**Supercritical antisolvent fractionation as a tool for enhancing
antiproliferative activity of mango seed kernel extracts against colon
cancer cells**

Diego Ballesteros-Vivas^{1,2a}, Gerardo Alvarez-Rivera^{2a}, Andrés Felipe García Ocampo¹, Sandra
Johanna Morantes³ Andrea del Pilar Sánchez Camargo¹, Alejandro Cifuentes², Fabián Parada-
Alfonso¹, Elena Ibáñez^{2*}

¹ High Pressure Laboratory, Department of Chemistry, Faculty of Science, Universidad
Nacional de Colombia, Carrera 30 #45-03, Bogotá D.C., 111321, Colombia.

² Laboratory of Foodomics, Institute of Food Science Research, CIAL, CSIC, Nicolas Cabrera
9, 28049 Madrid, Spain.

³ Unit of Basic Oral Investigation (UIBO), School of Dentistry, Universidad El Bosque, Av.
Carrera 9 #131 A-02, Bogotá D.C., 110121, Colombia.

^a These two authors contributed equally to this work.

***Corresponding author:**

Prof. Dr. Elena Ibáñez, Laboratory of Foodomics, Institute of Food Science Research, CIAL
(CSIC), Nicolas Cabrera 9, 28049 Madrid, Spain, e-mail: elena.ibanez@csic.es, Tel.: +34
910017956.

Keywords:

Mangifera indica L.; fruit by-products; Supercritical Antisolvent Fractionation; Pressurized-
liquid extraction; LC-q-TOF-MS; High-resolution mass spectrometry; antiproliferative
activity; HT-29 colon cancer cell line.

ABSTRACT

Mango seed kernel (MSK) has demonstrated to be a valuable source of bioactive phenolic compounds with antiproliferative activity against colon cancer cell lines. In this work, a phenolic MSK extract obtained by sequential pressurized liquid extraction (PLE) was submitted to supercritical antisolvent fractionation (SAF) to attain enriched fractions with improved inhibitory effect on HT-29 colon cancer cells. SAF was optimized using a Box-Behnken experimental design and considering as main factors: pressure, percentage of water in the feed extract and feed/sc-CO₂ flow rate ratio; selected response variable was the inhibitory cell proliferation (%) and the SAF yield (%) measured in both, raffinates and extracts. The highest inhibitory cell proliferation effect ($70.51 \pm 1.14\%$) was obtained in the extract using 50% of water v/v in the feed and 0.0625 feed/sc-CO₂ ratio at 15 MPa. This result represents an improvement in the antiproliferative activity by the SAF process compared to the original PLE extract. A targeted phytochemical profiling by LC-q-TOF-MS/MS followed by a multivariate statistical analysis of the observed bioactivity and the chemical composition evidenced the existence of a remarkable variability among the obtained SAF fractions, mainly due to characteristic compounds (e.g., mangiferin, homomangiferin, galloyl methylgallate, ellagic acid and methyl gallate), which might explain the enhanced antiproliferative activity of the optimal SAF extract.

1. INTRODUCTION

The search for alternative sources of health-promoting compounds has conducted to food by-products conversion into high added-value bio-products [1]. The valorisation strategies for food by-products involve the development of environmentally-friendly and sustainable procedures towards the selective recovery of bioactive compounds. These strategies integrate different approaches, including: i) theoretical models for solubility prediction (i.e. Hansen Solubility Parameters, COSMO-RS method) of target compounds in bio-based and suitable solvents, ii) process intensification and biorefinery based on compressed fluids extraction, iii) *in vitro* bioactive analysis (i.e. anticancer, antioxidant, anti-inflammatory activities), and iv) advanced analytical techniques such as liquid and gas chromatography (LC, GC) coupled to high resolution mass spectrometry (HRMS) for chemical characterization. These multi-analytical approaches play an essential role in the production of new functional food ingredients under a green Foodomics perspective [2,3].

Mango fruit (*Mangifera indica* L.) processing industry generates a large amount of solid by-products. They are mainly composed by peels and kernel seeds, which represent about 40% of the whole fruit [4]. Mango fruit by-products have been recognized as sources of bioactive compounds. In particular, mango seed kernel (MSK) contains several families of health-promoting compounds such as sterols, fatty acids, gallotanins, benzophenones, xanthonenes, flavonoids and phenolic acids [5-8]. The antiproliferative activity of MSK polyphenolics compounds has been well reported and their effects have shown promising anti-cancer potential against different cancer cell line models: breast, liver, leukemia, cervix, prostate, lung and colon [9-12]. The recovery of polyphenols from mango is usually performed employing methanol, ethanol, ethyl acetate, acetone, water and their mixtures as solvents in classical solid-liquid extraction techniques [13]. However, some drawbacks attributed to these techniques are the thermal and chemical degradation of labile compounds, as well as the residual of harmful

solvents affecting the quality and safety of the extract [14]. For these reasons, novel and more environmentally friendly extraction methods such as Pressurized Liquid Extraction (PLE) and Supercritical Fluid Extraction (SFE) have been used to obtain polyphenolic-rich extracts from mango [15,16]. Water, ethanol and carbon dioxide (CO₂) are the preferred solvents used in compressed fluid extraction techniques due to their physicochemical properties and their GRAS (Generally Recognized As Safe) character [17]. In addition, the solvation power and selectivity of these solvents and their mixtures can be tuned by temperature and pressure effects under sub- and supercritical conditions [18]. Indeed, thanks to the unique properties of supercritical carbon dioxide (sc-CO₂), it may act as a solvent, as an antisolvent or as a molecular mobility enhancer, allowing the development of different processes and applications [19].

In recent years, the anti-solvent properties of sc-CO₂ have been exploited for the development of precipitation and fractionation techniques such as the supercritical antisolvent fractionation (SAF). The SAF process consists of the continuous spraying of a polar liquid solution from a natural extract (Feed extract) in a co-current sc-CO₂. The lipophilic character of sc-CO₂ allows the precipitation of more polar components from the feed extract, whereas the less polar components remained dissolved and can be recovered by downstream pressure reduction [20]. In the upstream process, the feed extract can be obtained by any classical (i.e. Soxhlet, maceration) or emergent (i.e. PLE, microwave-assisted) extraction technique, however the solvent must be miscible with sc-CO₂. Ethanol and hydroethanolic mixtures have been the most employed solvents to obtain the feed extract, especially for polyphenols antisolvent fractionation purposes [21]. In this case, most polar polyphenolic compounds are practically insoluble in sc-CO₂, but are sufficiently soluble in a sc-CO₂ + EtOH mixture or in a sc-CO₂ + EtOH-H₂O mixture, leading to separation and/or distribution according to molecular weights and polarity [22]. The SAF technique has been successfully applied to produce nano- and microparticles of antioxidant polyphenolic compounds from mango by-products (peels and

pulp wastes) [23] and mango leaves [24,25]. Furthermore, SAF has also been employed to improve fractions' bioactivity compared to the original feed extract, correlating the compounds-enrichment with the observed bioactivity. In a previous study performed by our research group, Sánchez-Camargo et al. [26], fractionated a *Rosmarinus officinalis* L. extract by SAF process, enhancing the antiproliferative activity against cancer cells due to enrichment of carnosic acid and carnosol in the fractions.

In this context, the present work aims to optimize a SAF process to enhance the antiproliferative activity, against human colon adenocarcinoma cell line HT-29, of a MSK extract previously obtained in our research group, as a result of a two-step sequential PLE process [27]. Since no information beforehand was available on the main compounds responsible of the antiproliferative activity, a phytochemical profiling analysis of the original extract and the fractions obtained by SAF was carried out by liquid chromatography coupled to high resolution mass spectrometry (LC-HRMS) in order to comparatively evaluate the differential enrichment of the different fractions obtained by SAF. Finally, a multivariate analysis was performed in order to understand the potential relationships between the chemical composition of the fractions and the observed antiproliferative activity. This work is a step further in the sustainable valorisation strategy of mango by-products proposed by our research group under a Green Foodomics framework.

2. MATERIAL AND METHODS

2.1 Raw material and reagents

Sugar mango fruits, variety from Colombia, were acquired in the central supply marketplace (www.corabastos.com.co) in Bogotá D.C., Colombia in February 2018. Mango fruits pulping was carried out by mechanical process in order to separate the seed and pulp from the peel. After washing, the seeds were split into coat and kernel. 'Sugar MSK' was dried 48 h at room

temperature and ground using knife mill (Grindomix GM200-Retsch GmbH, Haan, Germany) and dry ice. The MSK powder was vacuum-packed and stored at -20 °C. CO₂ with a maximum purity of 99.8% was obtained from Linde (Colombia). Solvents employed were HPLC-grade. Acetonitrile, chloroform, *n*-heptane and ethanol and were acquired from VWR Chemicals (Barcelona, Spain). Ultrapure water was obtained from a Millipore system (Billerica, MA, USA). For the UHPLC-q-TOF-MS analyses, MS grade ACN and water from LabScan (Dublin, Ireland) were employed. 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT), RPMI-1640 cell culture medium, gentamicin (10 mg/mL), dimethyl sulfoxide (DMSO) and Triton X-100, were purchased from Sigma-Aldrich (Madrid, Spain). Fetal bovine serum (Gibco) and 0.05% and trypsin-EDTA (Gibco) were purchased from Thermo Fisher Scientific (Rockford, IL).

2.2 Pressurized liquid extraction (PLE)

The feed extract was obtained by a PLE process in two sequential steps using a commercial ASE 200 device as described in our previous work [27], where the extraction conditions (temperature and solvent composition) were optimized to maximize mangiferin content, extraction yield, total phenolic content, total flavonoid content and antioxidant activity. Briefly, dried and powdered samples of ‘Sugar MSK’ were mixed with sea sand (1:2 w/w) and loaded into the extraction cell (33 mL stainless steel). In the PLE-first step, the sample was defatted using *n*-heptane as solvent at 100 °C for 20 min. Subsequently, for the PLE-second step, ethanol was used as solvent at 150 °C for 20 min to obtain a polar extract, which was considered as feed extract. After the PLE process, three dilutions of the feed extract were prepared in ultrapure water according to the SAF experimental design. The dilutions were filtered (Whatman N° 1) under vacuum and stored in the dark at -20 °C until their use in SAF process.

2.3 Supercritical Antisolvent Fractionation (SAF)

SAF was performed in a home-build experimental unit (High Pressure Laboratory, Universidad Nacional de Colombia, Colombia) as shown in Figure 1. The main parts of the unit are: (P1) diaphragm type compressor for CO₂ pumping (Nova-Swiss 554-2121, Effretikon, Switzerland), (P2) liquid pump (Hitachi L-6210, Tokyo, Japan) for feed extract delivering, (S1) high pressure separator vessel with (HJ) electric heating jacket, (BPR) back pressure regulator valve (Pressure Tech LF540, Houston, TX, USA), (S2) low pressure separator vessel and (GF) CO₂ gas flowmeter. The SAF experiments were carried out in a continuous mode at 40 °C following this protocol: first, CO₂ was pumped at constant flow rate (13.7 g/min) into the S1 until the desired pressure conditions were achieved, subsequently the aqueous dissolution of the feed extract was pumped at proper flow rate (0.3 to 1.4 g/min) to attain the feed/sc-CO₂ flow rate ratio required according to the experimental design. The sc-CO₂ and feed extract were co-currently mixed in a T-connection device before entering S1. After mixing, the feed extract components with higher affinity for H₂O than for the sc-CO₂ + EtOH phase were concentrated and collected at the bottom of S1; this fraction was named *Raffinate*. The soluble components in the sc-CO₂ + EtOH were precipitated in the S2 after depressurization across the BPR; this fraction was named *Extract*. The SAF time was adjusted (10 to 60 min) depending on the flow rate of feed extract and its fed mass. Ethanol and water were removed from the fractions by rotary evaporation (Rotavapor R-300, Buchi Labortechnik AG, Flawil, Switzerland) or freeze-drying (Freeze dryer 77540, Labconco Corporation, Missouri, USA), respectively. The SAF yield was expressed as percentage from the ratio of the mass dry recovered in the separators and the dry PLE extract fed.

2.4 Experimental design

A Box-Behnken design (BBD) with response surface methodology (RSM) was applied to optimize the SAF conditions to fractionate the feed extract to obtain new extracts with enhanced antiproliferative activity against HT-29 cells. For this purpose, three factors were considered: pressure (A), percentage of water in the feed extract (B) and feed/sc-CO₂ flow rate ratio (C). The coded and real values of the experimental factors are summarized in Table 1. The BBD consisted of 15 randomized runs with three replicates at the central point. The experimental data was fitted with the following second order polynomial equation:

$$Y = \beta_0 + \beta_1 A + \beta_2 B + \beta_3 C + \beta_{1,1} A^2 + \beta_{2,2} B^2 + \beta_{3,3} C^2 + \beta_{1,2} AB + \beta_{1,3} AC + \beta_{2,3} BC \quad \text{Eq. (1)}$$

where Y is the response variable (Inhibitory cell proliferation or SAF yield), and β_0 , β_1 , β_2 , $\beta_{1,1}$, $\beta_{2,2}$, $\beta_{3,3}$, $\beta_{1,2}$, $\beta_{1,3}$ and $\beta_{2,3}$ are regression coefficients of variables for intercept, linear, quadratic, and interaction terms, respectively, and A , B and C are the independent variables.

2.5 Cell line and culture

Human colon adenocarcinoma cell line HT-29 was obtained from the American Type Culture Collection (ATCC, USA). The cells were ground and routinely maintained in RPMI 1640 medium, supplemented with Hepes 25 mM, L-glutamine 2.05 mM, 10% heat-inactivated fetal bovine serum and 25 µg/mL gentamicin in an incubator at 37°C, 5% CO₂, and 95% air humidity. The cultures at 80%–90% confluent were passed by trypsinization using trypsin-EDTA. Cell counts and viability were performed using a standard trypan blue cell counting technique. Cell passes were sub-cultured into new sterile culture flasks for further propagation.

2.6 Antiproliferative activity assay

The dried fractions were initially reconstituted in DMSO at 10 mg/mL and properly diluted in cell culture media to obtain an assay concentration of 28.67 µg/mL. This concentration corresponds to IC₅₀ (concentration of extract that causes 50% inhibition or cell death) of the

feed extract (PLE-extract), which was previously determined [27]. Cells in exponential growth phase (70-80% confluence) were seeded in 96-well plates at a 1.0×10^4 cells/well. The plates were incubated for 24 hours at 37°C to allow cell adhesion. Then, cells were treated with the fractions and the feed extract at the assay concentration and incubated for 72 h. The DMSO (0.1% v/v) safety on cells was also evaluated. Triton X-100 (1.0×10^{-4} to 1.0%) was used as positive control. After the treatment time, the culture medium was replaced by 100 µL of MTT solution (0.25 mg/mL RPMI 1640 medium) and the plates were incubated for 4 h. After this time, medium was discarded and 100 µL of DMSO were added to each well to dissolve the formazan crystals. The absorbance was measured at 570 nm using a microplate reader (Synergy HT, BioTek Instruments, Winooski, VT, USA). The inhibitory cell proliferation was expressed as percentage of dead cells relative to controls. Each experiment was performed as three independent tests with at least three replicates.

2.7 Phenolic profiling of SAF fractions by UHPL-q-TOF-MS/MS analysis

The phytochemical analysis of SAF fractions and feed extract was performed in an Agilent 1290 UHPLC system (Agilent Technologies, Santa Clara, CA, USA) equipped with an Agilent 6540 quadrupole-time-of-flight mass spectrometer (q-TOF MS). The elution was carried out at 30 °C on a Zorbax Eclipse Plus C18 column (2.1×100 mm, 1.8 µm particle diameter, Agilent Technologies, Santa Clara, CA). The mobile phases consisted of H₂O + formic acid 0.01% v/v (A) and acetonitrile + formic acid 0.01% v/v (B). The gradient elution was as follows: 0-30% B in 0-7 min, 30-80% B in 7-9 min, 80-100% B in 9-11 min, 100% B in 11-13 min and 0% B in 13-14 min. The flow rate was 0.5 mL/min and the sample injection volume was 5 µL. To obtain the MS and MS/MS spectra, the mass spectrometer was operated in the ESI negative mode (Agilent Jet Stream, AJS, Santa Clara, CA, USA). The tune method for MS analysis was as follows: capillary voltage, 4000 V; nebulizer pressure, 40 psi; drying gas flow rate, 10

L/min; gas temperature, 350 °C; skimmer voltage, 45 V; fragmentor voltage, 110 V. The MS and Auto MS/MS modes were set to acquire m/z values ranging between 50-1100 and 50-800, respectively, at a scan rate of 5 spectra per second. Agilent Mass Hunter Qualitative and Quantitative analysis software version B.07.00 were used for post-acquisition data processing.

2.7 Statistical analysis

The software STATISTICA 12 (Stat Soft, Inc., Tulsa, OK 74104, USA) was used for experimental design, data analysis and evaluation of the optimal SAF conditions using the desirability function. The suitability of the model was determined by regression coefficient (R^2), p -values, F -test values and the lack-of-fit test obtained from the analysis of variance (ANOVA). Estimate effect analysis and ANOVA were also used for the evaluation of the significance of independent variables' influence and interactions on the extraction of the phenolic compounds in the SAF process. The software Unscrambler® X v10.5 (CAMO Software AS, Oslo, Norway) was used for multivariate analysis. Principal component analysis (PCA) was applied to define both similarities and differences among the SAF fractions obtained from the experimental design, as well as to determine the relationship between antiproliferative activity and phenolic composition of the different extracts.

3. RESULTS AND DISCUSSION

3.1 PLE 'Sugar MSK' extraction: starting point

In a previous work [27], a polyphenols-enriched 'Sugar MSK' extract was obtained in our laboratory by a sequential PLE process, exhibiting inhibitory effect on HT-29 colon cancer cells ($IC_{50} = 28.67 \mu\text{g/mL}$ at 72h). Results showed that the PLE-extract had a selective antiproliferative effect towards HT-29 cells with a selectivity index of 3 respect to CCD-18Co cells (normal human colon fibroblasts). PLE-extract was obtained as a result of a RSM

optimization process, in which the optimal extraction conditions (150 °C and 100% EtOH) allowed to maximize the mangiferin content (13.27 mg/g dry basis), extraction yield (12.1 %), total phenolic content (143.79 mg Gallic acid equivalents/g dry basis), total flavonoid content (1.21 mg Quercetin equivalents/g dry basis) and antioxidant activity (EC_{50} 14.34 μ g/mL; 2.31 mM trolox g^{-1}). The profiling analysis of the PLE extract by UHPLC coupled to q-TOF-MS/MS revealed the presence of typical phenolic acids, flavonoids, characteristic xanthones, as well as a broad family of gallate derivatives and gallotannins.

In the present contribution, the PLE-extract was obtained under identical conditions of temperature and solvent composition, but scaling up the vessel extraction volume three-fold in order to increase the mass extract recovery. Extraction yield and inhibitory effect on cell proliferation were confirmed with the following results: $13.61 \pm 2.81\%$ and $53.26 \pm 2.83\%$ at IC_{50} value, respectively. The PLE-extract was considered as feed extract in the SAF process in order to obtain enriched fractions with improved antiproliferative activity, as described below.

3.2 Optimization of the SAF process for the antiproliferative activity improvement

According to previous works, the fractionation performance is dependent on the hydro-alcoholic composition of the feed extract, the feed solution to supercritical fluid ratio, the solids concentration in the feed and the pressure used for the antisolvent and separator stages [26, 28, 29]. For these reasons, the effects of pressure (8, 12 and 15 MPa), percentage of water in the feed extract (20, 30 and 50 v/v) and feed/sc-CO₂ flow rate ratio (0.0250, 0.0625 and 0.1000 w/w) were investigated in order to optimize the SAF process regarding the inhibitory cell proliferation capacity and extraction yield of selectively enriched fractions. For this purpose, a BBD was employed to evaluate the best combination of the factors for each response. Codified and real values of independent variables, as well as the resulting responses are summarized in Table 1. In each of the 15 experimental runs of BBD, the feed extract was separated into two

276 fractions according to its components' solubility in sc-CO₂ into a polar fraction called *Raffinate*
 277 (designated R1 to R15) and an ethanolic fraction called *Extract* (designated E1 to E15). The
 278 inhibitory cell proliferation of feed extract, raffinates and extracts was compared by one-way
 279 ANOVA. According to the interval plot (Figure 2) and Tukey's test, no significant differences
 280 ($p > 0.05$) among inhibitory cell proliferation (%) of raffinates were observed. Furthermore,
 281 the antiproliferative effect from raffinates (35.12 to 46.65%) was lower than that of the feed
 282 extract ($53.26 \pm 1.99\%$). Therefore, only the extract results were considered in the RSM
 283 analysis. The experimental data were fitted to various models and subsequent ANOVA (Table
 284 2) showed that inhibitory cell proliferation and yield were most suitably described with
 285 quadratic polynomial models. The goodness of models was evaluated by the coefficients of
 286 determination (R^2), which were close to 1 (0.9482-0.9723), indicating a high degree of
 287 correlation between the experimental and predicted values of the models. The models were
 288 highly significant and appropriately described the relationship between the responses and
 289 significant factors, since most terms showed high F -values and low p -values. In addition, the
 290 pure error values were low indicating reproducibility of the data obtained. However, the models
 291 showed lack-of-fit ($p < 0.05$) due to the similarity among the values of the central point of the
 292 design. The polynomial equation coefficients were calculated through RSM methodology in
 293 terms of coded units and are also provided in Table 2. Pareto charts and surface responses were
 294 plotted to display the influence of the SAF parameters on response variables (See Figure 3).
 295 Inhibitory cell proliferation was positively influenced by linear and quadratic effect of pressure,
 296 followed by water content in feed and the pressure-feed/sc-CO₂ interaction (Figure 3A). The
 297 feed/sc-CO₂ factor exhibited a negative effect on this response. In this sense, the highest
 298 inhibitory cell proliferation effect ($70.51 \pm 1.14\%$) was observed in the experiment 4 (E4) using
 299 50% of water v/v in the feed and 0.0625 feed/sc-CO₂ ratio at 15 MPa. This result represents an

improvement in the antiproliferative activity by SAF process, since E4 inhibited the cell proliferation 1.36-fold more than the original PLE extract at the same IC_{50} concentration.

On the other hand, yield was positively affected by both linear and quadratic effects of water percentage in feed and pressure (Figure 3B). This behaviour is also observed in the response surface plots since the yield was maximized at intermedium water in feed content and pressure conditions. In this sense, the feed components with more affinity for sc-CO₂ + EtOH phase were more easily separated from aqueous phase. In contrast, the highest and lowest level combinations for these factors caused a decrease in the yield of the extracts and the consequent increase in the raffinates. Thus, the highest yield in the extract (76.99%) was obtained employing 30% of water in feed, 0.0625 feed/sc-CO₂ ratio at 12 MPa, however under these conditions the inhibitory cell proliferation was $64.45 \pm 0.82\%$.

Based on the established regression model, a response optimization was performed using the Desirability function [30]. Inhibitory cell proliferation was considered and set to maximum. Profiles for predicted value estimated by desirability function are shown in Figure 4. The maximized response predicted was 71.57% under the following optimum conditions: 13.25 MPa, 50% of water v/v in the feed and 0.0625 feed/sc-CO₂ ratio. The desirability value was 1.00 indicating the highest maximisation degree for response optimization. However, the highest value observed ($70.51 \pm 1.14\%$) in BBD is within the confidence limits of predicted response (-CI 69.24 to +CI 73.91), therefore the values obtained by desirability approach under optimum conditions were not checked experimentally.

3.3 Phytochemical characterization of SAF fractions

The SAF fractions corresponding to extracts obtained in the above discussed experimental design were analysed by UHPLC-q-TOF-MS/MS and comparatively evaluated in terms of chemical composition with the feed extract. For this purpose, a targeted profiling analysis was

applied on the basis of the comprehensive phytochemical characterization of the phenolic fraction from MSK, previously reported by our group [27] . Thus, both the feed extract and the fractionated SAF extracts were shown to contain a broad variety of phenolic acids, flavonoids, benzophenones, C-glycosylxanones (e.g. mangiferin), gallic acid derivatives and gallotannins. Table S1 summarizes the q-TOF-MS/MS parameters used for the targeted profiling analysis of the main phenolic compounds previously identified in MSK extracts. Relative peak area for each compound was obtained in the MS mode using the precursor ion, i.e. the deprotonated molecular ion $[M-H]^-$, whereas one or two product ions, at least, were monitored in tandem mass spectrometry (MS/MS) mode for confirmation purposes. Extracted ion chromatograms (EIC) were obtained using a 20 ppm mass window, whereas a 5 ppm mass tolerance from the theoretical mass was set as positive identification criteria. The relative peak area values obtained for the target phenolic compounds identified both in feed and SAF extracts are shown in Table 3. Considering the diversity of isomeric forms, particularly for gallotannins, peak area values for these isomeric compounds were obtained as sum of each isomers' contribution.

Although no qualitative differences in terms of chemical composition was observed among the analysed samples, some significant quantitative changes could be observed. To better evaluate the differentially enriched compounds in the extracts obtained at different SAF conditions, a heatmap showing a colour code from high to low concentration ranging from dark red to dark blue is displayed in Figure 5. SAF extracts were grouped according to their phenolic content, as a result of a cluster analysis. Thus, four different groups can be clearly observed in the resulting heatmap; that is, Feed, E4, E1 and the rest of extracts. As can be seen, extract E4 and Feed are classified in opposite sides of the heatmap, showing the largest distance between samples, which indicate their different chemical composition compared to each other. On the other hand, E1 was significantly poorer in phenolic compounds, exhibiting the lowest phenolic

content among the SAF extracts. E1 was classified at the largest distance from E4, whereas the rest of extracts were grouped in a more homogeneous cluster between E1 and E4. In turn, the compounds of the extracts E1-E3 and E5-E15 are mainly grouped in two major clusters; that is, phenolics at higher concentration than feed extract (e.g. galloylquinic acid; maclurin glucosides; di-, tri-, tetra-, penta-, hexagalloyl glucose) and phenolics at lower levels compared to feed (e.g. ethyl and methylgallate, ellagic acid, mangiferin, homomangiferin, gallic acid, (epi)catechin, quercetin). However, the composition of E4 show some interesting discrepancies to the general trend followed by the majority of the extracts. For instance, phytochemicals such as mangiferin, homomangiferin, ellagic acid and galloyl methylgallate were found at notably higher levels in E4 compared to other extracts. The remarkable contrast shown by the heatmap between E4 and the feed extract evidence the differential enrichment in phenolic compounds obtained at the SAF experimental conditions corresponding to E4.

In order to explain the influence of the main SAF factors on the phenolic composition observed in the extracts, the individual compounds were considered as response variables in the quadratic model of the experimental design discussed in section 3.2. The influence of each factor was evaluated from their respective ANOVA, calculating the estimated linear, quadratic and interactions effects for each factor on the peak area of each phenolic compound. As can be seen in Table S2, the most important factor was the linear interaction between pressure and water percentage in feed (AB interaction). Maclurin-*C*-(*O*-galloyl)-glucoside, hexahydroxylated benzophenone isomer, ethyl trigallate, gallic acid and hexagalloyl glucose were some of the compounds most negatively affected by the AB interaction. In contrast, ethyl gallate, methyl gallate, ellagic acid, mangiferin and homomangiferin were positively affected by AB interaction. The linear effect of pressure was also important for ethyl gallate (negative), mangiferin (positive) and homomangiferin (positive). Furthermore, the water content in feed exerted simultaneously a negative effect on protocatechuic acid, iriflophenone glucoside and

4-Hydroxybenzoic acid, as well as a positive effect on homomangiferin. The quadratic effects of pressure, water in feed and feed/sc-CO₂ ratio showed a lower impact in the responses. Moreover, the feed/sc-CO₂ ratio and its linear interactions with the other factors showed very low effects on phenolic compounds distribution.

3.4 Correlation between antiproliferative activity and phenolic composition of SAF extracts

In an attempt to explain the enhanced bioactivity of the SAF extract E4 from the phytochemical composition point of view, a multivariate analysis was carried out. For this purpose, the different phenolic compounds identified in the targeted profiling analysis along with the antiproliferative activity values obtained for each extract of the SAF experiments were jointly evaluated in a principal component analysis (PCA). The proposed unsupervised multivariate analysis tool allows the evaluation of the compositional variability of the data in order to obtain the correlation structure between variables. According to the PCA results, graphically displayed in Figure 6, three dimensions are necessary for the complete explanation of the data variability (See Fig 6A-B scores and loadings, respectively). The three generated independent variables (PC1, PC2 and PC3) explained the 92% of the total variance. The first PC (PC1), accounting for a 47% of the variation, explains the difference between extract E1 and the rest of fractions. E1 has a high positive weight in PC1, exhibiting positive correlation with compounds such as iriflophenone, ethyl gallate and ellagic acid, whereas showing a strong negative correlation with more than a dozen compounds, which explain the weakness of this extract in terms of phenolic content (Fig. 6C). The second principal component (PC2) accounts for 36% of the variation, and this second variable allows explaining the difference between the PLE extract used as feed in the SAF process and the rest of the fractionated extracts. Feed extract shows a high negative weight on this second component and is strongly correlated with

(epi)catechin, protocatechuic acid, (epi)catechin gallate, quercetin, hexahydroxylated benzophenone, and galloyl diglucoside. These results are in good agreement with the higher content of these phenolic compounds in the Feed extract, as evidenced in the heatmap of Figure 5. Finally, a significant 9% of variance can still explain the remarkable difference between E4 and the rest of the fractions along the third principal component (PC3). Extract E4 shows an important contribution in the negative axis of PC3 and was shown to be the extract exhibiting the strongest correlation with the bioactivity, in agreement with the results obtained in the antiproliferative activity assay. As can be deduced from the correlation structure depicted in the biplot PCA of Figure 6D, the observed antiproliferative activity of extract E4 is in turn correlated with the presence of compounds such as mangiferin, homomangiferin, galloyl methylgallate, ellagic acid and methyl gallate, whereas quercetin is strongly non-correlated with the observed bioactivity.

These results obtained from this multivariate study show the existence of a notable variability in terms of chemical composition and bioactivity among the obtained SAF fractions, providing relevant information about the characteristic composition of the different groups samples, which might explain the antiproliferative activity of the extract E4.

In this regard, the anticancer potential of gallic acid derivatives has been recently reviewed [31]. Da Silva et al. [32] investigated the antiproliferative activity against ten human cancer cell lines of a series of galloyl derivatives including methyl gallate. This compound exhibited a significant growth inhibitory activity against HT-29 cells ($GI_{50} = 0.49$) and its powerful bioactivity was associated to galloyl group. Galloyl methyl gallate was another gallic acid derivative correlated with the inhibitory cell proliferation effect from E4 and its bioactivity could be related with the molecular mechanisms of gallic acid, however to the best of our knowledge the antiproliferative activity of this compound has not been described in the scientific literature. Ellagic acid has also exhibited antiproliferative activity against cancer

colon cell lines HTC116 and Caco-2 by cell cycle modulation, Bax translocation, caspase-8 activation and PCNA expression reduction [33]. Likewise, mangiferin is a xanthone-type compound mostly distributed in species from *Anacardaceae* family and its bioactivity, including antiproliferative and chemopreventive activities, has been widely documented [34-37]. Isomeric forms of mangiferin, such as homomangiferin, are also present in the mango organs and their co-occurrence has been linked to bioactivity from mango extracts [36].

4. CONCLUSIONS

In this work, an integrated extraction procedure involving supercritical antisolvent fractionation after sequential PLE was shown to be a suitable strategy to obtain phenolic rich extracts with improved inhibitory effect on HT-29 cells proliferation. The use of a co-current of pure sc-CO₂ with a water/ethanol mixture as solvent in the feed extract, along with a Box-Behnken experimental design for optimization and factors effect evaluation, makes the proposed strategy an efficient and green extractive approach with demonstrated fractionation capacity to valorize promising natural sources (e.g., MSK). Significant differences in the phenolics content could be observed from the phytochemical characterization and subsequent statistical analysis of the SAF extracts, evidencing a remarkable correlation between the antiproliferative activity and the content of characteristic xanthenes and gallic acid derivatives.

The results obtained herein can help to get a better understanding of MSK compounds responsible for the antiproliferative activity against HT-29 cells, thus providing the basis for the development of synergistic studies to investigate the interactions and contributions of the abovementioned phytoconstituents to the observed bioactivity.

450 **Acknowledgements**

451 This research was supported by COOPA20145, project from CSIC (Programa de Cooperación
452 Científica para el Desarrollo “i-COOP+”). G.A.-R. would like to acknowledge the Ministry of
453 Economy and Competitiveness for a “Juan de la Cierva” postdoctoral grant. The authors also
454 thank the support from the AGL2017-89417-R project.

References

- [1] A.P. Da Fonseca Machado, M. Rueda, G.F. Barbero, Á. Martín, M.J. Cocero, J. Martínez, Co-Precipitation of Anthocyanins of the Extract Obtained From Blackberry Residues By Pressurized Antisolvent Process, *J. Supercrit. Fluids*, 137 (2018) 81-92.
- [2] B. Gilbert-López, J.A. Mendiola, E. Ibáñez, Green foodomics. Towards a cleaner scientific discipline, *Trends Anal. Chem.*, 96 (2017) 31-41.
- [3] E. Ibáñez, A. Cifuentes, Green extraction techniques 2015, *Trends Anal. Chem.*, 71 (2015).
- [4] A. Nawab, F. Alam, M.A. Haq, A. Hasnain, Effect of guar and xanthan gums on functional properties of mango (*Mangifera indica*) kernel starch, *Int J Biol Macromol*, 93 (2016) 630-635.
- [5] M.H. Jahurul, I.S. Zaidul, K. Ghafoor, F.Y. Al-Juhaimi, K.L. Nyam, N.A. Norulaini, F. Sahena, A.K. Mohd Omar, Mango (*Mangifera indica* L.) by-products and their valuable components: a review, *Food Chem*, 183 (2015) 173-180.
- [6] C. Torres-León, R. Rojas, J.C. Contreras-Esquivel, L. Serna-Cock, R.E. Belmares-Cerda, C.N. Aguilar, Mango seed: Functional and nutritional properties, *Trends Food Sci Technol*, 55 (2016) 109-117.
- [7] A. Lopez-Cobo, V. Verardo, E. Diaz-de-Cerio, A. Segura-Carretero, A. Fernandez-Gutierrez, A.M. Gomez-Caravaca, Use of HPLC- and GC-QTOF to determine hydrophilic and lipophilic phenols in mango fruit (*Mangifera indica* L.) and its by-products, *Food Res Int*, 100 (2017) 423-434.
- [8] E. Dorta, M. González, M.G. Lobo, C. Sánchez-Moreno, B. de Ancos, Screening of phenolic compounds in by-product extracts from mangoes (*Mangifera indica* L.) by HPLC-ESI-QTOF-MS and multivariate analysis for use as a food ingredient, *Food Res Int*, 57 (2014) 51-60.

- 480 [9] A.S. Abdullah, A.S. Mohammed, A. Rasedee, M.E. Mirghani, Oxidative stress-mediated
481 apoptosis induced by ethanolic mango seed extract in cultured estrogen receptor positive
482 breast cancer MCF-7 cells, *Int J Mol Sci*, 16 (2015) 3528-3536.
- 483 [10] A.S. Abdullah, A.S. Mohammed, A. Rasedee, M.E. Mirghani, M.S. Al-Qubaisi, Induction
484 of apoptosis and oxidative stress in estrogen receptor-negative breast cancer, MDA-
485 MB231 cells, by ethanolic mango seed extract, *BMC Complement Altern Med*, 15 (2015)
486 45.
- 487 [11] F. Luo, Y. Fu, Y. Xiang, S. Yan, G. Hu, X. Huang, G. Huang, C. Sun, X. Li, K. Chen,
488 Identification and quantification of gallotannins in mango (*Mangifera indica* L.) kernel
489 and peel and their antiproliferative activities, *J Funct Foods*, 8 (2014) 282-291.
- 490 [12] B. Timsina, V.K. Nadumane, Mango seeds : a potential source for the isolation of
491 bioactive compounds with anti-cancer activity, *Int J Pharm Pharm Sci*, 7 (2015) 89-95.
- 492 [13] M.K. Ediriweera, K.H. Tennekoon, S.R. Samarakoon, A Review on
493 Ethnopharmacological Applications, Pharmacological Activities, and Bioactive
494 Compounds of *Mangifera indica* (Mango), *Evid Based Complement Alternat Med*, 2017
495 (2017) 6949835.
- 496 [14] K. Ameer, H.M. Shahbaz, J.-H. Kwon, Green Extraction Methods for Polyphenols from
497 Plant Matrices and Their Byproducts: A Review, *Compr Rev Food Sci Food Saf*, 16
498 (2017) 295-315.
- 499 [15] I.M. Prado, G.H.C. Prado, J.M. Prado, M.A.a. Meireles, Supercritical CO₂ and low-
500 pressure solvent extraction of mango (*Mangifera indica*) leaves: Global yield, extraction
501 kinetics, chemical composition and cost of manufacturing, *Food Bioprod Process*, 91
502 (2013) 656-664.

- 503 [16] M.P. Garcia-Mendoza, J.T. Paula, L.C. Paviani, F.A. Cabral, H.A. Martinez-Correa,
504 Extracts from mango peel by-product obtained by supercritical CO₂ and pressurized
505 solvent processes, *LWT-Food Sci Technol*, 62 (2015) 131-137.
- 506 [17] F.G. Calvo-Flores, M.J. Monteagudo-Arrebola, J.A. Dobado, J. Isac-Garcia, Green and
507 Bio-Based Solvents, *Top Curr Chem (Cham)*, 376 (2018) 18.
- 508 [18] M. Herrero, M. Castro-Puyana, J.A. Mendiola, E. Ibañez, Compressed fluids for the
509 extraction of bioactive compounds, *Trends Anal. Chem.*, 43 (2013) 67-83.
- 510 [19] I.A. Cuadra, F. Zahran, D. Martín, A. Cabañas, C. Pando, Preparation of 5-fluorouracil
511 microparticles and 5-fluorouracil/poly(l-lactide) composites by a supercritical CO₂
512 antisolvent process, *J. Supercrit. Fluids*, 143 (2019) 64-71.
- 513 [20] L. Martín, A. González-Coloma, R. Adami, M. Scognamiglio, E. Reverchon, G.D. Porta,
514 J.S. Urieta, A.M. Mainar, Supercritical antisolvent fractionation of ryanodol from *Persea*
515 *indica*, *J. Supercrit. Fluids*, 60 (2011) 16-20.
- 516 [21] J.T. Paula, I.M.O. Sousa, M.A. Foglio, F.A. Cabral, Selective fractionation of extracts of
517 *Arrabidaea chica* Verlot using supercritical carbon dioxide as antisolvent, *J. Supercrit.*
518 *Fluids*, 133 (2018) 9-16.
- 519 [22] Y.M. Monroy, R.A.F. Rodrigues, M.V.N. Rodrigues, F.A. Cabral, Fractionation of
520 ethanolic and hydroalcoholic extracts of green propolis using supercritical carbon
521 dioxide as an anti-solvent to obtain artepillin rich-extract, *J. Supercrit. Fluids*, 138 (2018)
522 167-173.
- 523 [23] M.a. Meneses, G. Caputo, M. Scognamiglio, E. Reverchon, R. Adami, Antioxidant
524 phenolic compounds recovery from *Mangifera indica* L. by-products by supercritical
525 antisolvent extraction, *J Food Eng*, 163 (2015) 45-53.

- 526 [24] M.C. Guamán-Balcázar, A. Montes, C. Pereyra, E.M. de la Ossa, Precipitation of mango
527 leaves antioxidants by supercritical antisolvent process, *J. Supercrit. Fluids*, 128 (2017)
528 218-226.
- 529 [25] M.C. Guamán-Balcázar, A. Montes, M.T. Fernández-Ponce, L. Casas, C. Mantell, C.
530 Pereyra, E. Martínez de la Ossa, Generation of potent antioxidant nanoparticles from
531 mango leaves by supercritical antisolvent extraction, *J. Supercrit. Fluids*, 138 (2018) 92-
532 101.
- 533 [26] A.P. Sánchez-Camargo, J.A. Mendiola, A. Valdés, M. Castro-Puyana, V. García-Cañas,
534 A. Cifuentes, M. Herrero, E. Ibáñez, Supercritical antisolvent fractionation of rosemary
535 extracts obtained by pressurized liquid extraction to enhance their antiproliferative
536 activity, *J. Supercrit. Fluids*, 107 (2016) 581-589.
- 537 [27] D. Ballesteros-Vivas, G. Alvarez-Rivera, S.J. Morantes-Medina, A.d.P. Sanchez
538 Camargo, E. Ibáñez, F. Parada-Alfonso, A. Cifuentes, An integrated approach for the
539 valorization of mango seed kernel: efficient extraction solvent selection, phytochemical
540 profiling and antiproliferative activity assessment, *J. Agric. Food Chem.*, Submitted
541 (2019).
- 542 [28] O.J. Catchpole, J.B. Grey, K.A. Mitchell, J.S. Lan, Supercritical antisolvent fractionation
543 of propolis tincture, *J. Supercrit. Fluids*, 29 (2004) 97-106.
- 544 [29] P. Kraujalis, P.R. Venskutonis, E. Ibáñez, M. Herrero, Optimization of rutin isolation from
545 *Amaranthus paniculatus* leaves by high pressure extraction and fractionation techniques,
546 *J. Supercrit. Fluids*, 104 (2015) 234-242.
- 547 [30] M.M. Feuereisen, M. Gamero Barraza, B.F. Zimmermann, A. Schieber, N. Schulze-
548 Kaysers, Pressurized liquid extraction of anthocyanins and biflavonoids from *Schinus*
549 *terebinthifolius* Raddi: A multivariate optimization, *Food Chem*, 214 (2017) 564-571.

- [31] S. Verma, A. Singh, A. Mishra, Gallic acid: Molecular rival of cancer, *Environ. Toxicol. Pharmacol.*, 35 (2013) 473-485.
- [32] M.M. da Silva, M. Comin, T.S. Duarte, M.A. Foglio, J.E. de Carvalho, M.C. do Vieira, A.S. Formagio, Synthesis, antiproliferative activity and molecular properties predictions of galloyl derivatives, *Molecules (Basel, Switzerland)*, 20 (2015) 5360-5373.
- [33] A.I. Yousef, O.S. El-Masry, E.H. Yassin, The anti-oncogenic influence of ellagic acid on colon cancer cells in leptin-enriched microenvironment, *Tumour Biol*, 37 (2016) 13345-13353.
- [34] M. Imran, M.S. Arshad, M.S. Butt, J.H. Kwon, M.U. Arshad, M.T. Sultan, Mangiferin: a natural miracle bioactive compound against lifestyle related disorders, *Lipids Health Dis*, 16 (2017) 84.
- [35] M. Adam, K. Piotr, G. Edyta, W. Dorota, Mangiferin – a Bioactive Xanthonoid, not only from Mango and not just Antioxidant, *Mini-Rev. Med. Chem.*, 13 (2013) 439-455.
- [36] R.K. Khurana, R. Kaur, S. Lohan, K.K. Singh, B. Singh, Mangiferin: a promising anticancer bioactive, *Pharm Pat Anal*, 5 (2016) 169-181.
- [37] P. Rajendran, T. Rengarajan, N. Nandakumar, H. Divya, I. Nishigaki, Mangiferin in cancer chemoprevention and treatment: pharmacokinetics and molecular targets, *J Recept Signal Transduct Res*, 35 (2015) 76-84.

575 **Figure captions**

576 **Figure 1.** Schematic representation of the experimental SAF unit.

577 **Figure 2.** Interval plot for inhibitory cell proliferation from feed extract and SAF fractions.

578 **Figure 3.** Standardized Pareto charts and response surface plots of SAF extracts showing
579 independent variables effects on A) inhibitory cell proliferation and B) yield.

580 **Figure 4.** Desirability value and predicted response value in response-optimization.

581 **Figure 5.** Heatmap showing the distribution of the main phenolic components (P1-P33) in the
582 Feed and SAF extracts (E1-E15).

583 **Figure 6.** Principal component analysis (PCA) of feed and SAF extracts. A) 3D-Scores plot,
584 B) 3D-loadings plot, C) PC1 – PC2 biplot, D) PC2-PC3 biplot. The first three components of
585 the PCA show 92% of the total variance: 47% for PC 1, 36% for PC 2 and 9% for PC 3.