soufan O, Li C, Caraus I, Li S, Bourque G, Wishart DS, Xia J (2018) MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. Nucleic Acids Res 46:1–9

Consentino L, Lambert S, Martino C, Jourdan N, Bouchet P, Witczak J, Castello P, El-Esawi M, Corbineau F, Harlingue A, Ahmad M, Ahmad M (2015) Blue-light dependent reactive oxygen species formation by *Arabidopsis* cryptochrome may define a novel evolutionarily conserved signaling mechanism. N Phytol 206:1450–1462

Counet C, Ouwerx C, Rosoux D, Collin S (2004) Relationship between procyanidin and flavor contents of cocoa liquors from different origins. J Agric Food Chem 52:6243–6249

D'Souza RN, Grimbs S, Behrends B, Bernaert H, Ullrich MS, Kuhnert N (2017) Origin-based polyphenolic fingerprinting of theobroma cacao in unfermented and fermented beans. Food Res Int 99:550–559

De Araujo QR, Gattward JN, Almoosawi S, das Silva M, Dantas PA, De Araujo Júnior QR (2013) Cocoa and human health: from head to foot—a review. Crit Rev Food Sci Nutr 56:1–12

De Sá M, Justino V, Spranger MI, Zhao YQ, Han L, Sun BS (2014) Extraction yields and anti-oxidant activity of proanthocyanidins from different parts of grape pomace: effect of mechanical treatments. Phytochem Anal 25:134–140

Di Mattia CD, Sacchetti G, Mastrocola D, Serafini M (2017) From cocoa to chocolate: the impact of processing on in vitro antioxidant activity and the effects of chocolate on antioxidant markers in vivo. Front Immunol 8:1–7

Dixon R, Xie D, Sharma S (2005) Proanthocyanidins—a final frontier in flavonoid research? N Phytol 165:9–28

EFSA (2012) Scientific opinion on the substantiation of a health claim related to cocoa flavanols and maintenance of normal endothelium-dependent vasodilation pursuant to Article 13(5) of Regulation (EC) No 1924/2006 1. EFSA J 10:1–21

El-Esawi M, Arthaut LD, Jourdan N, D'Harlingue A, Link J, Martino CF, Ahmad M (2017) Blue-light induced biosynthesis of ROS contributes to the signaling mechanism of *Arabidopsis* cryptochrome. Sci Rep 7:1–9

Elwers S, Zambrano A, Rohsius C, Lieberei R (2009) Differences between the content of phenolic compounds in criollo, forastero and trinitario cocoa seed (*Theobroma cacao* L.). Eur Food Res Technol 229:937–948

Falleh H, Ksouri R, Boulaaba M, Guyot S, Abdelly C, Magné C (2012) Phenolic nature, occurrence and polymerization degree as marker of environmental adaptation in the edible halophyte *Mesembryanthemum edule*. S Afr J Bot 79:117–124

Franceschi P, Dong Y, Strupat K, Vrhovsek U, Mattivi F (2012) Combining intensity correlation analysis and MALDI imaging to study the distribution of flavonols and dihydrochalcones in golden delicious apples. J Exp Bot 63:1123–1133

Fu B, Ji X, Zhao M, He F, Wang X, Wang Y, Liu P, Niu L (2016) The influence of light quality on the accumulation of flavonoids in tobacco (*Nicotiana tabacum* L.) leaves. J Photochem Photobiol B 162:544–549

Gallego A, Henao Ramírez AM, Urrea Trujillo AI, Atehortúa Garcés L (2016) Polyphenols distribution and reserve substances analysis in cocoa somatic embryogenesis. Acta Biol Colomb 21:335–345

Gallego AM, Rojas LF, Parra O, Rodriguez HA, Rivas JCM, Urrea AI, Atehortúa L, Fister AS, Guiltinan MJ, Maximova SN (2018) Transcriptomic analyses of cacao cell suspensions in light and dark provide target genes for controlled flavonoid production. Sci Rep 8:1–14

Ghasemzadeh A, Jaafar HZE, Rahmat A, Wahab PEM, Abd Halim MR (2010) Effect of different light intensities on total phenolics and flavonoids synthesis and anti-oxidant activities in young ginger varieties (*Zingiber officinale* Roscoe). Int J Mol Sci 11:3885–3897

Gris EF, Mattivi F, Ferreira EA, Vrhovsek U, Pedrosa RC, Bordignon-Luiz MT (2011) Proanthocyanidin profile and antioxidant capacity of Brazilian *Vitis vinifera* red wines. Food Chem 126:213–220

Gu F, Tan L, Wu H, Fang Y, Xu F, Chu Z, Wang Q (2013) Comparison of cocoa beans from China, Indonesia and Papua New Guinea. Foods 2:183–197

Gutiérrez-Salmean G, Meaney E, Ceballos-Reyes G, Villareal F (2015) Chocolate and cardiovascular health. In: Chocolate and health: chemistry, nutrition and therapy. Royal Society of Chemistry, Cambridge, pp. 132–145

Hammerstone JF, Lazarus SA, Mitchell AE, Rucker R, Schmitz HH (1999) Identification of procyanidins in cocoa (*Theobroma cacao*) and chocolate using high-performance liquid chromatography/mass spectrometry. J Agric Food Chem 47:490–496

Hammerstone JF, Lazarus SA, Schmitz HH (2000) Chocolate: modern science investigates an ancient medicine procyanidin content and variation in some commonly consumed foods 1. Am Soc Nutr Sci 130:2086–2092

Han Z-P, Liu R-L, Cui H-Y, Zhang Z-Q (2013) Microwave-assisted extraction and LC/MS analysis of phenolic antioxidants in sweet apricot (*Prunus armeniaca* L.) kernel skins. J Liq Chromatogr Relat Technol. https://doi.org/10.1080/10826076.2012.717057

Hernández I, Alegre L, Van Breusegem F, Munné-Bosch S (2009) How relevant are flavonoids as antioxidants in plants? Trends Plant Sci 14:125–132

Hoppe A (2012) What mRNA abundances can tell us about metabolism. Metabolites 2:614–631

Kakkar S, Bais S (2014) A review on protocatechuic acid and its pharmacological potential. ISRN Pharmacol. https://doi.org/10.1155/2014/952943

Khan M, Mohammad F (2011) Role of secondary metabolites in defense mechanisms of plants. Biol Med 3:232–249

Krähmer A, Engel A, Kadow D, Ali N, Umaharan P, Kroh LW, Schulz H (2015) Fast and neat—determination of biochemical quality parameters in cocoa using near infrared spectroscopy. Food Chem 181:152–159

Kumar A, Dangi V (2016) Electromagnetic spectrum and its impact on human life. Int J All Res Educ Sci Methods (IJARESM) 4:67–72

Latif R (2013) Chocolate/cocoa and human health: a review. Neth J Med 71:63–68

Lee KW, Kim YJ, Lee HJ, Lee CY (2003) Cocoa has more phenolic phytochemicals and a higher antioxidant capacity than teas and red wine. J Agric Food Chem 51:7292–7295

Li H, Lin Y, Chen X, Bai Y, Wang C, Xu X, Wang Y, Lai Z (2018) Effects of blue light on flavonoid accumulation linked to the expression of MiR393, MiR394 and MiR395 in Longan embryogenic calli. PLoS ONE 13:e0191444

Lin GM, Lin HY, Hsu CY, Chang ST (2016) Structural characterization and bioactivity of proanthocyanidins from indigenous cinnamon (*Cinnamomum osmophloeum*). J Sci Food Agric 96:4749–4759

Liu Y, Shi Z, Maximova S, Payne MJ, Guiltinan MJ (2013) Proanthocyanidin synthesis in theobroma cacao: genes encoding anthocyanidin synthase, anthocyanidin reductase, and leucoanthocyanidin reductase. BMC Plant Biol 13:1–19

Lou S, Ho C (2016) Phenolic compounds and biological activities of small-size citrus: Kumquat and calamondin. J Food Drug Anal 25:1–14

Magi E, Bono L, Di Carro M (2012) Characterization of cocoa liquors by GC–MS and LC–MS/MS: focus on alkylpyrazines and flavanols. J Mass Spectrom 47:1191–1197

Magrone T, Russo MA, Jirillo E (2017) Cocoa and dark chocolate polyphenols: from biology to clinical applications. Front Immunol 8:1–13

Manivannan A, Soundararajan P, Halimah N, Ko CH, Jeong BR (2015) Blue LED light enhances growth, phytochemical contents, and antioxidant enzyme activities of *Rehmannia glutinosa* cultured in vitro. Hortic Environ Biotechnol 56:105–113

Miketova P, Schram KH, Whitney J, Li M, Huang R, Kerns E, Valcic S, Timmermann BN, Rourick R, Klohr S (2000) Tandem mass spectrometry studies of green tea catechins. Identification of three minor components in the polyphenolic extract of green tea. J Mass Spectrom 35:860–869

Mock H, Wray V, Beck W, Metzger J, Strack D (1993) Coumaroylaspartate from cell suspension cultures of *Arabidopsis thaliana*. Phytochemistry 34:157–159

Moscatiello R, Baldan B, Navazio L (2013) Plant cell suspension cultures. In: H. Press (ed) Plant mineral nutrients. York, pp. 77–93

Nakabayashi R, Yonekura-Sakakibara K, Urano K, Suzuki M, Yamada Y, Nishizawa T et al (2014) Enhancement of oxidative and drought tolerance in *Arabidopsis* by overaccumulation of antioxidant flavonoids. Plant J 77:367–379

Natsume M, Osakabe N, Yamagishi M, Takizawa T, Nakamura T, Miyatake H, Hatano T, Yoshida T (2000) Analyses of polyphenols in cacao liquor, cocoa, and chocolate by normal-phase and reversed-phase HPLC. Biosci Biotechnol Biochem 64:2581–2587

Navarro M, Zamora W, Quesada S, Azofeifa G, Alvarado D, Monagas M (2017) Fractioning of proanthocyanidins of *Uncaria tomentosa*. Composition and structure–bioactivity relationship. Antioxidants 6:1–13

Niemenak N, Rohsius C, Elwers S, Omokolo Ndoumou D, Lieberei R (2006) Comparative study of different cocoa (*Theobroma cacao* L.) clones in terms of their phenolics and anthocyanins contents. J Food Compos Anal 19:612–619

Oracz J, Zyzelewicz D, Nebesny E (2015) The content of polyphenolic compounds in cocoa beans (*Theobroma cacao* L.), depending on variety, growing region, and processing operations: a review. Crit Rev Food Sci Nutr 55:1176–1192

Pancaningtyas S (2015) Study of the presence and influence of phenolic compounds in callogenesis and somatic embryo development of cocoa (*Theobroma cacao* L.). Pelita Perkeb 31:14–20

Patras MA, Milev BP, Vrancken G, Kuhnert N (2014) Identification of novel cocoa flavonoids from raw fermented cocoa beans by HPLC–$MS^n$. Food Res Int 63:353–359

Pedroso RCN, Branquinho NAA, Hara ACBAM, Costa AC, Silva FG, Pimenta LP, Silva MLA, Cunha WR, Pauletti PM, Januario AH (2017) Impact of light quality on flavonoid production and growth of *Hyptis marrubioides* seedlings cultivated in vitro. Braz J Pharmacogn 27:466–470

Pereira-Caro G, Borges G, Nagai C, Jackson MC, Yokota T, Crozier A, Ashihara H (2013) Profiles of phenolic compounds and purine alkaloids during the development of seeds of *Theobroma cacao* Cv. Trinitario. J Agric Food Chem 61:427–434

Plumb GW, De Pascual-Teresa S, Santos-Buelga C, Cheynier V, Williamson G (1998) Antioxidant properties of catechins and proanthocyanidins: effect of polymerisation, galloylation and glycosylation. Free Radic Res 29:351–358

Rai A, Umashankar S, Rai M, Kiat LB, Bing JA, Swarup S (2016) Coordinate regulation of metabolites glycosylation and stress hormones biosynthesis by TT8 in *Arabidopsis*. Plant Physiol. https://doi.org/10.1104/pp.16.00421

Ramirez-Estrada K, Vidal-Limon H, Hidalgo D, Moyano E, Goleniowski M, Cusidó RM, Palazon J (2016) Elicitation, an effective strategy for the biotechnological production of bioactive high-added value compounds in plant cell factories. Molecules. https://doi.org/10.3390/molecules21020182

Rigaud J, Escribano-Bailon MT, Prieur C, Souquet JM, Cheynier V (1993) Normal-phase high-performance liquid chromatographic separation of procyanidins from cacao beans and grape seeds. J Chromatogr A 654:255–260

Rockenbach II, Jungfer E, Ritter C, Santiago-Schübel B, Thiele B, Fett R, Galensa R (2012) Characterization of flavan-3-ols in seeds of

grape pomace by CE, HPLC–DAD–$MS^n$ and LC–ESI–FTICR–MS. Food Res Int 48:848–855
Rodriguez A, Infante D (2011) Characterization in silico of flavonoids biosynthesis in *Theobroma cacao* L. Netw Biol 1:34–45
Rojas LF, Londoño J, Gallego AM, Herrera AL, Aguilera C, Atehortúa L (2008) Total polyphenols analysis of mature seeds and tissue cultures of some colombian cocoa varieties. Actual Biol 30:117–123
Rojas LF, Gallego A, Gil A, Londono J, Atehortua L (2015) Monitoring accumulation of bioactive compounds in seeds and cell culture of *Theobroma cacao* at different stages of development. In Vitro Cell Dev Biol Plant 51:174–184
Rua AMG, Rojas LF, Trujillo AIU, Zuleta OP, Alvarez CDC, Garces LA (2017) A rational approach for the improvement of biomass production and lipid profile in cacao cell suspensions. Bioprocess Biosyst Eng 40:1479–1492
Sampaio BL, Bara MTF, Ferri PH, da Costa SantosS, de Paula JR (2011) Influence of environmental factors on the concentration of phenolic compounds in leaves of *Lafoensia pacari*. Braz J Pharmacogn 21:1127–1137
Sánchez-Rabaneda F, Jáuregui O, Casals I, Andrés-Lacueva C, Izquierdo-Pulido M, Lamuela-Raventós RM (2003) Liquid chromatographic/electrospray ionization tandem mass spectrometric study of the phenolic composition of cocoa (*Theobroma cacao*). J Mass Spectrom 38:35–42
Schlotterbeck G, Ross A, Dieterle F, Senn H (2006) Metabolic profiling technologies for biomarker discovery in biomedicine and drug development. Pharmacogenomics 7:1055–1075
Spranger I, Sun B, Mateus AM, de Freitas V, Ricardo-da-Silva JM (2008) Chemical characterization and antioxidant activities of oligomeric and polymeric procyanidin fractions from grape seeds. Food Chem 108:519–532
Stark T, Hofmann T (2005) Isolation, structure determination, synthesis, and sensory activity of *N*-phenylpropenoyl-L-amino acids from cocoa (*Theobroma cacao*). J Agric Food Chem 53:5419–5428
Stark T, Hofmann T (2006) Application of a molecular sensory science approach to alkalized cocoa (*Theobroma cacao*): structure determination and sensory activity of noenzymatically C-glycosylated flavan-3-ols. J Agric Food Chem 54:9510–9521
Taulavuori E, Taulavuori K, Holopainen JK, Julkunen-Tiitto R, Acar C, Dincer I (2017) Targeted use of LEDs in improvement of production efficiency through phytochemical enrichment. J Sci Food Agric 97:5059–5064
Taulavuori K, Pyysalo A, Taulavuori E, Julkunen-Tiitto R (2018) Responses of phenolic acid and flavonoid synthesis to blue and blue-violet light depends on plant species. Environ Exp Bot 150:183–187
Tohge T, Perez de Souza L, Fernie AR (2018) On the natural diversity of phenylacylated-flavonoid and their in planta function under conditions of stress. Phytochem Rev 17:279–290
Tomas-Barberán FA, Cienfuegos-Jovellanos E, Marín A, Muguerza B, Gil-Izquierdo A, Cerdá B, Zafrilla P, Morillas J, Mulero J, Ibarra A, Pasamar MA, Ramón D, Espín JC (2007) A new process to develop a cocoa powder with higher flavonoid monomer content and enhanced bioavailability in healthy humans. J Agric Food Chem 55:3926–3935
Treutler H, Tsugawa H, Porzel A, Gorzolka K, Tissier A, Neumann S, Balcke GU (2016) Discovering regulated metabolite families in untargeted metabolomics studies. Anal Chem 88:8082–8090
Tsugawa H, Cajka T, Kind T, Ma Y, Higgins B, Ikeda K, Kanazawa M, VanderGheynst J, Fiehn O, Arita M (2015) MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. Nat Methods 12:523–526
Wang SY, Chen CT, Wang CY (2009) The influence of light and maturity on fruit quality and flavonoid content of red raspberries. Food Chem 112:676–684
Wang L, Nägele T, Doerfler H, Fragner L, Chaturvedi P, Nukarinen E, Bellaire A, Huber W, Weiszmann J, Engelmeier D, Ramsak Z, Gruden K, Weckwerth W (2016) System level analysis of cacao seed ripening reveals a sequential interplay of primary and secondary metabolism leading to polyphenol accumulation and preparation of stress resistance. Plant J Cell Mol Biol 87:318–332
Weinert CH, Wiese S, Rawel HM, Esatbeyoglu T, Winterhalter P, Homann T, Kulling SE (2012) Methylation of catechins and procyanidins by rat and human catechol-*O*-methyltransferase: metabolite profiling and molecular modeling studies. Drug Metab Dispos 40:353–359
Wollgast J, Anklam E (2000) Review on polyphenols in *Theobroma cacao*: changes in composition during the manufacture of chocolate and methodology for identification and quantification. Food Res Int 33:423–447
Zhang L, Tai Y, Wang Y, Meng Q, Yang Y, Zhang S, Yang H, Zhang Z, Li D, Wan X (2017) The proposed biosynthesis of procyanidins by the comparative chemical analysis of five *Camellia* species using LC–MS. Sci Rep 7:1–10
Zhao J, Dixon RA (2009) The "ins" and "Outs" of flavonoid transport. Trends Plant Sci 15:72–80
Zhao H-Y, Fan M-X, Wu X, Wang H-J, Yang J, Si N, Bian B-L (2013) Chemical profiling of the Chinese herb formula Xiao-Cheng-Qi decoction using liquid chromatography coupled with electrospray ionization mass spectrometry. J Chromatogr Sci 51:273–285
Zoratti L, Karppinen K, Luengo Escobar A, Häggman H, Jaakola L (2014) Light-controlled flavonoid biosynthesis in fruits. Front Plant Sci 5:534



## Affiliations

**Adriana M. Gallego[1,10] · Luisa F. Rojas[2] · Héctor A. Rodriguez[1,3] · Conrado Mora[4] · Lucía Atehortúa[1] · Aura I. Urrea[1] · Mark J. Guiltinan[5,6,7] · Siela N. Maximova[6,7] · Emmanuel Gaquerel[8] · Martha Zuluaga[9] · Natalia Pabón-Mora[10]**

[1] Grupo de Biotecnología, Instituto de Biología, Universidad de Antioquia, Medellín, Colombia

[2] Grupo de Biotecnología, Escuela de Microbiología, Universidad de Antioquia, Medellín, Colombia

[3] Corporación para Investigaciones Biológicas and Departamento de Ciencias Agronómicas, Facultad de Ciencias Agrarias, Universidad Nacional de Colombia, UNALMED-CIB, Medellín, Colombia

[4] Compañía Nacional de Chocolates, Medellín, Colombia

[5] Department of Biology, The Pennsylvania State University, University Park, PA, USA

[6] Department of Plant Science, The Pennsylvania State University, University Park, PA, USA

[7] The Huck Institutes of the Life Sciences, The Pennsylvania State University, 422 Life Sciences Building, University Park, PA, USA

[8] Institut de Biologie Moléculaire des Plantes (IBMP) CNRS, UPR 2357, Université de Strasbourg, 67000 Strasbourg, France

[9] Universidad Nacional de Colombia, Manizales, Colombia

[10] Grupo Evo-Devo en Plantas, Instituto de Biología, Universidad de Antioquia, Medellín, Colombia