



UNIVERSIDAD
DE GRANADA



Dpto. de Química Analítica
Prof. Fermín Capitán García

APLICACIONES ANALÍTICAS DE LAS "HUELLAS DACTILARES" CROMATOGRÁFICAS EN CIENCIAS DE LOS ALIMENTOS Y DE LA SALUD

Tesis Doctoral
Lucía Valverde Som

Directores:
Dr. Luis Cuadros Rodríguez
Dra. Cristina Ruiz Samblás

para optar al título de:
Doctora en Química
Granada

Editor: Universidad de Granada. Tesis Doctorales

Autora: Lucía Valverde Som

ISBN: 978-84-9163-744-8

URI: <http://hdl.handle.net/10481/49166>

AGRADECIMIENTOS

AGRADECIMIENTOS

"Un duelo está finalizado cuando lo que queda es el sentimiento inequívoco de AGRADECIMIENTO" Carlos Odrizola

Todo en esta vida es movimiento y tiene un principio y un fin.

Ha llegado la hora de terminar con esta etapa de mi vida, terminar este proceso de crecimiento laboral y personal que he conseguido durante estos cuatro años de Tesis. Por lo que quiero empezar AGRADECIENDO todo lo aprendido, todo lo madurado, todo lo experimentado.

Primero, agradecer a la persona más importante de mi vida: **YO**. Quiero agradecerme todo el trabajo que he hecho, toda mi evolución, todo el cambio, recorriendo mi camino y saltando todos los obstáculos que se han podido poner en él. Gracias a mi constancia, a no desistir, a ser valiente, a ser FUERTE, a ser responsable... he sabido afrontarlos, pudiendo así culminar esta Tesis Doctoral.

"Discúlpeme, no le había reconocido: he cambiado mucho" Oscar Wilde

Han sido muchas las personas que me han acompañado en este arduo camino, y que han contribuido a que mejore como persona, como química y como profesional investigadora.

Lo primero de todo, agradecer **a mis directores de tesis**, ellos han hecho posible que adquiriera muchas competencias en el ámbito profesional, me han enseñado a desenvolverme ante las dificultades, me han acompañado en la elaboración de cada una de las partes experimentales y de tratamiento de datos de esta Tesis Doctoral durante los cuatro años que llevo realizándola.

Luis, gracias por todo lo que me has enseñado. Gracias por proponerme hacer la Tesis Doctoral, por darme la oportunidad de participar en todos los proyectos que he podido. Por todas las oportunidades que me has dado para aprender todo lo que sé en el mundo laboral; por permitirme aprender de tu orden; de cómo trabajar en un laboratorio de calidad; de cómo poder llegar al

público, sea científico o no; cómo elaborar un proyecto, como corregirlo... Gracias por enseñarme a generar conocimiento, a conocer otro punto de vista... Gracias por la confianza que has puesto en mí durante todos estos años, me has dado tareas para aumentar mi formación y me has dado la oportunidad de aprender a gestionar las situaciones cuando me han venido grandes. Gracias por tu paciencia ante mis dudas y ante mis lluvias de ideas, de inteligencia, donde mi cerebro estaba tan activo que casi ni me entendías... Gracias por esos días de poesía y por esos momentos de charla más allá de lo profesional. GRACIAS.

Cristina, una y mil gracias por todo. Creo que me queda poco que decirte que no sepas... Gracias por todo lo que me has enseñado del laboratorio. Has sido un gran apoyo para mí, un referente. Empecé teniendo mil dudas, preguntándote a cada momento, equivocándome... y me has ido dando las herramientas para corregirlo, para saber que no pasaba nada y para así APRENDER. Gracias por apoyar mis ideas y darme otros puntos de vista que han hecho que mi mente se abra más al conocimiento. Gracias por tu dedicación, paciencia y apoyo. Llegó el momento en el que te hiciste doctora y te marchaste del labo. Fue entonces un paso más que tenía que dar para crecer, era el momento de aplicar todo lo que me habías enseñado, de gestionar los problemas del equipo, de hablar con técnicos, con comerciales, y de estar un poco "al mando" de todo. Gracias por estar ahí, por seguirme en cada paso, aun cuando hemos estando lejos. GRACIAS. Gracias por formar parte de mi vida profesional y personal y... por seguir estando ahí.

Gracias **a mi familia**. Lo primero de todo **a mis padres, Paco y María Eugenia**. Gracias por apoyarme en cada uno de mis sueños. Gracias por estar ahí cuando no he sabido que hacer y darme la oportunidad de estudiar, de continuar con mi formación y de darme ese aliento en cada uno de los agobios que he tenido. Gracias por animarme para que hiciera la Tesis Doctoral. Gracias por estar ahí siempre, por apoyarme, darme ánimos y fuerza. "A por ellos, que son pocos y cobardes"... Y ahora... **a mi Naniiii, Esther**, gracias por estar ahí, por ser la mejor hermana del mundo mundial!! Por hablarme desde la barriga, por cuidarme siempre, por apoyarme, por escucharme. Quería ser como tú y gracias a eso... me puse manos a la obra y me puse a estudiar, aunque me costara trabajo. Siempre he visto cómo te has esforzado, con constancia siempre sigues adelante y solucionas los problemas. Gracias por ayudarme cuando no he podido avanzar, aunque solo bastara con... Nani, esto no lo entiendo; ¿el que?; ahm! no, ya ya ya seee... Gracias por ayudarme con todo, gracias por animarme a seguir. Gracias por ser tú. GRACIAS.

Gracias **a mi grupo de investigación, AnAMA**, por acogerme siempre como una más. Gracias **Antonio**, has sido un gran compañero y acompañante de este camino, por esos momentos de café, de videos, de despejes y de risas. Gracias **Brigui** por esos momentos de té y de ver qué pasa con el SICA y por poder compartir alguna locura...¿quién dijo boxeo? :P. Gracias **Domingo** por tus explicaciones sobre algunas lecciones de precolumnas...GRACIAS. Gracias **al triángulo de las bermudas, Ana y Raquel**. Gracias por esos momentos de apoyo en los que todo se ve negro, en los que parece que no merece la pena investigar, en los que rechazan una y otra vez tú trabajo. Gracias por estar ahí apoyándome y sobre todo, gracias por los buenos momentos que pasamos de risas, por esos desayunos improvisados para parar un poco, por esa música con mensajes subliminares para desconectar...por las fotos y los videos...Gracias por todos esos momentos! Y aunque de forma esporádica, **a la cuarta jinete (Raquel Rovi) de las jinetes del apocalipsis** :P. GRACIAS. También gracias a **Fidel**, el nuevo integrante del grupo, gracias por todo el apoyo que me has dados en estos meses finales, por aguantar todas mis preocupaciones y mis estados de ánimo en el labo, y sobre todo, gracias por cuidarme con esos desayunos y meriendas acompañados del "ahorcado" :P. GRACIAS.

Ahora es el turno de agradecer **a los compañeros del departamento de Química Analítica**. A **Alegría**, por estar siempre ahí, por ayudarme sobre todo con el estudio del colesterol y la parte en la que hemos trabajado juntas del proyecto SensOLEO, me has ayudado y enseñado otra forma de trabajar. Gracias por esas preguntas para ver como "vender" mi producto, el método analítico, ¿qué mejor?. Y... por esos momentos apostando y preguntándole a la gran bola del conocimiento ;). Por esos cafelitos o desayuno mañanero actualizándonos la vida. GRACIAS. A **Ignacio** siempre preocupado de que descansara los fines de semana, controlando que llevara la maleta siempre al labo... Gracias a **Loli y Ana García Campaña**, por preocuparos por mí. Gracias a **Natalia**, por estar ahí cuando he necesitado ayuda. A **Alberto Zafra**, por esos momentos de corta-columna...y a cada uno de los que formáis el departamento, por estar siempre ahí, dándome ánimos y lo que he necesitado.

A **Chiara, Federico, Francesca y Carlo** por acogerme durante la estancia doctoral en su departamento y aportarme otro modo de trabajar, así como ratos de charlas y desconexión.

Gracias **a mi otra familia**, el mejor minigrupo de la historia :P. **Alma, Ana, Antonio, Cristina, Nacho y Rosa**. En especial a **Ana** y sobre todo a **Antonio**. Gracias por enseñarme tanto, gracias por apoyarme siempre,

gracias por estar ahí, gracias por cuidarme, gracias por confiar en mí, gracias por creer en mí, aun cuando yo no creía en mí misma...gracias por darme ánimos para seguir, para fomentar mi constancia y darle la espalda a la desidia. Gracias a **Carolina** y a **Laura**, porque sin vosotras aún estaría más perdida, aún sería más pequeña, me hubiera costado mucho saber buscar y conseguir la valentía, la fuerza y el poder que tengo ahora. GRACIAS. Gracias **al resto de compis** y **maestros** que la vida me ha regalado, la vida os ha puesto ahí para aprender de vosotros. GRACIAS por todo lo que me habéis enseñado, por esos llantos, por esas risas y por esas quedadas y reencuentros. GRACIAS.

Gracias a **Amor, Alice** y **Gaia**, que me acogieron como parte de su familia en mi estancia en Turín, gracias por los helados, piratas del caribe y todos esos momentos de risas y música con Alice. GRACIAS.

Gracias **a esa tribu!!** Que me ha ayudado a despejarme y a liberar mi mente de sus preocupaciones... Sobre todo a **Estefanía**, por todos los momentos :).

Gracias **a las locas del dibujo técnico. María**, gracias por apoyarme en que escribiera la Tesis, en darle caña, por darme un poco de tu energía, ganas y fuerzas, por esa constancia. Ya sabes que me encanta eso de ti, y me has transmitido un pocooo. GRACIAS.

Gracias a mis compis de licenciatura, mis **quimicuchos**, por todos los momentos, de risas y de concentración, que hemos vivido hasta ser QUÍMICOS, en especial por la primera cena de Navidad en el Burger King. GRACIAS. Gracias sobre todo **a las buenorras de corpus y alma** y **Fermín**. Me habéis ayudado mucho en esta última etapa académica, habéis sido un gran apoyo, sabiendo que estabais ahí en todo momento y dándome ánimos. **Tamara**, a pesar de estar casi a 900 km de distancia siempre te siento al lado, GRACIAS.

Gracias **a mi otra familia**...si, otra más, gracias de nuevo a **Cristina**, a **Fran** y sobre todo, al pequeño **Leo**. Gracias por acogerme como una más y gracias por todos esos momentos compartidos. Leo me ha enseñado a ser feliz sin más, a ser feliz solo mirándolo, a ser feliz y que se me olviden todos los problemas, todos los agobios y todas las preocupaciones de la Tesis.

GRACIAS A TODOS.

"El preocuparnos no le quita problemas al mañana, le quita fuerzas al hoy" J.R.R. Tolkien.

ÍNDICE

ÍNDICE

RESUMEN	21
ABSTRACT	22
PRESENTACIÓN	25
Donde empieza todo	26
En esta Tesis Doctoral	28
Contextualización	29
Problema general	36
Hipótesis	37
Objetivos	38
Artículos científicos	39
Otros trabajos	40
CAPÍTULO 1. Determinación de precursores y derivados del colesterol en muestras de sangre	43
1.1. PRESENTACIÓN	45
1.2. INTRODUCCIÓN	46
Colesterol	46
Metabolitos – Oxiesteroles	48
Enfermedad de Alzheimer	50
Instrumentación	51
1.3. MUESTRAS	54
Muestras de plasma de ratón	54
Muestras de suero de cabra	55

1.4. PROCESO EXPERIMENTAL	56
1.5. ARTÍCULO: Separation and determination of some of the main cholesterol-related compounds in blood by gas chromatography-mass spectrometry (selective ion monitoring mode)	60
Abstract	60
Keywords	61
1. Introduction	61
2. Material and methods	66
2.1. Chemicals and reagents	66
2.2. Standards solutions	67
2.3. Samples and sample preparation	67
2.4. Extract sample treatment	69
2.5. GC-(EI/IT)MS instrument and method conditions	69
3. Results and discussion	72
3.1. Optimization of chromatographic and detection conditions	72
3.2. Validation of the developed method	77
3.3. Application of the method to the analysis of biological samples	82
4. Conclusions	86
Acknowledgments	86
Ethical conduct of research	87
Author contributions	87
Conflicts of interest	87
References	88

1.6. CONCLUSIONES	93
1.7. COMUNICACIÓN A CONGRESO	95
ANEXO I. ESTUDIOS COMPLEMENTARIOS NO PUBLICADOS	96
I.1. MUESTRAS INICIALES	96
I.2. ESPECTROS DE MASAS	100
I.3. RECTAS DE CALIBRADO	106
I.4. HOJA DE CONTROL DE PATRONES Y CONTROL DE CALIDAD	109
CAPÍTULO 2. "Huellas dactilares" cromatográficas	113
2.1. PRESENTACIÓN	115
2.2. INTRODUCCIÓN	116
Señal analítica	116
Señal específica	116
Señal inespecífica	117
Datos analíticos	117
"Huella dactilar" instrumental y técnicas instrumentales	118
Quimiometría	121
Aplicaciones	128
2.3. ARTÍCULO 1: <i>Chromatographic fingerprinting: An innovate approach for food 'identification' and food authentication - A tutorial</i>	131
Abstract	132
Keywords	133
1. Introduction	133
2. Food authentication: targets and approaches. Current perspectives	138

2.1. Food metabolomics, foodomics and food authentication	144
2.2. Food authentication and food 'identitation' by fingerprinting	147
3. Chromatographic fingerprinting for food authentication: framework	148
3.1. Chromatographic signals and analytical data	149
3.2. Chromatographic profiles and chromatographic fingerprints	152
3.3. Chromatographic fingerprints methodology	156
4. Applications of chromatographic fingerprinting on food 'identitation' and food authentication	160
4.1. Gas chromatography	161
4.2. Liquid chromatography	168
4.3. Comprehensive 2D chromatography	169
5. Concluding remarks	175
Acknowledgments	176
Appendix A. List of acronyms	176
References	180
2.4. ARTÍCULO 2: Sobre el uso de 'huellas dactilares' cromatográficas en la autenticación de alimentos	193
1. Autenticación de alimentos	193
1.1. Enfoques y metodologías	194
2. Huellas dactilares cromatográficas: marco metodológico	198
2.1. Señales cromatográficas y datos analíticos	200

2.2. Perfiles cromatográficos y huellas dactilares cromatográficas	205
2.3. Métodos y técnicas quimiométricas	207
2.4. 'Identificación' de alimentos	210
3. Aplicaciones de las huellas dactilares cromatográficas en 'identificación' y autenticación de alimentos	212
Referencias	214
2.5. CONCLUSIONES	218
2.6. COMUNICACIONES A CONGRESOS	219
CAPÍTULO 3. Monitorización de la estabilidad de patrones de aceite de oliva para análisis sensorial.....	221
3.1. PRESENTACIÓN	223
3.2. INTRODUCCIÓN	224
Aceite de oliva virgen	224
Análisis sensorial	225
Materiales de referencia certificados para aceite de oliva	227
Compuestos responsables del flavor	229
Métodos analíticos para el estudio de la estabilidad	230
Métodos de extracción para la determinación de volátiles	232
Instrumentación	233
3.3. MUESTRAS	236
3.4. PROCESO EXPERIMENTAL	238
Pre-tratamiento de los datos	241
Gráficos de control	243
3.5. ARTÍCULO 1: Multivariate approaches for chemical chromatographic control of the stability of the organoleptic quality of olive oil reference materials for sensory analysis -	254

Part I: Framework and fundamentals	
Abstract	255
Keywords	255
1. Introduction	255
2. Fundamentals	259
Abbreviations used	273
Acknowledgments	274
References	275
3.6. ARTÍCULO 2: Multivariate approaches for chemical chromatographic control of the stability of the organoleptic quality of olive oil reference materials for sensory analysis - Part II: Applications	281
Abstract	282
Keywords	282
1. Introduction	283
2. Material and methods	287
2.1. Olive oil reference materials	287
2.2. Sensory analysis	288
2.3. Analysis of VOCs in the samples	289
2.4. Quality control solution	290
2.5. GC-FID instrument and method conditions	290
2.6. HS-SPME and injection conditions	291
2.7. Chemometrics - Data treatment	292
3. Results and discussion	293
3.1. Sensory analysis	293
3.2. Chromatographic analysis	294

Abbreviations used	305
Acknowledgments	306
References	307
3.7. CONCLUSIONES	312
3.8. COMUNICACIONES A CONGRESOS	313
ANEXO II. ESTUDIOS COMPLEMENTARIOS NO PUBLICADOS	315
II.1. HOMOGENEIDAD DE LOS MRC	315
II.2. SELECCIÓN DE PICOS CARACTERÍSTICOS (MARCADORES)	318
II.3. MONITORIZACIÓN DE LA ESTABILIDAD	357
CAPÍTULO 4. Evaluación de la madurez de la aceituna	369
4.1. PRESENTACIÓN	371
4.2. INTRODUCCIÓN	372
Formación de compuestos orgánicos volátiles	372
Maduración de la aceituna	375
Instrumentación	377
4.3. MUESTRAS	381
4.4. PROCESO EXPERIMENTAL	385
Tratamiento de los datos	387
4.5. ARTÍCULO: <i>Combined untargeted and targeted fingerprinting with comprehensive two-dimensional chromatography for volatiles and ripening indicators in olive oil</i>	396
Abstract	397
Key-words	398
1. Introduction	398
2. Materials and methods	401

2.1. Reference compounds and solvents	401
2.2. Olive oil samples	401
2.3. Head-space solid phase micro extraction sampling devices and conditions	402
2.4. Comprehensive two-dimensional gas chromatographic system (GC×GC-MS) set-up and analytis conditions	404
2.5. Raw data acquisition and GC×GC data handling	405
2.6. Profiling and advanced fingerprinting work-flow	405
3. Results and discussion	408
3.1. Targeted analysis and samples discrimination.....	410
3.2. Untargeted analysis	420
3.3. Retrospective analysis and definition of reliable chemical indexes of ripening	424
3.4. Fingerprinting by image features approach	430
4. Conclusions	433
Acknoledgments	434
References	435
4.6. CONCLUSIONES	441
4.7. COMUNICACIÓN A CONGRESO	442
CONCLUSIONES FINALES	445
Principales aportaciones	447
Perspectivas futuras	448
Valoración personal	449

RESUMEN/ABSTRACT

RESUMEN

Esta Tesis Doctoral describe y desarrolla diferentes aplicaciones en el ámbito de ciencias de la salud y de los alimentos. Las aplicaciones se centran en: (i) muestras de sangre para el estudio del colesterol y sus metabolitos como posibles biomarcadores de la enfermedad del Alzheimer; (ii) muestras de aceite de oliva virgen para la monitorización de la estabilidad organoléptica, así como del grado de madurez de la aceituna, a partir de la fracción de los compuestos orgánicos volátiles.

El objetivo principal es el desarrollo de metodologías analíticas innovadoras que conlleven el empleo de diferentes técnicas separativas como son la cromatografía de gases o cromatografía de gases bidimensional, acopladas a detectores de espectrometría de masas, con diferentes tipos de analizadores, así como detectores de ionización de llama. A través de dichas técnicas es posible obtener información analítica útil para resolver los problemas relacionados con cada línea de investigación.

Por tanto, se describen diferentes métodos analíticos y de tratamiento de datos para resolver los problemas de cada aplicación. Se aplica la cromatografía clásica de detección y cuantificación de los analitos para el estudio centrado en las ciencias de la salud; y la metodología de "huella dactilar" cromatográfica, junto con quimiometría (gráficos de control multivariable en cromatografía unidimensional y enfoques dirigido (*target*) y no dirigido (*untargeted*) de las regiones de los picos detectados en la cromatografía bidimensional), para el ámbito de ciencias de los alimentos.

ABSTRACT

This Doctoral Thesis describes and develops different applications in health and food sciences. The applications are focused on: (i) the amount of several precursors of cholesterol and oxysterols in blood samples as possible biomarkers of Alzheimer's disease; (ii) the control of the organoleptic stability of olive oil and the ripeness of the olive fruit by mean of monitoring the volatile organic compounds profile.

The main objective is the development of innovative analytical methodologies which use different separation techniques as gas chromatography and comprehensive two-dimensional gas chromatography. Both techniques coupled to different detectors as: mass spectrometry (with different analyzers) and flame ionization detector. All these techniques give enough analytical information to solve the problems of each research field.

Thus, different analytical methods and data treatment to find a solution to each challenge have been described. The classic chromatography is used to detect and quantify the analytes of interest in health science. However, the chromatographic fingerprinting methodology in conjunction with chemometric (multivariate control charts by one-dimensional chromatography and *targeted* and *untargeted* approaches based on peak-region by comprehensive two-dimensional chromatography) is used in food science.

PRESENTACIÓN

PRESENTACIÓN

Esta Tesis Doctoral está regida por el Real Decreto 99/2011, por el que se regulan las enseñanzas oficiales de doctorado, modificado por el Real Decreto 534/2013 en lo relativo a la evaluación y defensa de la Tesis Doctoral.

En el artículo 5 del Real Decreto 99/2011, se especifica que:

La tesis doctoral consistirá en un trabajo original de investigación elaborado por el candidato en cualquier campo del conocimiento. La tesis debe capacitar al doctorando para el trabajo autónomo en el ámbito de la I+D+i.

y se indican las competencias básicas que ha de adquirir el doctorando:

- a) *Comprensión sistemática de un campo de estudio y dominio de las habilidades y métodos de investigación relacionados con dicho campo.*
- b) *Capacidad de concebir, diseñar o crear, poner en práctica y adoptar un proceso sustancial de investigación o creación.*
- c) *Capacidad para contribuir a la ampliación de las fronteras del conocimiento a través de una investigación original.*
- d) *Capacidad de realizar un análisis crítico y de evaluación y síntesis de ideas nuevas y complejas.*
- e) *Capacidad de comunicación con la comunidad académica y científica y con la sociedad en general acerca de sus ámbitos de conocimiento en los modos e idiomas de uso habitual en su comunidad científica internacional.*
- f) *Capacidad de fomentar, en contextos académicos y profesionales, el avance científico, tecnológico, social, artístico o cultural dentro de una sociedad basada en el conocimiento.*

Así como la capacidad profesional en ámbitos que requieran creatividad e innovación, que le será otorgado tras la obtención del título de Doctor y que supone el haber adquirido las capacidades y destrezas personales para:

- a) *Desenvolverse en contextos en los que hay poca información específica.*
- b) *Encontrar las preguntas claves que hay que responder para resolver un problema complejo.*

-
- c) Diseñar, crear, desarrollar y emprender proyectos novedosos e innovadores en su ámbito de conocimiento.*
 - d) Trabajar tanto en equipo como de manera autónoma en un contexto internacional o multidisciplinar.*
 - e) Integrar conocimientos, enfrentarse a la complejidad y formular juicios con información limitada.*
 - f) La crítica y defensa intelectual de soluciones.*

Esta Tesis Doctoral se enmarca en el Programa Oficial de Doctorado en Química, bajo la línea de investigación "Metodologías de obtención de información analítica en sistemas reales" dentro de la Escuela de Doctorado en Ciencias, Tecnologías e Ingenierías de la Universidad de Granada.

Dónde empieza todo...

Esta Tesis Doctoral ha sido llevada a cabo en el seno de la Unidad de Cualimetría y Metrología Química (CMQ) y del grupo de investigación Análisis en Alimentación y Medio Ambiente (AnAMA, FQM-232), y ha sido realizada en el Departamento de Química Analítica, en la Facultad de Ciencias de la Universidad de Granada, bajo la tutela y dirección del Dr. Luis Cuadros Rodríguez y la Dra. Cristina Ruiz Samblás.

Las muestras utilizadas durante la Tesis Doctoral proceden de diferentes proyectos científicos, que se citan a continuación, en los que la doctoranda ha trabajado y colaborado y que han sido concedidos al grupo de investigación, AnAMA.

- "Potencial de diferentes plataformas GC-MS de aplicación en metabolómica: dos casos de estudio". Proyecto financiado por el Campus de Excelencia Internacional (CEI BioTic) Granada (Ref.: CEI2013-P-17).
- "Certificación y distribución de nueve materiales de referencia certificados (MRC) para análisis organoléptico de aceite de oliva (SensOLEO-MRC 2015)". Proyecto financiado por la Consejería de

Agricultura, Pesca y Desarrollo Rural (CAPDER), Junta de Andalucía (Ref.: OTRI 3637-2015).

- "Evolución de la calidad del aceite en función de la maduración de la aceituna en cultivos del Altiplano de Granada". Colaboración que se llevó a cabo con "GDR Altiplano de Granada".

Además, una parte significativa de esta Tesis Doctoral, y que constituye un capítulo de la misma, ha sido realizada durante una estancia doctoral de tres meses de duración en el "Dipartimento di Scienza e Tecnologia del Farmaco" de la Universidad de Turín (Italia) bajo la supervisión de la Dra. Chiara E. Cordero. Para la realización de esta estancia se recibió una ayuda por parte de la Universidad de Granada y el CEI BioTic Granada bajo el nombre "Convocatoria de movilidad internacional de jóvenes investigadores de programas de doctorado universidad de Granada y CEI BioTic Granada. Curso académico 2014/2015".

En esta Tesis Doctoral

Antes de desarrollar cada uno de los capítulos en los que se describe la investigación llevada a cabo, se hace una contextualización de la Tesis Doctoral, donde se menciona de forma resumida de donde "nace" el desarrollo de la misma y que fue objeto del plan de investigación que dio paso a la incorporación de la doctoranda en el Programa de Doctorado en Química.

En la parte de contextualización, se resumen las dos líneas de investigación que se abarcan en la Tesis y se va haciendo un recorrido por cada uno de los capítulos que se van a presentar.

Tras esto, se indican el problema, hipótesis y objetivos generales de la misma, incluyendo un listado de los artículos científicos que implican el culmen de cada una de las investigaciones desarrolladas. Así como un listado de artículos científicos y proyectos complementarios en los que se ha trabajado.

Seguidamente, se exponen un total de cuatro capítulos, cada uno de los cuales incluye una breve presentación, indicando los objetivos concretos del propio capítulo, para pasar a una introducción que pone en contexto el estudio, las muestras utilizadas, así como el proceso experimental llevado a cabo. A continuación aparecen el artículo o artículos científicos que derivan de la investigación y que muestran los resultados obtenidos, para finalizar con las conclusiones específicas de dicho capítulo. Por último, se muestra un listado de las comunicaciones a congresos que derivan de cada capítulo.

Algunos capítulos incluyen anexos en los que se muestran los estudios complementarios que no fueron publicados en dicha investigación.

Finalmente, se presentan las conclusiones finales, y se destacan las principales aportaciones de la Tesis, atendiendo al interés para la comunidad científica. Además, se exponen nuevas aportaciones que permitirían llevar a cabo investigaciones futuras y una valoración de todo lo que conlleva esta Tesis, a nivel científico y personal.

Contextualización

Esta Tesis Doctoral está centrada en el desarrollo de diversos métodos analíticos enfocados en obtener información global y particular de metabolitos en muestras en el **ámbito de la salud y de la alimentación**, pudiendo dividir la Tesis en sendas líneas de investigación.

Se denomina **metabolitos** a los compuestos resultantes de las reacciones enzimáticas que tienen lugar en los organismos. Aquellos que tienen una función esencial en el crecimiento o reproducción del organismo donde se producen se denominan metabolitos primarios, a diferencia de lo que sucede con los metabolitos secundarios. Aunque aún no se conocen bien los factores que dan lugar a la producción de metabolitos secundarios, sí parece bien establecida la influencia de los alimentos, específicamente de algunos nutrientes [1].

En el ámbito de las biociencias, existe un gran interés por identificar los diferentes metabolitos, así como descifrar las rutas metabólicas que dan lugar a los mismos. Y de aquí surge la importancia de la **química analítica** ya que facilita los fundamentos y las metodologías que permiten obtener información tanto cualitativa como cuantitativa sobre los metabolitos, y utiliza el conocimiento generado para dar solución a diferentes problemas que se presentan, fundamentalmente relacionados con las ciencias de la salud y los alimentos.

Durante la década de los 90 y como consecuencia de los avances científicos, se produce una revolución en las técnicas y métodos utilizados en el campo de la biología molecular. Debido a esto, se acuñan nuevos conceptos y aparecen nuevas disciplinas entre las que destacan las llamadas **ciencias ómicas**, como la

[1] U. Roessner (2007). "The chemical challenge of the metabolome". En: S.G. Villas-Bôas, U. Roessner, M.A.E. Hansen, J. Smedsgaard, J. Nielsen (Eds.), *Metabolome analysis. An introduction*, John Wiley & Sons, pp. 15-38.

genómica, la proteómica y la metabolómica [2]. Existen dos objetivos principales en estas ciencias: (i) la caracterización completa de los compuestos en un organismo, tejido o célula, es la parte analítica de las ciencias ómicas; (ii) el establecimiento de relaciones entre la presencia del compuesto, caracterizado analíticamente, y su intervención en los procesos vitales [3].

La **metabolómica** se basa en la conjunción de técnicas separativas con técnicas de detección para estudiar el conjunto de los metabolitos que sintetiza un sistema biológico [3,4]. Los metabolitos son los productos finales de los procesos de regulación celular debido a la respuesta a los cambios genéticos y ambientales de los sistemas biológicos. De manera que se elaboran los perfiles de múltiples metabolitos y sus cambios causados por factores como el medio ambiente, la dieta, el estilo de vida, los efectos genéticos y farmacéuticos en todo el organismo [5].

Existen diferentes **tipos de enfoques en el estudio de la metabolómica** [4,6,7]:

- **Análisis dirigido de metabolitos.** Detección y cuantificación de un metabolito específico o un grupo pequeño de ellos.
- **Perfiles metabolómicos o de metabolitos.** Identificación y cuantificación aproximada de un conjunto de metabolitos relacionados, debido a su similitud estructural o funcional, o por estar involucrados en vías metabólicas específicas. Su objetivo es detectar la mayor cantidad posible de metabolitos, independientemente de si son conocidos o no.

-
- [2] J. Nielsen (2007). "Metabolomics in functional genomics and systems biology". En: S.G. Villas-Bôas, U. Roessner, M.A.E. Hansen, J. Smedsgaard, J. Nielsen (Eds.), *Metabolome analysis. An introduction*, John Wiley & Sons, pp. 3-14.
- [3] J.K. Nicholson, E. Holmes, J.C. Lindon (2007). "Metabonomics and metabolomics techniques and their applications in mammalian systems". En: J.C. Lindon, J.K. Nicholson, E. Holmes (Eds), *The handbook of metabonomics and metabolomics*, Elsevier, pp. 1-33.
- [4] K. Dettmer, P.A. Aronov, B.D. Hammock (2007). "Mass spectrometry-based metabolomics", *Mass Spectrometry Reviews*, 51-78.
- [5] O. Fiehn (2002). "Metabolomics – the link between genotypes and phenotypes", *Plant Molecular Biology*, 48, 155-171.
- [6] A. Krastanov (2010). "Metabolomics – The state of art". *Biotechnology & Biotechnological Equipment*, 1537-1543.
- [7] B. Álvarez-Sánchez, F. Priego-Capote, M.D. Luque de Castro (2010). "Metabolomics analysis I. Selection of biological samples and practical aspects preceding sample preparation". *Trends in Analytical Chemistry*, 29, 111-119.

- **"Huella dactilar" metabolómica.** Es un análisis directo de extractos de muestras para su clasificación sin que sea necesaria la identificación de los metabolitos. Se trata de un estudio rápido y de alto rendimiento.
- **Caracterización del metaboloma.** Análisis cualitativo y cuantitativo del conjunto completo de metabolitos que se pueden encontrar en un sistema biológico.

A continuación se ponen en contexto las dos líneas de investigación:

I. Aplicación de la metabolómica en la ciencia de la salud

Los estudios metabolómicos tienen gran relevancia en el campo de la salud. Como muestras de estudio se utilizan biofluidos, y extractos de tejidos o células, ya que proporcionan una visión integrada de los sistemas biológicos de los mamíferos [3]. A través del estudio en células, se puede conocer sistemas celulares tumorales, entre otros aspectos con los que se relaciona la metabolómica en relación con la salud. También se puede obtener información para ensayos clínicos mediante muestras de orina y sangre (suero o plasma), y a través de ellos se puede concluir el diagnóstico de una enfermedad y el seguimiento de los medicamentos que se han recetado en una terapia [7].

A modo de ejemplo, existe gran interés por el estudio de los metabolitos del colesterol presentes en suero y en otros fluidos por estar relacionados con lesiones cardiovasculares y arteriosclerosis [8]. Desde hace tiempo, se sospecha su implicación en los desórdenes lipídicos, y parecen estar involucrados en importantes desórdenes degenerativos, tales como el Alzheimer, osteoporosis, degeneración muscular, etc. [9]. Ciertos esteroides están relacionados con la biosíntesis del colesterol, que por su parte, se metaboliza en oxisteroides de

[8] M. Mulder (2009). "Sterols in the central nervous system", *Current Opinion in Clinical Nutrition and Metabolic Care*, 12, 152-158.

[9] C. Ibáñez, A. Valdés, V. García-Cañas, C. Simó (2014). "Metabolomics in the study of Alzheimer's disease". En: V. García-Cañas, A. Cifuentes, C. Simó (Eds.), *Applications of advanced omics technologies: from genes to metabolites*. *Comprehensive analytical chemistry*. Vol. 64, Elsevier, pp. 249-278.

cadena lateral con la ayuda de enzimas específicas denominadas oxigenasas [10,11].

Así, el **primer capítulo** de esta Tesis Doctoral, se centra en explicar esta línea de investigación. La detección y cuantificación del perfil de metabolitos, se lleva a cabo de forma clásica obteniendo una buena resolución cromatográfica para la identificación y cuantificación de los picos y así expresar los resultados a través del: área del pico característico, concentración, límite de detección, límite de cuantificación, reproducibilidad, repetibilidad, etc.

A continuación y antes de dar paso a la siguiente línea de investigación, el **capítulo dos** trata de explicar los avances en el uso de la "huella dactilar" instrumental a través del tratamiento de datos quimiométrico, sin necesidad de identificar ni cuantificar los metabolitos presentes en las muestras.

La metodología de "**huellas dactilares**" instrumentales, se basa en considerar la señal global suministrada y registrada por un instrumento analítico, p.e.: un cromatograma [12]. Dicho registro refleja de forma implícita la composición química. El tipo de "huellas dactilares" se puede dividir en dos grupos, dependiendo del carácter de la información asociada a ella: (a) **inespecíficas**: obtenidas de una señal en la que no se puede diferenciar la información individual sobre los diferentes componentes; y (b) **específicas**: las señales están lo suficientemente bien resueltas o contienen información explícita, como para poder obtener información de componentes individuales o de grupos formados por compuestos de una misma familia.

Para una aplicación efectiva de las "huellas dactilares" cromatográficas se recurre a la quimiometría. La **quimiometría**, según la definición de la Sociedad Internacional de Quimiometría, es la disciplina química que usa métodos matemáticos y estadísticos para diseñar o seleccionar procedimientos óptimos y

[10] W.J. Griffiths, Y. Wang (2011). "Analysis of oxysterol metabolomes", *Biochimica et Biophysica Acta*, 1811, 784-799.

[11] Y. Wang, W.J. Griffiths (2008). "Steroids, sterols and the nervous system". En: W.J. Griffiths (Ed.), *Metabolomics, metabonomics and metabolite profiling*, RSC Biomolecular Sciences, pp.71-115.

[12] O. Fiehn (2001). "Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks", *Comparative and Functional Genomics*, 2, 155-168.

experimentos y proporcionar la máxima información química mediante el análisis de los datos químicos [13]. Más reciente, la IUPAC establece la definición como la ciencia que relaciona las medidas realizadas en un proceso o sistema químico al estado del sistema a través de la aplicación de métodos matemáticos o estadísticos [14].

Existe un gran número de **herramientas quimiométricas** [15,16]: diseño de experimentos, reconocimiento de pautas, etc. Haciendo uso de las técnicas de reconocimiento de pautas, se puede identificar relaciones y vínculos entre objetos (muestras) o grupos de estas, previamente definidas, que permiten su agrupamiento y clasificación según la similitud de la "huella dactilar" cromatográfica, de sus espectros, imágenes, etc. Tanto la quimiometría clásica como la más reciente derivada de la inteligencia artificial (minería de datos) usan técnicas de reconocimiento de pautas ("pattern recognition", PR) para llevar a cabo el tratamiento de datos multivariados. Este tipo de tratamientos suelen tener como objetivo principal, el clasificar, es decir distribuir unos objetos (materiales) entre una serie de clases, buscando una pauta típica de cada clase de datos. Una clase es un conjunto de objetos que poseen una o varias propiedades comunes.

Las técnicas de reconocimiento de pautas se dividen generalmente en dos grandes grupos que engloban el reconocimiento de pautas supervisadas (sPR) y el reconocimiento de pautas no supervisadas (nsPR) [16,17]. En el primero previamente se conocen las clases, lo que el PR permite la asignación de nuevos objetos a las mismas; y en el segundo grupo, no se conocen las pautas y el PR permitirá conocer si existen agrupaciones entre los objetos.

-
- [13] R.Learidi (2003). "Chemometrics in data analysis", En: M. Lees (Ed.), Food authenticity and traceability, CRC Press, pp.315-336.
- [14] D.B. Hibbert (2016). "Vocabulary of concepts and terms in chemometrics", (IUPAC Recommendations 2016), Pure and Applied Chemistry, 88, 407-443.
- [15] R.G. Brereton (2007). "Applied chemometrics for scientists", Wiley a John Wiley and Sons, Ltd., cap. 2, pp. 9-62.
- [16] R.G. Brereton (2007). "Applied chemometrics for scientists", Wiley a John Wiley and Sons, Ltd., cap. 5, pp. 145-191.
- [17] R.G. Brereton (2009). "Chemometrics for pattern recognition", Wiley a John Wiley and Sons, Ltd., cap. 3, pp.47-106.

II. Aplicación de la metabolómica en la ciencia de los alimentos

En el caso de los productos alimenticios, la composición química determina los aspectos de valor nutricional, sabor, fragancia, apariencia, etc., así como los aspectos negativos, como la toxicidad [18]. La metabolómica permite, por tanto, conocer y comprender todos los detalles moleculares responsables de las propiedades específicas de alimentos y bebidas. Además, la composición de los metabolitos en los alimentos se puede usar para controlar la calidad de los mismos y, por tanto, es posible evitar fraudes debidos a adulteraciones alimentarias que se realizan para reducir el precio de la producción y obtener un beneficio extra [19].

Esta aplicación de metabolómica en alimentos recientemente se ha denominado **alimentómica** (término que proviene de la traducción literal del inglés "foodomics") y también es conocida con el nombre de **nutrimetabolómica** [20,21]. Esta disciplina combina las ciencias ómicas, para elaborar perfiles de compuestos alimenticios, con herramientas como la bioinformática (detección de biomarcadores para la calidad o seguridad alimentaria), ensayos in vitro/in vivo (bioactividad) o ensayos clínicos para determinar los posibles efectos saludables.

A modo de ejemplo, la caracterización analítica de ciertos metabolitos en el aceite de oliva es importante para controlar la calidad, debido a los numerosos fraudes y adulteraciones que se llevan a cabo de este producto en el sector olivarero [22]. La calidad comercial de un alimento es el conjunto de aquellas características o atributos individuales de un producto que son significativos para determinar el grado de aceptación que aprecia el consumidor. En el caso del

-
- [18] E. Ibáñez, A. Cifuentes (2014). "Foodomics: food science and nutrition in the postgenomic era". En: V. García-Cañas, A. Cifuentes, C. Simó (Eds.), *Applications of advanced omics technologies: from genes to metabolites*. Comprehensive analytical chemistry. Vol. 64, Elsevier, pp. 395-440.
 - [19] G.P. Danezis, C.A. Papachristidis, C.A. Georgiou (2016). "Elemental fingerprinting". En: C.A. Georgiou, G.P. Danezis (Eds.), *Food authentication. Management, analysis and regulation*, Wiley Blackwell, pp. 85-116.
 - [20] A. Cifuentes (2009). "Food analysis and foodomics". *Journal of Chromatography A*, 1216, 7109-7110.
 - [21] A. Cifuentes (2013). "Foodomics: principles and applications". En: A. Cifuentes (Ed.), *Foodomics. Advanced mass spectrometry in modern food science and nutrition*, Wiley a John Wiley & Sons, Inc., pp. 1-13.
 - [22] M. Jee (2002). "Adulteration and authentication of oils and fats: an overview". En: M. Jee (Ed.), *Oils and fats authentication*, CRC Press, pp. 1-24.

aceite de oliva, se ha de controlar la calidad y monitorizar su estabilidad para asegurar que se mantienen sus propiedades organolépticas y saludables y, por tanto, la calidad del mismo, en el momento de la venta al público.

La clasificación del aceite de oliva por categorías de calidad (virgen extra, virgen y lampante) se realiza mediante análisis sensorial, además de una serie de parámetros físico-químicos.

Dentro de esta línea de investigación se desarrollan los **capítulos tres y cuatro**, sobre la monitorización de la estabilidad de patrones de aceite de oliva para análisis sensorial a partir de medidas cromatográficas de la fracción de compuestos volátiles, así como la evolución de la madurez de la aceituna y su importancia con respecto a la calidad del aceite de oliva. En ambos capítulos se ha trabajado con toda la información instrumental contenida en la "huella dactilar" cromatográfica de la fracción volátil.

El conjunto de esta Tesis Doctoral, conduce a la obtención de información analítica en sistemas reales, línea de investigación escogida dentro del Programa Oficial de Doctorado en Química y donde se enmarca la misma.

Problema general

En la actualidad existe gran interés en estudiar los metabolitos en muestras de biofluidos (como suero u orina) y de aceites vegetales (especialmente en aceite de oliva) con el objetivo de solventar problemas concretos en los ámbitos de la salud humana y la alimentación.

Así, en el caso del estudio de metabolitos en biofluidos, se pretende relacionar una determinada cantidad de metabolito o relaciones de ellos, con una enfermedad concreta. Esto es en el caso de que exista relación entre los metabolitos de una ruta metabólica y dicha enfermedad en estudio.

Sin embargo, para el caso del estudio de metabolitos en muestras de aceite de oliva, el objetivo no es relacionarlos con los beneficios que puedan presentar en la salud del consumidor, si no relacionarlos con la calidad del mismo que se ofrece al consumidor. Para el estudio de la calidad del aceite de oliva, es necesario conocer los metabolitos presentes en el aceite que sean característicos a una categoría determinada y variedad de aceituna.

En ambos casos se pretende solventar el mismo problema final, que consiste en poder caracterizar los metabolitos presentes en muestras de sangre y de aceite de oliva. Por tanto, se van a desarrollar métodos analíticos que impliquen el uso de la información asociada al gran número de compuestos presentes en este tipo de muestras a través de técnicas separativas, como es la cromatografía de gases, acopladas a diferentes detectores.

Hipótesis

La "huella dactilar" instrumental, en conjunción con los métodos quimiométricos, pueden dar solución a los problemas que se plantean. A través de la "huella dactilar" cromatográfica es posible obtener información sobre propiedades específicas relacionadas con organismos o alimentos, sin que sea necesario resolver, identificar y cuantificar los diferentes metabolitos que aparecen en el cromatograma.

En el ámbito sanitario es posible crear modelos de predicción de enfermedades, siempre y cuando los metabolitos presentes en fluidos biológicos tengan relación con una enfermedad concreta.

En el ámbito alimentario es posible estudiar la calidad del aceite de oliva virgen a partir de la estabilidad de la muestra. Si la "huella dactilar" no sufre modificación a lo largo del tiempo, se podrá considerar que tampoco lo hará el flavor que caracteriza al aceite de oliva en estudio. Por otro lado, es posible identificar compuestos y encontrar marcadores responsables de la madurez de la aceituna a partir de la "huella dactilar" cromatográfica.

En el caso que sea necesario cuantificar metabolitos concretos de dicha "huella dactilar", se recurrirá a la optimización del método de análisis para obtener una mayor resolución analítica y cuantificar los metabolitos de interés.

Objetivos

La Tesis Doctoral se centra en la caracterización y/o determinación de metabolitos relacionados con dos líneas de investigación: (i) ciencias de la salud y (ii) ciencias de los alimentos; ambas basadas en el uso de "huellas dactilares" cromatográficas para llevar a cabo el tratamiento de datos quimiométrico.

El principal objetivo es desarrollar metodologías analíticas innovadoras que conlleven el empleo de diferentes técnicas separativas como son la cromatografía de gases o cromatografía de gases bidimensional, acopladas a detectores de espectrometría de masas, con diferentes tipos de analizadores, así como detectores de ionización de llama. A través de dichas técnicas es posible obtener información analítica útil para resolver problemas relacionados con las dos líneas de investigación en estudio:

(I) Línea de investigación en el ámbito de las ciencias de la salud.

Esta línea de investigación se centra en muestras de suero y trata de estudiar muestras en pacientes sanos y enfermos que relacionen una enfermedad concreta con la concentración de diferentes metabolitos.

(II) Línea de investigación en el ámbito de las ciencias de los alimentos.

Esta línea de investigación se centra en muestras de aceites de oliva y trata de estudiar la calidad del aceite de oliva a partir de los compuestos orgánicos volátiles. Para ello surge la necesidad de monitorizar la estabilidad en patrones de aceite de oliva para análisis sensorial, así como estudiar la evolución de la madurez de la aceituna mediante la selección de marcadores.

Los objetivos específicos se presentan en cada uno de los capítulos donde se desarrolla la investigación llevada a cabo.

Artículos científicos (por fecha de publicación, aceptación o envío):

- ❖ L. Cuadros-Rodríguez, A. González-Casado, L. Valverde-Som, E. Pérez Castaño, C. Ruiz-Samblás, (2016), **Chromatographic fingerprinting: an innovative approach for food 'identification' and food authentication – A tutorial**, *Analytica Chimica Acta*, 909, 09-23 (DOI: 10.1016/j.aca.2015.12.042).
- ❖ F. Magagna, L. Valverde-Som, C. Ruiz-Samblás, L. Cuadros-Rodríguez, S.E. Reichenbach, C. Bicchi, C. Cordero (2016), **Combined untargeted and targeted fingerprinting with comprehensive two-dimensional chromatography for volátiles and ripening indicators in olive oil**, *Analytica Chimica Acta*, 936, 245-258 (DOI: 10.1016/j.aca.2016.07.005).
- ❖ L. Valverde-Som, A. Carrasco-Pancorbo, S. Sierra, S. Santana, C. Ruiz-Samblás, N. Navas, J.S. Burgos, L. Cuadros-Rodríguez (enviado, 22 Septiembre 2017), **Separation and determination of some of the main cholesterol-related compounds in blood by gas chromatography-mass spectrometry (selective ion monitoring mode)**, *Separations*.
- ❖ L. Valverde-Som, C. Ruiz-Samblás, F.P. Rodríguez-García, L. Cuadros-Rodríguez (enviado, 13 Noviembre 2017), **Multivariate approaches for chemical chromatographic control of the stability of the organoleptic quality of olive oil reference materials for sensory analysis – Part I: Framework and fundamentals**, *Journal of the Science of Food and Agriculture*.
- ❖ L. Valverde-Som, C. Ruiz-Samblás, F.P. Rodríguez-García, L. Cuadros-Rodríguez (enviado, 13 Noviembre 2017), **Multivariate approaches for chemical chromatographic control of the stability of the organoleptic quality of olive oil reference materials for sensory analysis – Part II: Applications**, *Journal of the Science of Food and Agriculture*.

Otros trabajos no incluidos en la Tesis Doctoral:

Artículos científicos (por fecha de publicación, aceptación o envío)

- ❖ V. Olivares-Arias, L. Valverde-Som, V. Quirós-Rodríguez, R. García-Romero, N. Muñoz, M. Navarro-Alarcón, C. Carbera-Vique (2015), **Nickel in food and influencing factors in its level, intake, bioavailability and toxicity: a review**, CyTA – Journal Food, 87-101 (DOI: 10.1080/19476337.2014.917383).
- ❖ E. Pérez-Castaño, C. Ruiz-Samblás, S. Medina-Rodríguez, V. Quirós-Rodríguez, A.M. Jiménez-Carvelo, L. Valverde-Som, A. González-Casado, L. Cuadros-Rodríguez (2015), **Comparison of different analytical classification scenarios: application for the geographical origin of edible palm oil by sterolic (NP)HPLC fingerprinting**, Analytical Method, 7, 4192-4201 (DOI: 10.1039/C5AY00168D).
- ❖ M. Delgado-Aguilar, L. Valverde-Som, L. Cuadros-Rodríguez (enviado, 11 Septiembre 2017), **Solver, an Excel application to solve the difficulty in applying different regression methods**, Chemometrics and Intelligent Laboratory Systems.

Proyectos de investigación (por fecha de elaboración)

- ❖ "Certificación y distribución de doce materiales de referencia certificados (MRC) para análisis organoléptico de aceite de oliva (SensOLEO-MRC 2013)". Proyecto financiado por la Consejería de Agricultura, Pesca y Medio Ambiente (CAPMA), Junta de Andalucía (Ref.: OTRI 3315-2013).
- ❖ Proyecto INSTRUMENTAL: "Estudio de las diferentes tecnologías analíticas actuales aptas para completar o construir una alternativa futura al método comunitario de análisis sensorial denominado 'panel test'". Proyecto financiado por la Organización Interprofesional del Aceite de Oliva Español, (Ref.: OTRI 3501-2014).
- ❖ "Desarrollo de aplicaciones analíticas de la plataforma UHPLC-(Orbitrap)MS en los ámbitos de la metabolómica y los biofármacos". Convenio de colaboración financiado por Thermo-Fisher Scientific S.L.U., (Ref.: OTRI 3503-2014).
- ❖ "Análisis de los datos experimentales recopilados durante el desarrollo del proyecto: Estudio de la evolución de los ésteres alquílicos en el aceite de oliva virgen extra". Proyecto financiado por la Organización Interprofesional del Aceite de Oliva Español, (Ref.: OTRI 3501-02-2015).
- ❖ "Desarrollo de un sistema de gestión de la calidad para el funcionamiento de un panel analítico de cata científica para el análisis sensorial de los alimentos a ser certificados, que reúna los requisitos establecidos en la norma UNE-EN ISO/IEC 17025". Proyecto financiado por Fundación Qvalytech Alimentación, (Ref.: OTRI 3680-2015).
- ❖ Proyecto InterOLEO-MRC 2016: "Certificación y distribución de cinco materiales de referencia certificados (MRC) para análisis físico-químicos de aceite de oliva". Proyecto financiado por la Consejería de Agricultura, Pesca y Desarrollo Rural (CAPDER), Junta de Andalucía, (Ref.: OTRI 3755-2016).



CAPÍTULO 1

Determinación de precursores y derivados del colesterol en muestras de sangre



CAPÍTULO 1

1.1. PRESENTACIÓN

Este capítulo recoge los resultados obtenidos del estudio dedicado a la **separación, identificación y cuantificación de los principales compuestos de la ruta metabólica del colesterol**, presentes en la circulación sanguínea, y que están relacionados con algunas enfermedades características de desórdenes degenerativos como la enfermedad de Alzheimer.

El proyecto en el que se desarrolló este capítulo, se denomina: "Potencial de diferentes plataformas GC-MS de aplicación en metabolómica: dos casos de estudio" financiado por el Campus de Excelencia Internacional de Granada (CEI BioTic Granada). El proyecto estaba dividido en dos actividades, que fueron desarrolladas en colaboración por diferentes investigadores: (i) Departamento de Química Analítica, Universidad de Granada; (ii) Departamento de Pediatría, Universidad de Granada; (iii) Leiden University Medial Center (LUMC), Universidad de Leiden; (iv) Imperial Collage, Universidad de Londres; (v) Neuron Bio, empresa I+D de Granada.

La actividad uno bajo el título "Análisis exploratorio de orina humana tras ingesta de aceite de oliva mediante GC-MS para estudiar el metabolismo de compuestos fenólicos" se desarrolló en otro grupo de investigación y no tiene cabida en esta Tesis. Y la actividad dos, fue la desarrollada por el grupo AnAMA, bajo el título "Determinación de colesterol y sus metabolitos hidroxilados en muestras de plasma", que permitió adquirir las muestras utilizadas para este estudio.

Las muestras fueron suministradas por la empresa Neuron Bio y el **objetivo** de la investigación fue optimizar y desarrollar un método analítico para la detección y cuantificación de precursores y derivados del colesterol en muestras de plasma de ratón.

1.2. INTRODUCCIÓN

Colesterol

El **colesterol** es un esteroide abundante en mamíferos y es producto y precursor de moléculas relevantes presentes en la circulación sanguínea y el sistema nervioso, tales como esteroides hormonales, ácidos biliares, oxisteroles, e intermedios en la vía biosintética del colesterol [1,2]. El colesterol también está disponible a través de la dieta.

La figura 1.1 muestra la **estructura del colesterol** indicando el número de carbonos para conocer la posición de los compuestos resultantes de él y su nomenclatura.

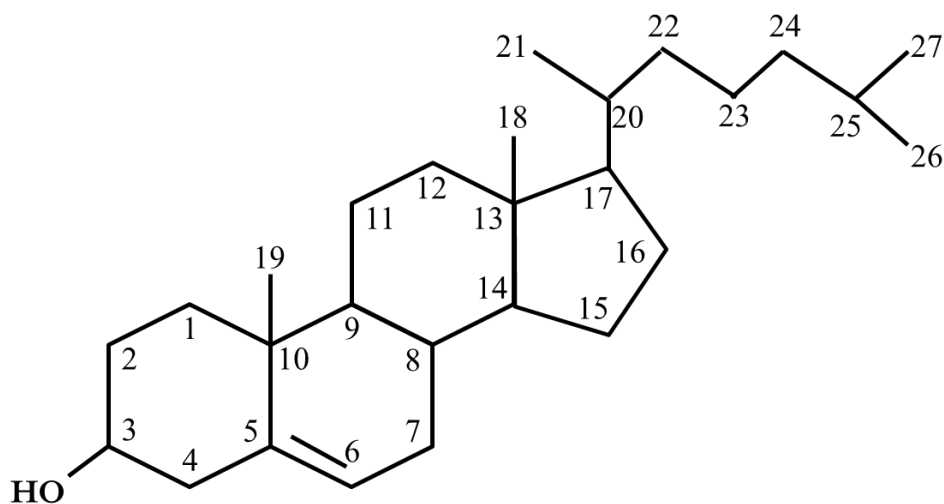


Figura 1.1. Estructura del colesterol, también llamado (3 β)-colest-5-en-3-ol.

- [1] J.G. McDonald, D.D. Smith, A.R. Stiles, D.W. Rusell (2012). "A comprehensive method for extraction and quantitative analysis of sterols and steroids from human plasma", *Journal of Lipid Research*, 53, 1399-1409.
- [2] C. Garenc, P. Julien, E. Levy (2010). "Oxysterols in biological systems: the gastrointestinal tract, liver, vascular Wall and central nervous system", *Free Radical Research*, 44, 47-73.

En los seres humanos, el colesterol puede ser sintetizado por la mayoría de las células. El hígado es el órgano principal en el que se produce [3].

La **ruta de síntesis de colesterol** es una de las más largas y complejas del metabolismo. En el organismo, el colesterol deriva de la biosíntesis de *novo*. Esta vía es en la que los nucleótidos se sintetizan a partir de precursores de bajo peso molecular [3].

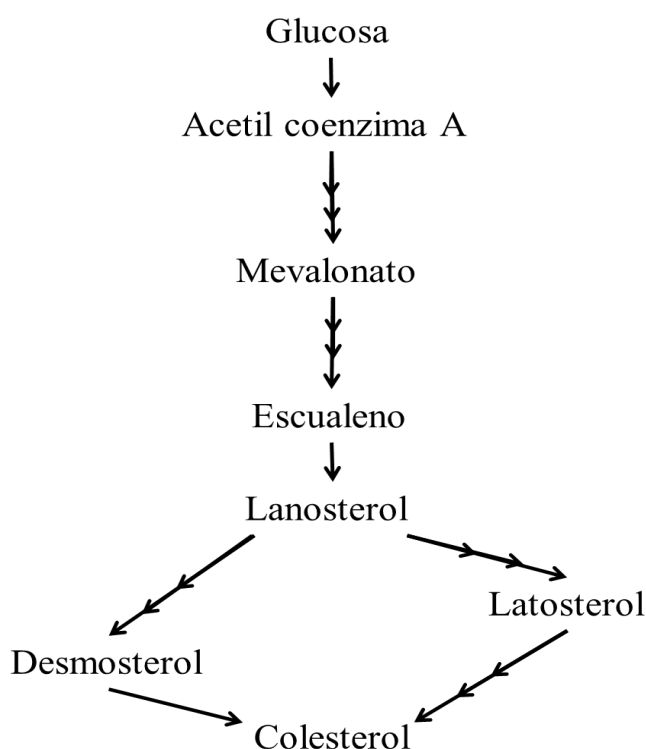


Figura 1.2. Esquema de la ruta de síntesis del colesterol.

La figura 1.2 muestra de forma esquemática la ruta de síntesis del colesterol, donde las tres flechas indican que existen compuestos intermedios que se han omitido y cada flecha indica una o varias reacciones enzimáticas, en el caso en el que se muestren una o tres flechas, respectivamente. Este esquema es un

[3] Y. Wang, W.J. Griffiths (2008). "Steroids, sterols and the nervous system". En: W.J. Griffiths (Ed.), *Metabolomics, metabonomics and metabolite profiling*, RSC Biomolecular Sciences, pp.71-115.

resumen de la información extraída de bibliografía [3,4,5,6] con el objetivo de centrarnos en los metabolitos de interés.

Esta ruta muestra los compuestos precursores del colesterol que pueden cobrar importancia en su relación con una enfermedad determinada y por tanto establecerse como posibles marcadores.

Metabolitos – Oxiesteroles

A partir del colesterol se producen compuestos **derivados oxidados**, que se denominan **oxiesteroles**. Estos se generan enzimáticamente, por autooxidación, o por reacción con ozono.

En analogía a estos, los neuroesteroles son esteroides y oxiesteroides sintetizados o encontrados en el sistema nervioso [7]. Estos oxiesteroides están implicados en procesos inflamatorios asociados con el desarrollo de lesiones como la aterosclerosis y en el mal repliegue de proteínas como ocurre en enfermedades neurodegenerativas.

Los oxiesteroides desempeñan un papel crucial como: (i) reguladores de la expresión de genes implicados en la biosíntesis de lípidos y esteroides; (ii) sustratos para la formación de ácidos biliares; y (iii) mediadores del transporte inverso del colesterol, a través de los cuales el exceso de colesterol se

-
- [4] M.J. Gouveia, P.J. Brindley, L.L. Santos, J.M.C. da Costa, P. Gomes, N. Vale (2013). "Mass spectrometry techniques in the survey of steroid metabolites as potential disease biomarkers: A review", *Metabolism clinical and experimental*, 62, 1206-1217.
 - [5] H. Kölsch, R. Heun, F. Jessen, J. Popp, F. Hentschel, W. Maier, D. Lütjohann (2010). "Alterations of cholesterol precursor levels in Alzheimer's disease", *Biochimica et Biophysica Acta*, 945-950.
 - [6] Y. Wang, K.M. Sousa, K. Bodin, S. Theofilopoulos, P. Sacchetti, M. Hornshaw, G. Woffendin, K. Karu, J. Sjövall, E. Arenas, W.J. Griffiths (2009). "Targeted lipidomic analysis of oxysterols in the embryonic central nervous system", *Molecular BioSystems*, 5, 529-541.
 - [7] C. Ibáñez, A. Valdés, V. García-Cañas, C. Simó (2014). "Metabolomics in the study of Alzheimer's disease". En: V. García-Cañas, A. Cifuentes, C. Simó (Eds.), *Applications of advanced omics technologies: from genes to metabolites*. Comprehensive analytical chemistry. Vol. 64, Elsevier, pp. 249-278.

disuelve al hígado para la excreción [8,9]. Sin embargo, la sobreabundancia de los oxisteroles puede causar diferentes enfermedades, tales como problemas cardiovasculares, degeneración de retina, enfermedad inflamatoria intestinal, aterosclerosis y trastornos neurodegenerativos [10,11,12].

Estos metabolitos se encuentran presentes en sangre (suero o plasma), líquido cefalorraquídeo y tejido cerebral [13].

El colesterol es transportado al cerebro por la apolipoproteína E. En condiciones normales no se produce ningún intercambio de colesterol sobre la barrera hematoencefálica [12,14]. Sin embargo, se puede dar este intercambio (del cerebro a la sangre) a partir del oxisterol de cadena lateral 24(S)-hidroxicolesterol, que se genera a través de la 24(S)-hidroxilasa, denominada CYP46, y que se encuentra solo en células neuronales [15]. Para conseguir la **homeostasis del colesterol en el cerebro**, se produce un **intercambio de la misma cantidad de 24(S)- y 27-hidroxicolesterol entre el cerebro y la sangre**, como se muestra en la figura 1.3.

-
- [8] C. Garenc, P. Julien, E. Levi (2010). "Oxysterols in biological systems: the gastrointestinal tract, liver, vascular Wall and central nervous system", *Free Radical Research*, 44, 47-73.
- [9] I. Björkhem, M. Heverin, V. Leoni, S. Meaney, U. Diczfalusy (2006). "Oxysterols and Alzheimer's disease", *Acta neurologica Scandinavica Supplementum*, 185, 43-49.
- [10] J. Terao (2014). "Cholesterol hydroperoxides and their degradation mechanism". En: Y. Kato (Ed.), *Lipid hydroperoxide-derived modification of biomolecules*, Springer, pp. 83-91.
- [11] G. Poli, F. Biasi, G. Leonarduzzi (2013). "Oxysterol in the pathogenesis of major chronic diseases", *Redox Biology*, 1, 125-130.
- [12] M. Mulder (2009). "Sterols in the central nervous system", *Current Opinion in Clinical Nutrition and Metabolic Care*, 12, 152-158.
- [13] L. Baila-Rueda, A. Cenarro, M. Cofán, I. Orera, S. Barcelo-Batllo, M. Pocoví, E. Ros, F. Civeira, C. Nerín, C. Domeño (2013). "Simultaneous determination of oxysterols, phytosterols and cholesterol precursors by high performance liquid chromatography tandem mass spectrometry in human serum", *Analytical Methods*, 5, 2249-2257.
- [14] M. Heverin, N. Bogdanovic, D. Lütjohann, T. Bayer, I. Pikuleva, L. Bretillon, U. Diczfalusy, B. Winblad, I. Björkhem (2004). "Changes in the levels of cerebral and extracerebral sterols in the brain patients with Alzheimer's disease", *ASBMB Journal of Lipid Research*, 45, 186-193.
- [15] I. Björkhem (2013). "Five decades with oxysterols", *Biochimie*, 95, 448-454.

El colesterol que existe en el cerebro es independiente de su ingesta, puesto que es sintetizado por la biosíntesis del *novo* en el propio cerebro. Este balance entre el cerebro y la sangre de los oxisteroles 24(S)- y 27-hidroxicolesterol, es un proceso importante y su mal funcionamiento (como la mayor permeabilidad de la barrera hematoencefálica) puede influir en enfermedades neurodegenerativas [14,15].

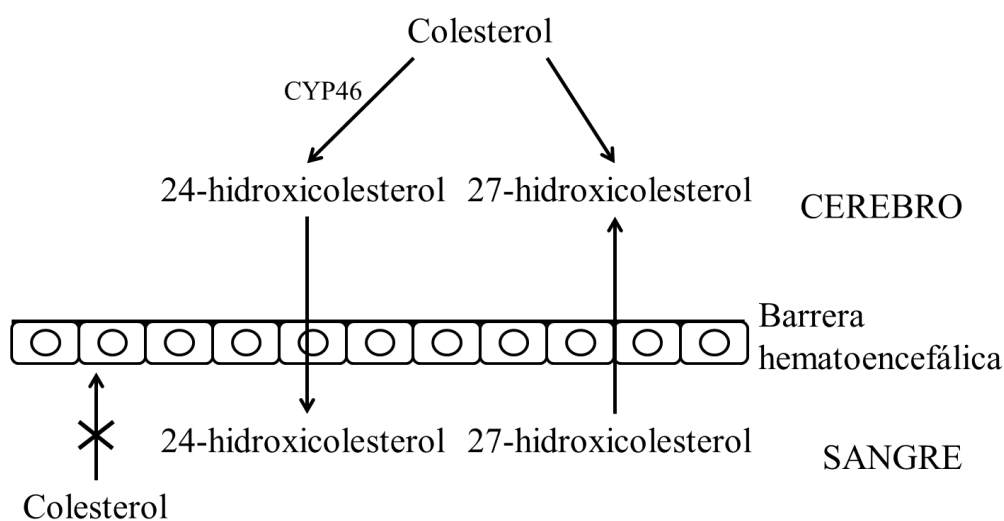


Figura 1.3. Esquema del transporte oxisteroles de cadena lateral del cerebro a la sangre y viceversa (figura adaptada [16]).

Enfermedad de Alzheimer

Una de las enfermedades neurodegenerativas que se pueden producir por una desregulación del metabolismo del colesterol (debido al transporte inadecuado del colesterol a través de diferentes oxisteroles entre el cerebro y la sangre), es la **enfermedad de Alzheimer** [14,17].

La enfermedad de Alzheimer afecta al sistema nervioso central, que lleva a una destrucción progresiva y atrofia de la corteza central, especialmente a regiones relacionadas con funciones mentales superiores. A medida que la

[16] W.Y. Ong, J.H. Kim, X. He, P. Chen, A.A. Farooqui, A.M. Jenner (2010). "Changes in brain cholesterol metabolome after excitotoxicity", *Mol Neurobiol*, 41, 299-313.

[17] E. Trushina, M.M. Mielke (2014). "Recent advances in the application of metabolomics to Alzheimer's disease", *Biochimica et Biophysica Acta-Molecular Basic of Disease*, 1842, 1232-1239.

patología avanza, se ven afectadas las capacidades ordinarias como la memoria y la atención conduciendo a un síndrome de demencia [7].

Este capítulo de Tesis se centra en el **estudio de posibles marcadores en sangre de dicha enfermedad**, por ser uno de los objetivos empresariales en la empresa Neuron Bio. Esta empresa se dedica a la búsqueda de fármacos para la prevención y el tratamiento de enfermedades neurodegenerativas, y especialmente frente a la enfermedad de Alzheimer.

El objetivo de usar **muestras de sangre**, es encontrar un **método no invasivo** para la detección de dicha enfermedad.

Instrumentación

La caracterización de los metabolitos en el organismo, tejidos, o células requieren de técnicas analíticas avanzadas para su separación y detección. Los oxisteroles se han determinado por cromatografía de líquidos y cromatografía de gases, normalmente acoplados a detectores de espectrometría de masas y también por resonancia magnética nuclear [18,19]. El espectrómetro de masas como detector proporciona un método ideal para el perfil de esteroides [7].

La técnica analítica seleccionada para la determinación de dichos analitos es la **cromatografía de gases acoplada con espectrometría de masas con impacto electrónico como sistema de ionización y analizador de trampa de iones como analizador de masas (GC-(EI/IT)MS)** (ver figura 1.4).

El equipo que se muestra en la imagen es un cromatógrafo de gases VARIAN GC 3800 equipado con un inyector split/splitless acoplado a un espectrómetro

[18] W.J. Griffiths, Y. Wang (2011), "Analysis of oxysterols metabolomes", *Biochimica et Biophysica Acta*, 1811, 784-799.

[19] M.J. Gouveia, P.J. Brindley, L.L. Santos, J.M. Correia da Costa, P. Gomes, N. Vale (2013), "Mass spectrometry techniques in the survey of steroid metabolites as potential disease biomarkers: A review", *Metabolism*, 62, 1206-1217.

de masas de trampa de iones VARIAN (IT)MS 4000, equipado con una fuente de impacto electrónico. Así como un automuestreador robotizado CombiPal CTC ANALYTICS.



Figura 1.4. Equipo de cromatografía de gases acoplado a un espectrómetro de masas con analizador de trampa de iones, usado para el análisis de las muestras.

Un **espectrómetro de masas**, presenta varias partes: una fuente de ionización, un analizador de masas, un detector y un sistema de obtención de datos [20].

Las moléculas de las muestras son ionizadas en la **fente de ionización**, en este caso de **impacto electrónico**. El impacto electrónico implica el bombardeo de las moléculas de una muestra con electrones de alta energía, normalmente 70eV, como resultado se generan iones, normalmente cationes, y electrones libres debido a la energía de impacto. En muchos casos los iones moleculares son inestables y se fragmentan dando productos más estables. Esta técnica de ionización funciona bien para muchas moléculas en fase gaseosa, pero introduce un inconveniente, ya que la fragmentación puede ser tal que no se observe el ión molecular.

Los iones generados son separados en función de la relación masa-carga (m/z) en el analizador y a continuación son detectados. En este caso el **analizador**

[20] Y. Wang, W.J. Griffiths (2008). "Mass spectrometry for metabolite identification". En: W.J. Griffiths (Ed.), *Metabolomics, metabonomics and metabolite profiling*, RSC Biomolecular Sciences, pp.71-115.

es de **trampa de iones**. La trampa de iones está formada por tres electrodos hiperbólicos: un electrodo con forma de anillo y dos electrodos por encima y por debajo del anillo. En la figura 1.5 se observa el analizador del equipo.

Los iones son retenidos o expulsados de la trampa en función de su relación m/z , cuando todos los iones salen de la trampa, pasan al **detector** y la señal es transmitida a un sistema de **recolección de datos** dando lugar al espectro de masas.

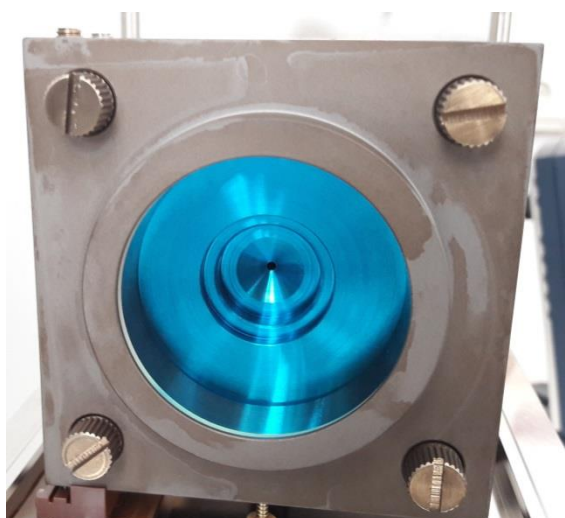


Figura 1.5. Trampa de iones del equipo VARIAN (IT)MS 4000.

Los **modos de adquisición** del espectro son *full scan* (o de barrido completo), SIM (*single ion monitoring*) y MS/MS. El modo *full scan* consiste en medir todo el espectro de masas completo comprendido entre dos límites de relación m/z que se seleccionan a priori en el método. El modo **SIM** consiste en enfocar, por segmentos seleccionados (con un determinado tiempo de retención), uno o varios iones de interés sobre un único analito o varios. El modo **MS/MS** consiste en enfocar un ión concreto y romperlo en iones más pequeños para confirmar que se trata de la molécula de interés, también se hace por segmentos.

1.3. MUESTRAS

La empresa Neuron Bio proporcionó **dos grupos de muestras**: (i) **plasma de ratón**, con el objetivo de detectar, identificar y cuantificar en ellas los analitos de interés en tres grupos de ratones; así como (ii) **suero de cabra**, como blanco, que fue dopado a diferentes concentraciones de una mezcla de patrones que contenían los analitos de interés.

Los **analitos** de interés que se seleccionaron como **posibles biomarcadores** de la **enfermedad de Alzheimer** fueron: (i) lanosterol, latosterol, desmosterol, como precursores del colesterol; (ii) colesterol; (iii) 7-ketocolesterol, 7 α -hidroxicolesterol, 7 β -hidroxicolesterol, 24(S)-hidroxicolesterol, 25-hidroxicolesterol, 27-hidroxicolesterol, como oxisteroles del colesterol. La figura 1.6 muestra la estructura de cada uno de esos compuestos.

Muestras de plasma de ratón

Son tres los grupos de ratones que se seleccionaron: seis ratones con 9 meses de edad, transgénicos, que son ratones en los que se le ha modificado algún gen genéticamente. Estos fueron divididos en dos grupos, tres de ellos que a los que no se les realizó ningún tratamiento y se nombraron con la codificación APP^{swe}; los tres ratones restantes, alimentados con una dieta enriquecida en estatina, desarrollado por la empresa Neuron Bio [21], y codificados con el nombre NST0037. Además otros tres ratones con 9 meses de edad no transgénicos, con la codificación Non-Tg.

[21] S. Campoy, S. Sierra, B. Suarez, M.C. Ramos, J. Velasco, J.S. Burgos, J.L. Adrio (2010). "Semisynthesis of novel monacolin J derivatives:hypocholesterolemic and neuroprotective activities", *Journal of Antibiotics*, 63, 499-505.

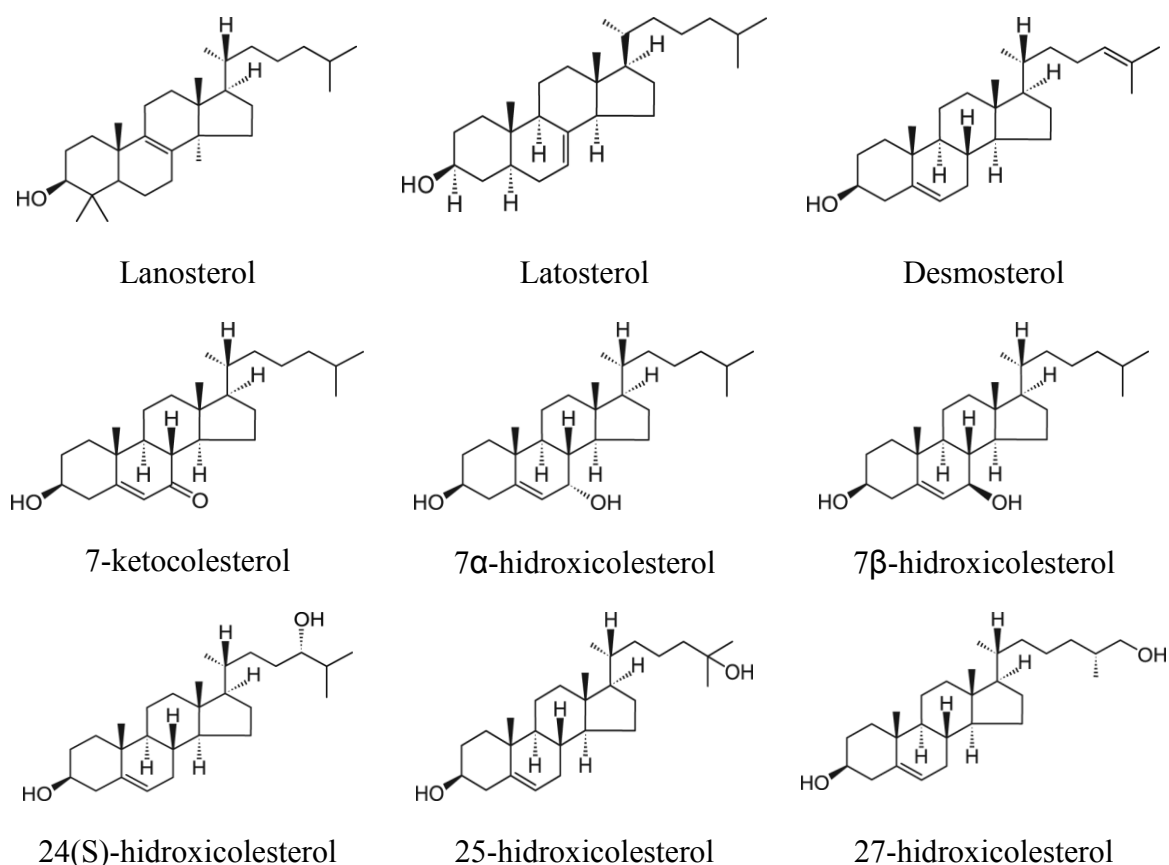


Figura 1.6. Estructura química de los metabolitos de interés.

Muestras de suero de cabra

El suero de cabra fue utilizado como blanco, al que fue dopado con diferentes concentraciones (0.5, 1.0, 2.5 y 5.0 $\mu\text{L}/\text{mL}$) de una mezcla de patrones que contenían los analitos de interés. En este caso se hicieron dos grupos: (i) un primer grupo en el que se dopaba el suero de cabra antes del proceso de extracción; (ii) un segundo grupo de suero de cabra dopado tras el proceso de extracción. El objetivo era estudiar el efecto matriz y la recuperación obtenida tras el proceso de extracción.

1.4. PROCESO EXPERIMENTAL

Las muestras de suero y plasma sanguíneo requieren de un proceso de extracción que permita eliminar las proteínas. Este proceso de extracción fue llevado a cabo en la empresa Neuron Bio bajo un procedimiento interno, que queda esquematizado en el apartado siguiente donde aparece el artículo científico resultado de esta investigación.

Los metabolitos de interés presentan baja volatilidad, por lo que es necesaria una **reacción de derivatización**, con el objetivo de poder ser analizados en el cromatógrafo de gases. Como agente derivatizante se selecciona el **BSTFA + TMCS** (N,O-bis(trimethylsilyl)trifluoroacetamide y trimethylchlorosilane), que reemplaza un hidrógeno activo por el trimetilsilil, de manera que los compuestos se vuelven más volátiles y menos polares. El BSTFA causa menos interferencias cromatográficas que otros agentes silanizantes [22]. Se selecciona este agente derivatizante junto con TMCS porque además de silanizar, cataliza la reacción de derivatización. De esta manera se logra la vaporización de la muestra sin descomposición térmica.

Para llevar a cabo la reacción de derivatización, se evapora todo el disolvente de la muestra, en este caso acetato de etilo, con nitrógeno, con un equipo como el que se muestra en la figura 1.7.

Cuando el disolvente es eliminado (figura 1.8) se añade el mismo volumen de agente derivatizante, y se deja reaccionar durante una hora a temperatura ambiente.

Las muestras utilizadas son muestras que poseen una matriz compleja, que puede calcinarse en el inyector y obstruir la columna cromatográfica, por lo

[22] D. Saraiva, R. Semedo, M.C. Castilho, J.M. Silva, F. Ramos (2011). "Selection of the derivatization reagent – The case of human blood cholesterol, its precursors and phytosterols GC-MS analyses", *Journal of Chromatography B*, 879, 3806-3811.

que en el equipo se usa un **liner de vidrio con malla de lana**, para que se queden los restos de la calcinación en la malla y evitar que pase suciedad a la columna cromatográfica. La problemática que existe en este tipo de muestras, es que el colesterol en sangre aparece a una elevada concentración con respecto a los metabolitos de interés para su selección como posibles marcadores de la enfermedad de Alzheimer. Por tanto, se plantea la necesidad de recurrir al modo SIM para identificar y cuantificar de forma más clara los analitos de interés.



Figura 1.7. Evaporación con nitrógeno del acetato de etilo en muestras.

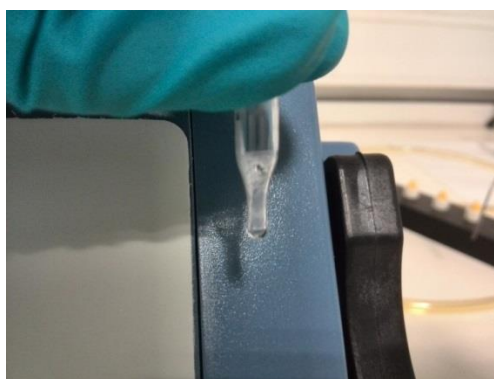


Figura 1.8. Evaporación con nitrógeno del acetato de etilo en muestras.

Cada uno de los pasos llevados a cabo para la optimización del método cromatográfico, se describe en el apartado siguiente de artículo científico

(apartado 3.1 *optimization of chromatographic and detection conditions* del artículo científico). Como innovación con este tipo de muestras se introduce una **pre-columna** con sílica desactivada (1 m × 0.25 mm i.d.) con conexión press-fit, así como una **columna capilar** con 5% difenil-95% dimetil polisiloxano de fase estacionaria, DB-5MS (30 m × 0.25 mm i.d. × 0.25 μm).

A través de la fase estacionaria, la muestra, volatilizada en el inyector, pasa por la pre-columna y la columna cromatográfica, donde **los compuestos son separados** en función de su polaridad quedando o no retenidos en la fase estacionaria. Los analitos de interés pasan al **espectrómetro de masas** donde los analitos son detectados en relación a su m/z.

En las tablas 1.1, 1.2 y 1.3 se muestran las condiciones del método analítico utilizado para la obtención de los cromatogramas.

Tabla 1.1. Condiciones del cromatógrafo de gases.

Gas portador	Helio a 1.2 mL/min		
Modo inyección	Splitless		
Temperatura inyección	250 °C		
Volumen inyección	1 μL		
Pre-columna	Silica desactivada (1 m × 0.25 mm i.d.)		
Columna cromatográfica	DB-5MS (30 m × 0.25 mm i.d. × 0.25 μm), 5% difenil-95% dimetil polisiloxano		
Rampa de temperatura GC	Temperatura (°C)	Rampa (°C/min)	Tiempo* (min)
	200	0.00	0.50
	300	20.00	10.00
	Tiempo de análisis		15.50

* Tiempo que se mantiene la temperatura final a la que llega el gradiente.

Tabla 1.2. Condiciones del espectrómetro de masas.

Sistema de ionización	Impacto electrónico, 70 eV
Modo Scan	SIM
Temperatura línea de transferencia	280 °C
Temperatura trampa de iones	250 °C

Tabla 1.3. Condiciones de cada uno de los segmentos seleccionados en modo SIM en el espectrómetro de masas.

Segmento	Metabolito diana	Tiempo de retención (min)	Iones seleccionados (m/z)
1	7 α -hidroxicolesterol	4.00 – 8.98	456 – 458
2	Desmosterol	8.98 – 9.64	159, 253, 351
3	Latosterol y 7 β -hidroxicolesterol	9.64 – 10.20	213, 255, 353, 443, 456 – 458
4	Lanosterol	10.20 – 11.31	395 – 396
5	24(S)-hidroxicolesterol	11.31 – 11.99	255, 323, 413, 441
6	25-hidroxicolesterol	11.99 – 12.25	131
7	7-ketocolesterol	12.25 – 12.70	367, 382, 472, 545
8	27-hidroxicolesterol	12.70 – 13.60	161, 255, 417, 456, 546
9	Betulina (P.I.)	13.60 – 15.50	189, 203

El resultado de esta investigación dio lugar a la escritura y envío de un artículo científico que se presenta a continuación.

1.5. ARTÍCULO

Separation and determination of some of the main cholesterol-related compounds in blood by gas chromatography-mass spectrometry (selective ion monitoring mode)

(Artículo enviado a revision en Septiembre 2017 a la revista: Separations)



1 *Type of the Paper (Article)*

2 **Separation and determination of some of the main**
3 **cholesterol-related compounds in blood by gas**
4 **chromatography-mass spectrometry (selective ion**
5 **monitoring mode)**

6 Lucia Valverde-Som ^{1,*}, Alegría Carrasco-Pancorbo ¹, Saleta Sierra ², Soraya Santana ², Cristina
7 Ruiz-Samblás ¹, Natalia Navas ¹, Javier S. Burgos ² and Luis Cuadros-Rodríguez ¹

8 ¹ Department of Analytical Chemistry, Faculty of Science, University of Granada, C/ Fuentenueva s/n,
9 E-18071, Granada, Spain

10 ² Neuron Bio, P.T.S. Granada, C/ Avicena 4, E-18016, Granada, Spain

11 * Correspondence: luciavs@ugr.es; Tel.: +34-958-24-0797

Abstract

Oxysterols are metabolites produced in the first step of cholesterol metabolism which is related to neurodegenerative disorder and can be detected by several biological fluid analysis as blood, plasma, serum, or cerebrospinal fluid. In this study, some cholesterol precursors and oxysterols have been analysed by gas chromatography coupled to mass

spectrometry. The selected cholesterol-related compounds were desmosterol, lathosterol, lanosterol, 7 α -hydroxycholesterol, 7 β -hydroxycholesterol, 24(S)-hydroxycholesterol, 25-hydroxycholesterol, 7-ketocholesterol, 27-hydroxycholesterol. These relevant cholesterol-related compounds show similar structure and the interfering effect produced by high concentrations of cholesterol was avoided by applying selected ion monitoring strategy (SIM). The method has been developed an optimised, validated and different analytical performance characteristics, such as linearity index, detection and quantification limits, selectivity and matrix effect, precision (repeatability) and trueness (recovery factor) for each cholesterol-related compounds have been evaluated. The mice's plasma was chosen to determine as representative commodity matrix to apply the proposed methods for the determination of these compounds in real samples.

Keywords

Cholesterol-related compounds, GC-(IT)MS, selective ion monitoring mode.

1. Introduction

Oxysterols play a crucial role as: (a) regulators of expression of genes involved in lipid and sterol biosynthesis; (b) substrates for the formation of bile acids; and (c) mediators of reverse cholesterol

transport, through which the excess of cholesterol is returned to the liver for excretion [1,2]. However, oxysterols' overabundance could cause different diseases, such as cardiovascular problems, retina degeneration, inflammatory bowel disease, atherosclerosis and neurodegenerative disorders [3,4,5].

Oxysterols can be found in several chemical forms and they are enzymatically produced in the first steps of cholesterol (Chol) metabolism [6,7]. For example, Alzheimer's disease (AD) pathology is initiated (or accelerated) by a dysregulation of the metabolism of Chol, sphingolipids, and fatty acids, hypercholesterolemia is an important risk factor for AD [8,9]. These metabolites are present in blood (serum or plasma), cerebrospinal fluid, and brain tissue. Desmosterol, lanosterol and lathosterol are the precursors of the Chol synthesis. Chol is metabolized into side chain oxysterols -24(S)-hydroxycholesterol, 25-hydroxycholesterol and 27-hydroxycholesterol- during its hepatic conversion [10]. Chol is transported into the brain by apolipoprotein E; under normal conditions, no exchange of cholesterol occurs over the blood-brain barrier. However, it can be transported from the human brain into the form of side-chain oxidized oxysterol, as, for instance, 24(S)-hydroxycholesterol and 27-hydroxycholesterol [5,11,12]. As a consequence, the presence of these oxysterols has been related to neurological disorders and they have proposed as potential AD biomarkers [9,13].

Over the last decades, significant scientific developments and cutting-edge technologies have made possible a technical revolution in molecular

biological methods. Metabolomics is technology devoted to the study of metabolites, which are final products of cell regulation processes due to genetic changes and environmental of biological system [14].

Different approaches can be used in metabolomics [15]: (i) metabolite targeted analysis:-detection, identification and quantification of a single metabolite or a small group of specific metabolites; (ii) metabolite/metabolic profiling: identification and approximately quantification of a set of metabolites, which are studied together because of its structural or functional similarity, or because they are involved in specific metabolic pathways; (iii) metabolic fingerprinting: direct analysis of crude samples; it is a rapid analysis in which the identification of metabolites is not necessary and the goal is to classify phenotypes based on a metabolite pattern; and (iv) characterization of the complete metabolome (metabolomics): qualitative and quantitative analysis of the whole metabolome of a biological system.

The characterization of the metabolites in organisms, tissue or cell logically implies the use of both separation and detection advanced analytical techniques. As far as oxysterols are concerned, they have been determined by liquid chromatography (LC) or gas chromatography (GC), mainly coupled to mass spectrometry (MS), as well as by nuclear magnetic resonance (NMR) [6,16]. The last have a great potential to select proper biomarkers in order to diagnose associated diseases [13,17].

To determine the oxysterols in biological fluids and tissues, there are different ways to prepare the sample; they obviously depend on the matrix, analyte/s and technique/s to be used. Blood plasma has been separated to blood serum by using ethylenediaminetetraacetic acid (EDTA) and further centrifugation of the blood sample [18,19]. Then, the extraction of sterols from blood plasma can be carried out by alkaline hydrolysis. In this way, lipids such as phospholipids and triglycerides are degraded and their possible interferences could be avoided. Lastly, solid phase extraction (SPE) has been utilized to separate sterols from other substances [18,20].

A derivatisation reaction is usually required prior to GC analysis in order to increase the volatility of the oxysterols; N,O-bis(trimethylsilyl)trifluoroacetamide/trimethylchloroside (BSTFA:TMCS) is a very widely used derivatisation reagent. However, N-methyl-N-(trimethylsilyl)-trifluoroacetamide-1,4-dithioerythritol-trimethyliodosilane (MSTFA:DTE:TMIS) has been proposed as the most efficient derivatisation reagent based on the specificity/sensitivity ratio in GC-MS [21,22].

In spite of LC is recognised as a chromatographic technique which does not need a previous derivatisation of the samples, in the specific case of determining oxysterols the quantification is more difficult if the analytes are not derivatised [22]. The use of N,N-dimethyl-glycine, picolinyl esters or Girard P reagents has been described in LC-MS [23,24]. The derivatised compounds are more polar than the native

sterols and, consequently, the ionisation in the mass spectrometer is improved.

The number of papers describing LC-MS analytical methodologies for the oxysterol determination in biological matrices has lately increased to the detriment of GC-MS methods [25,26,27]. However, it is well-known that the use of GC-MS [19,28] has some important advantages, good peak shape and high chromatographic resolution, if compared with LC-MS. The use of MS in tandem (MSⁿ) could improve the sensitivity of the methods [29,30], but the MSⁿ experiencing has new and greater methodological difficulties.

The aim of this study was to develop a reliable and fit-for-purpose bioanalytical method for some the determination of Chol-related compounds in mice's blood plasma which were selected by the biotechnology-based company Neuron Bio as potential biomarkers of AD. Thus, this is an 'ad-hoc' developed method on demand to be applied in the analytical laboratory using the available chromatographic equipment.

The analytical method was developed by using GC-MS in selected ion monitoring (SIM) mode. An ion trap (IT) mass analyser was used to control the selection and MS separation of each Chol-related compound, using the distinct analytical m/z signal which can be measured for 7 α -hydroxycholesterol (7 α OHChol), desmosterol (Desm), lathosterol (Latho), 7 β -hydroxycholesterol (7 β OHChol), lanosterol (Lano), 24(S)-hydroxycholesterol (24OHChol), 25-hydroxycholesterol (25OHChol), 7-ketocholesterol (7KetoChol), 27-hydroxycholesterol

(27OHChol), and betulin as internal standard (IS). Each one of them is separate in an individual segment, with the exception of Latho and 7 β OHChol, which were detected within the same segment. This strategy has allowed the detection and quantification of Chol-related compounds with similar structure in mouse blood plasma. The interfering effect produced by high concentrations of cholesterol (which could mask the chromatographic peaks of the nearest analytes within the chromatogram) was avoided by applying the SIM strategy and defining specific segments or analytical windows to better detect each analyte.

2. Material and methods

2.1. Chemicals and reagents

Methanol and ethyl acetate (EA) were HPLC-grade by Panreac (Spain). BSTFA (N,O-bis(trimethylsilyl)trifluoroacetamide) with 1% TMCS (trimethylchloroside) was obtained from Sigma-Aldrich (United States) as derivatisation reagent. Solid standards of analytes were supplied by Enzo Life Sciences, United States (24OHChol, >98%), MP Biomedicals, France (25OHChol, >98%), Sigma-Aldrich, United States (Chol, \geq 99%; Latho, \geq 93%; Desm, \geq 84%; and 7KetoChol, \geq 90%) and Avanti Polar Lipids, United States (Lano, >99%; 7 α OHChol, >99%; 7 β OHChol, >99%; and 27OHChol, >99%). In addition, betulin, 98% purity, were supplied by Avanti Polar Lipids (United States) was used as IS.

All of them, as well the corresponding derived solutions, were stored in the dark at -20°C .

2.2. Standards solutions

A multi-standard stock solution containing $100\ \mu\text{g}/\text{mL}$ of each analyte (Chol, Desm, Lano, Latho, 7KetoOH, 7 α OHChol, 7 β OHChol, 24OHChol, 25OHChol and 27OHChol) was prepared in methanol weighing the appropriate amount of the solid pure standards.

In the same way, $100\ \mu\text{g}/\text{mL}$ stock standard solution of betulin was also prepared in methanol.

2.3. Samples and sample preparation

The study was carried out on two samples groups: (i) mice's blood plasma as test samples and (ii) goat's blood serum as quality control samples.

Six 9 month-old APPswe mice (transgenic) and three 9 month-old non-transgenic (Non-Tg) littermate mice of both genders were used. They were maintained in temperature and light-controlled rooms in Neuron Bio facilities. Three APPswe mice were fed with NST0037 (a novel statin developed by Neuron Bio [31])-enriched diet and three APPswe and three Non-Tg littermate mice with control diet from 9 month-old to 15 month-old. In the rest of the paper, we are going to denote each type of mice by using letters A, B, and C: A are non-transgenic mice,

Non-Tg; B are transgenic mice, APP^{swe}; and C are transgenic mice, APP^{swe} which were fed with NST0037.

The 15 month-old mice were anesthetized with sodium pentobarbital (300 mg/kg), their blood collected and the plasma obtained and frozen at -80°C.

Each extract was prepared applying the following procedure: 200 µL of blood sample were added on a glass tube, together with 40 µL of internal standard and 0.5 mL of EA. The tube was shaken on a vortex (1 min), was sonicated (50 W) during 15 min and centrifuged for 5 min at 10000 rpm (approx. 8400 'g'), and then 100 µL of the organic phase was transferred to an eppendorf vial and stored in dark atmosphere at -20°C until the analysis was performed.

The set of quality control samples was constituted by 5 extracts in EA of the blood serum from a single goat. Each extract (300 µL) was prepared in the same way that the test samples, starting from 600 µL of blood sample. As indicated above, the extracts were stored in the dark at -20°C until the analysis. Notice that the sample treatment removes the proteins. Thus, the extract matrices from plasma and serum are similar and the serum samples can be used as commodity representative of plasma sampler in order to carry out an effective analytical quality control.

The final concentration of IS was 40 µg/mL in all the extracts.

2.4. Extract sample treatment

Quality control (QC) standard solutions were prepared by adding 0.5, 1.0, 2.5 and 5.0 $\mu\text{g/mL}$ of each analyte into goat blood serum extract. To carry out an efficient monitoring of the analytical method, two types of QC samples were used (spiked before or after the extraction).

In every case, 50 μL of the sample extract (both test and QC) was transferred to a chromatographic vial containing a 250 μL insert. The solvent was evaporated with N_2 and, thereafter, an aliquot of 50 μL of derivatisation reagent was added. The reaction mixture was left at room temperature over 1 h.

An overall scheme showing the preparation of both sample groups (test samples and quality control samples) can be found in Figure 1.

2.5. GC-(EI/IT)MS instrument and method conditions

All separations were performed by using a VARIAN GC 3800 gas chromatograph (PA, USA) equipped with a split/splitless injector coupled to a VARIAN (IT)MS 4000 (ion trap) mass spectrometer (PA, USA) equipped with an electron impact (EI) source. 1 μL of each sample was introduced into the GC equipment using a robotized system autosampler module (CombiPal, CTC ANALYTICS, Switzerland). Splitless injection mode was selected and the injector temperature was held at 250°C.

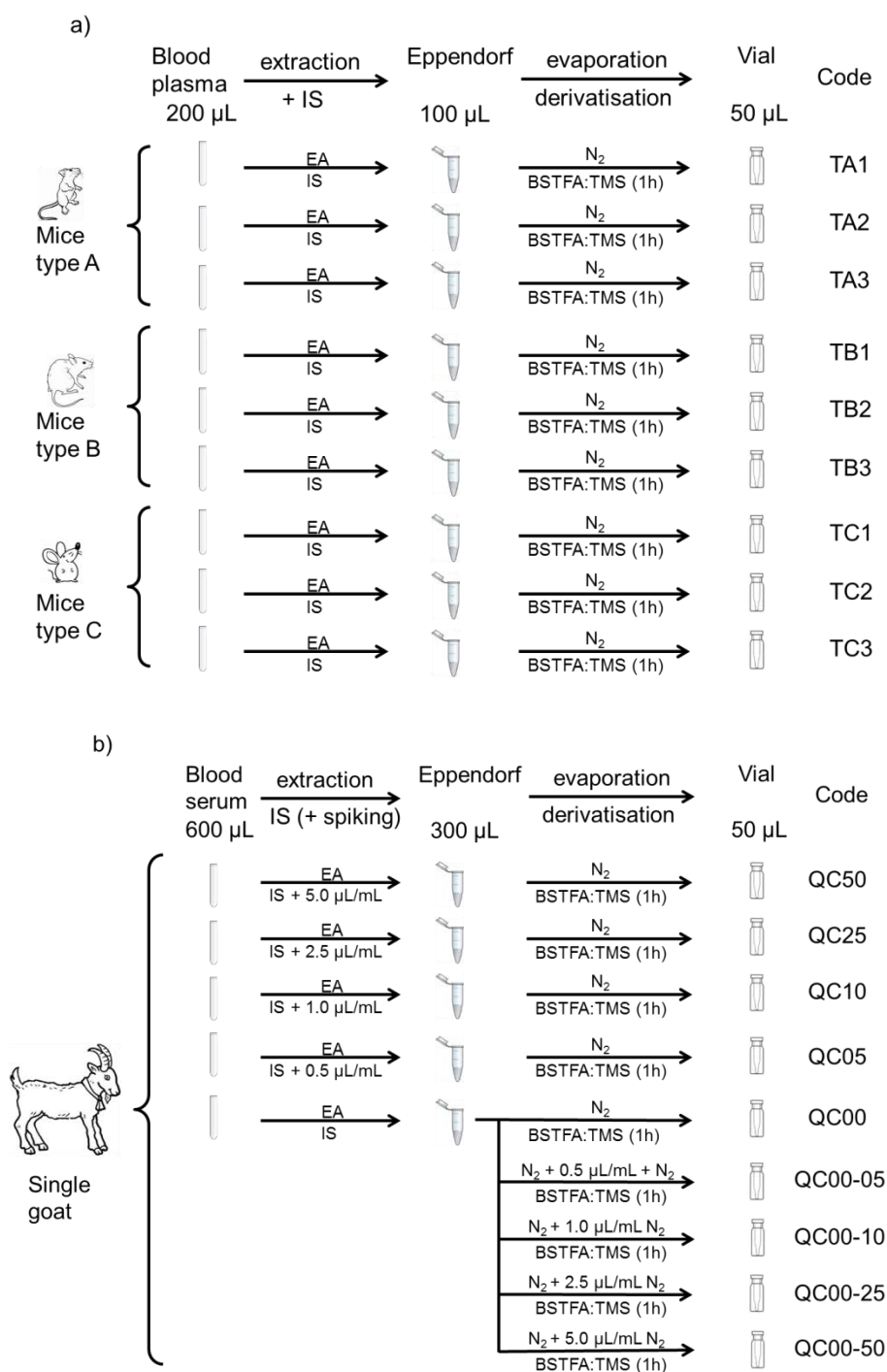


Figure 1. Scheme including the set of test samples (a) and quality control samples (b). In every vial the code of the sample was indicated. The sample codes for the test samples (T) identify the types of mice (A, B or C) and the number of specimens (1, 2 or 3). In the same way, the quality control samples (QC) indicate the spiked concentration (50, 25, 10, 05 or 00) in μ g/mL. (EA: ethyl acetate; IS: internal standard; BSTFA/TMS: silylation reagent).

The analytical column was a capillary column coated with 5% diphenyl-95% dimethyl polysiloxane stationary phase (Agilent Technologies DB-5MS; 30 m × 0.25 mm i.d. × 0.25 μm film thickness, maximum temperature 350°C; Agilent Technologies J&W, Santa Clara, USA). An innovation was used including a pre-column with silica deactivated (1 m × 0.25 mm i.d.) with press-fit connections (Agilent Technologies, Santa Clara, USA).

The column oven temperature of 200°C was held for 0.50 min, then programmed at 20°C/min to 300°C and maintained at this temperature for 10 min, giving a total run time of 15.50 min. Helium (99.995%) was used as the carrier gas and its flow rate was 1.2 ml/min. This method was optimized from a previously proposed method [32].

The mass spectrometric conditions were as follows: the ion source temperature was held at 250°C during the GC-MS runs. The transfer-line temperature was maintained at 280°C. The electron energy was 70 eV and the emission current 10 μA. In full-scan mode, average spectra were acquired in the m/z range of 50-600 m/z and were recorded at a scan speed of 1.20 s. Scan control, data acquisition, and processing were performed by version 6.9 MS Workstation (VARIAN) software data system.

Selected ion monitoring (SIM) mode was selected and was carried out by monitoring 456-458 m/z for 7αOHChol (from 4.00 min to 8.98 min); monitoring 159, 253 and 351 m/z for Desm (from 8.98 min to 9.64 min); 213, 255, 353, 443 and 456-458 m/z for Latho and 7βOHChol (from

9.64 min to 10.20 min); 395-396 m/z for Lano (from 10.20 min to 11.31 min); 255, 323, 413 and 441 for 24OHChol (from 11.31 min to 11.99 min); 131 m/z for 25OHChol (from 11.99 min to 12.25 min); 367, 382, 472 and 545 m/z for 7KetoChol (from 12.25 min to 12.70 min); 161, 255, 417, 456, 546 m/z for 27OHChol (from 12.70 min to 13.60 min); 189, 203 m/z for betulin (from 13.60 min to 15.50 min). Peak identification was based on comparison with standards of retention times and mass spectra ions.

3. Results and discussion

3.1. Optimization of chromatographic and detection conditions

The starting *GC* conditions were provided by Neuron Bio. Further *GC-MS* optimization was carried out in 5 stages to improve the initial conditions.

In a first stage of the optimization, each analyte was individually injected in *GC-MS* (full scan mode) to determine its retention time and its characteristic m/z. The found values for each of them and IS are included in Table 1. As far as the m/z values are concerned, it is important to say that the *MS* signals appear in decreasing order of intensity.

After characterizing the individual behaviour of each analyte, we proceeded to analyse a multi-standard solution (full scan mode); an example is given in Figure 2.

Table 1. Cholesterol precursors and oxysterols characterized by retention time (tR) and m/z.

Standard	m/z	tR (min)
7 α -hydroxycholesterol	456, 457	8.73
cholesterol	368, 329, 133, 353, 145, 458	9.14
desmosterol	253, 351, 159, 456	9.49
lathosterol	255, 458, 213, 443, 353	9.74
7 β -hydroxycholesterol	456, 457	9.92
lanosterol	395, 396	10.68
24(S)-hydroxycholesterol	145, 323, 413, 159, 441, 255, 546	11.82
25-hydroxycholesterol	131, 145	12.17
7-ketcholesterol	472, 367, 382, 545, 161	12.40
27-hydroxycholesterol	456, 417, 161, 546, 255, 441, 129	13.06
betulin (IS)	189, 203, 496	14.47

It is very well-known that the concentration of cholesterol is quite high in plasma blood; therefore, to simulate this situation, we prepared a standard solution containing 2000 $\mu\text{g}/\text{mL}$ of Chol as well as 5 $\mu\text{g}/\text{mL}$ of each one the rest of analytes. The result is illustrated in Figure 3, where it is possible to observe that even with very high concentration levels of Chol within the sample under study Desm (which elutes just after Chol) could be properly detected. However, Figure 3 also demonstrated that when Chol is present at high concentrations, the signal of the other analytes is very low in comparison, making difficult

their appropriate determination. Thus, the next step of the optimization focused on improving the MS conditions trying the use of SIM detection.

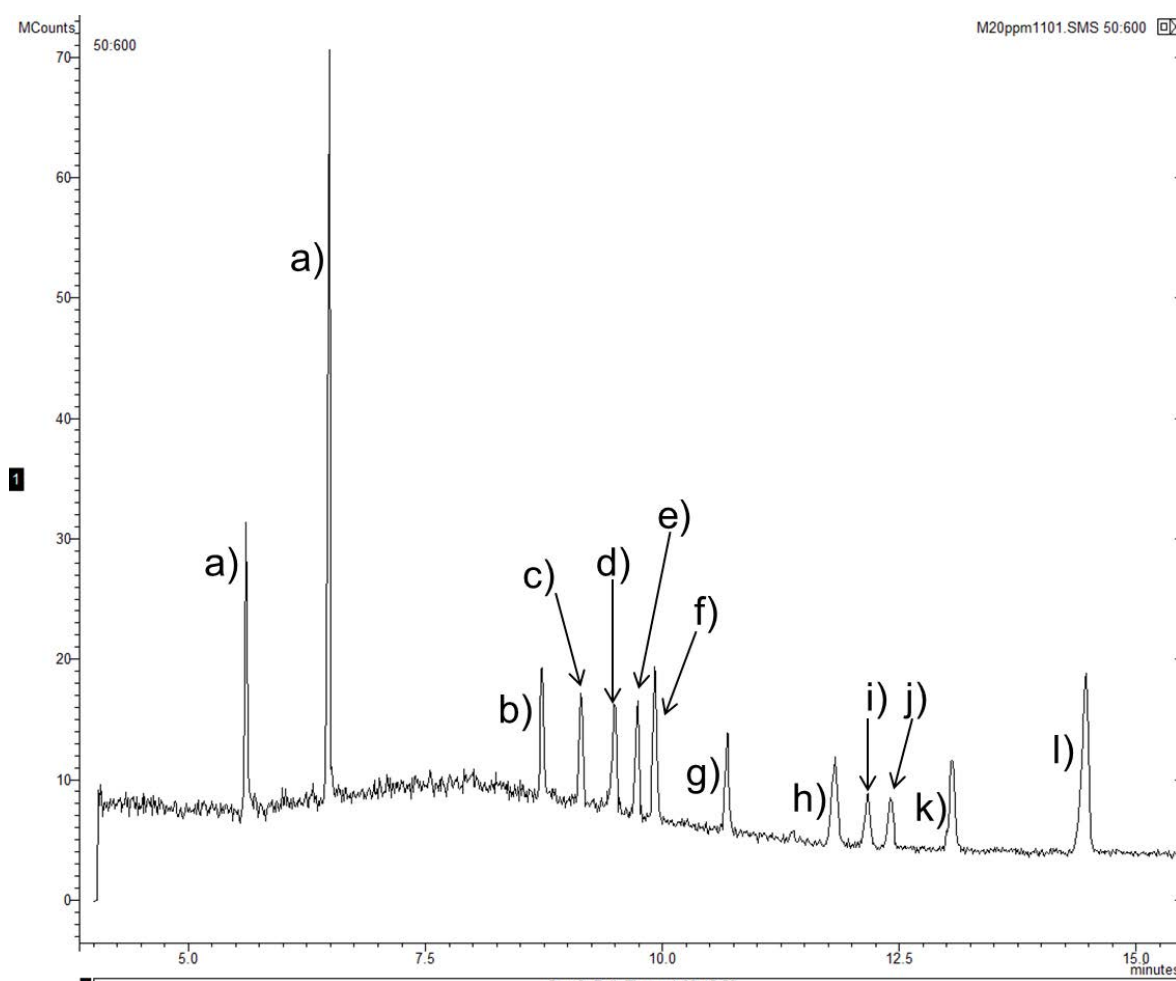


Figure 2. Chromatogram of 20 $\mu\text{g/mL}$ from multi-standard solution to check the retention time and the ions of each Chol-related compounds. The standards: a) BSTFA/TMS (silylation reagent); b) 7 α -hydroxycholesterol; c) cholesterol; d) desmosterol; e) lathosterol; f) 7 β -hydroxycholesterol; g) lanosterol; h) 24(S)-hydroxycholesterol; i) 25-hydroxycholesterol; j) 7-ketocholesterol; k) 27-hydroxycholesterol; l) betulin (IS).

SIM operational mode was selected and we created specific segments or analytical windows, to better detect each analyte (using specific MS

parameters and looking for characteristic m/z signals). A multi-standard solution containing $10 \mu\text{g/mL}$ of each analyte was injected into the GC-MS system (SIM mode). The result is shown in Figure 4.

IT was used as mass analyser and it exhibits the capability of carrying out MS/MS experiments; however, considering the harsh fragmentation achieved by the EI source, it is enough for SIM.

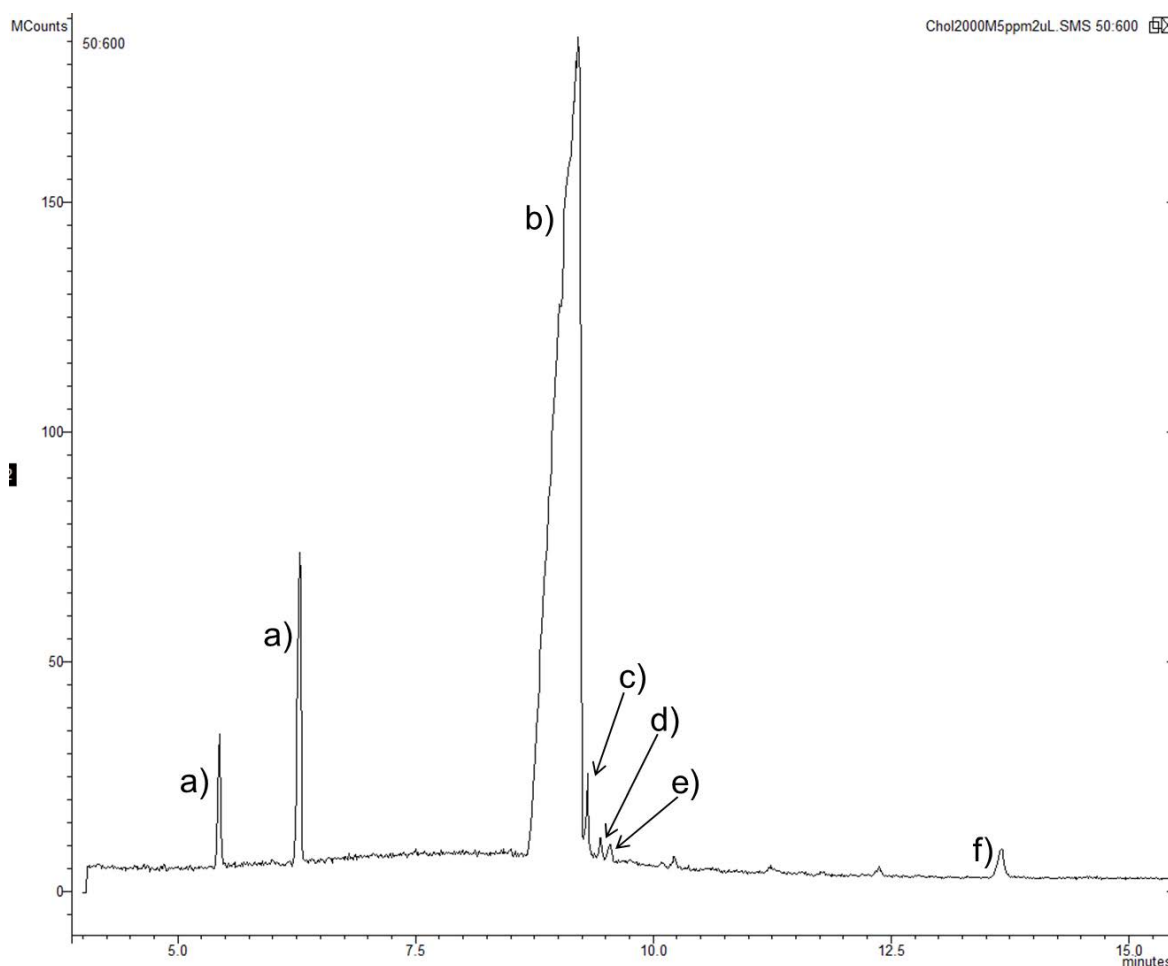


Figure 3. Chromatogram of standard solution to simulate of concentration from cholesterol ($2000 \mu\text{g/mL}$) in blood. The standards: a) BSTFA/TMS (silylation reagent); b) cholesterol; c) desmosterol; d) lathosterol; e) 7β -hydroxycholesterol; f) betulin (IS).

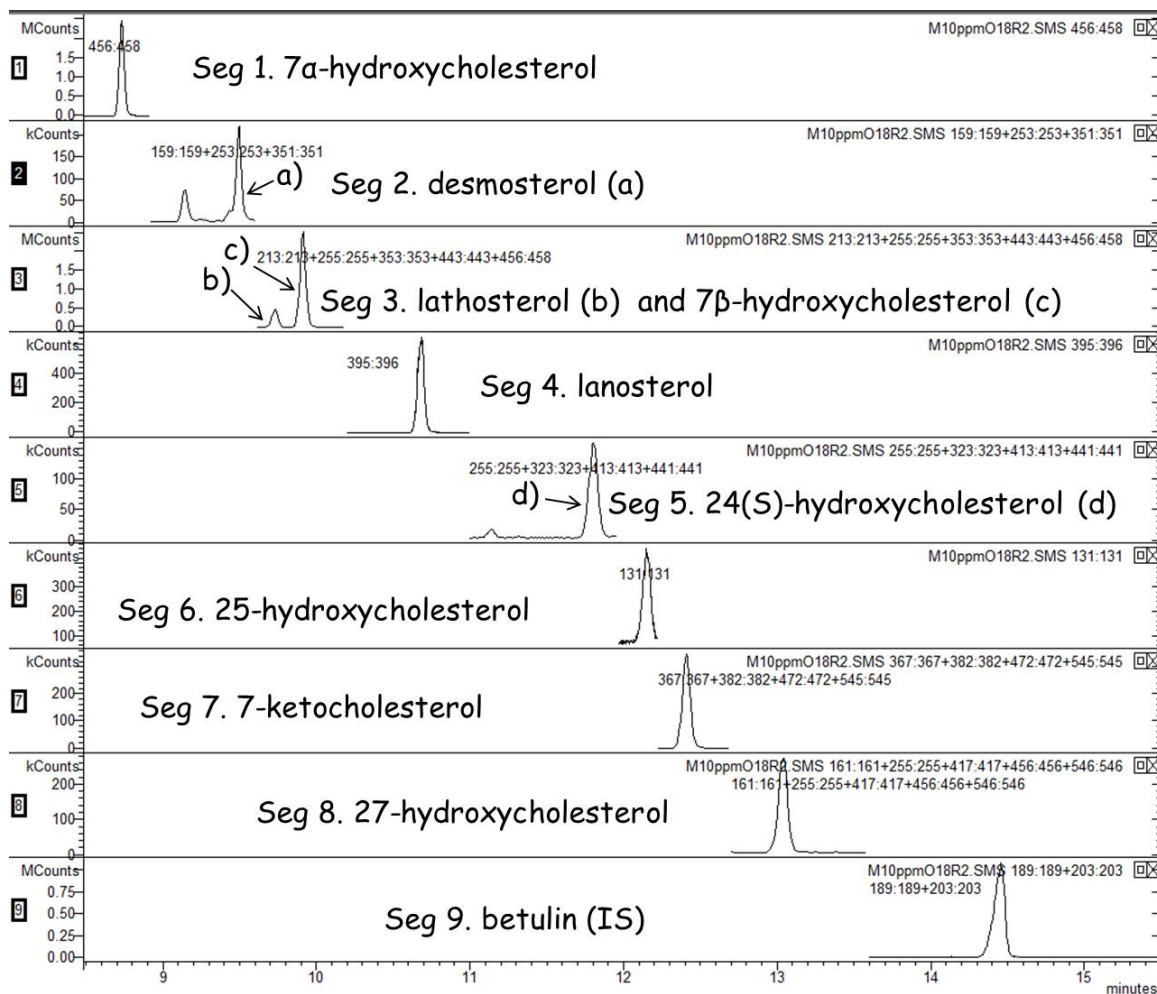


Figure 4. 9 segments (seg) which each Chol-related compounds and internal standard in SIM mode of 10 $\mu\text{g}/\text{mL}$ from multi-standard solution to improve the detection.

As a result, 9 analytes of interest plus the IS were separately detected by using 9 specific segments. The segments were optimized and defined taking into account the information included in Table 1 (retention times and m/z responses). SIM led to very promising results in terms of improvement of the signal to noise ratio (S/N), sensitivity, detection and quantification limits and some other important analytical validation characteristics.

After optimising the GC-MS conditions, some biological samples were injected to prove the suitability of the new methodology in relatively long sequences (including biological samples, which can be sometimes quite dirty).

After 25 analyses, the need of a pre-column was quite evident to reduce the noise which is increasing as the injection number is increasing. A silica phase with the same internal diameter than the column (in this case 0.25 mm) was selected. Pre-column had 1 m length and was connected to the column with press-fit connections. Once pre-column was installed the S/N will remain constant. The information that Table 1 includes regarding retention time was achieved with pre-column connected within the system.

3.2. Validation of the developed method

Multi-standard solutions and quality control samples were used to carry out the validation of the method. Table 2 shows all the analytical parameters during the validation of the developed method.

Multi-standard solutions were prepared at different concentration levels (0.1, 0.2, 0.5, 1.0, 2.5 and 5.0 µg/mL from a multi-standard solution containing 100 µg/mL of each analyte) to build the calibration curves. 40 µg/mL of betulin was added for each solution. 50 µL of the standard solution was transferred to a vial containing a 250 µL insert. The solvent was evaporated with N₂ and was added 50 µL of

derivatisation reagent. The reaction mixture was left at room temperature over 1 h.

Calibration curves, one for each Chol-related, were fitted by least-squares regression from relative area signal ($\text{area}_{\text{Chol-related}}/\text{area}_{\text{IS}}$) vs Chol-related concentration. The intercept was estimated as zero for all the analytes. For this reason, the relative areas were directly proportional to the concentration of each analyte.

So, it was possible to calculate a response factor (RF) for each compound, which is constant and independent of the concentration of each analyte and is shown in Table 2. In order to assess the goodness of fitting, determination coefficients and linearity indexes [33] were calculated for each calibration curve; the second one is included in Table 2 and was upper 98% in all the cases, except for 25OHChol (96.5%), which had higher detection limit and, therefore, low concentrations could not be detected.

One of the calibration levels was also used as stability control and was injected every 5 analyses over different sequences. The RF for each analyte was calculated from those injections and the stability of its value checked out over the sequences, establishing $\pm 10\%$ (%RSD) as a limit. 7KetoChol was the only compounds found out of control, after an important number of injections, but in any case, the %RSD was higher than 20%.

Detection and quantification limits were calculated considering the standard deviation calculated from the lower sample concentration, and

multiplying it by 3.28 and 10, respectively. Detection and quantification limits for 25OHChol were higher, fact which can be explained because only one m/z was characteristic of this analyte. 7 α OHChol, 7 β OHChol and Desm had lower detection and quantification limits, with 0.01 and 0.03 $\mu\text{g/mL}$, respectively, in both cases.

Spiked goat's blood serum samples were prepared to test the matrix effect and control the repeatability of the overall analytical process. To check the matrix effect, calibration curves were prepared spiking (at different levels (0.5, 1.0, 2.5 and 5.0 $\mu\text{g/mL}$)) -after the extraction- goat's blood serum with the Chol-related compounds (samples codes: QC00-05, QC00-10, QC00-25, QC00-50); the slope of these calibration curves was compared with those achieved by using external standard calibration. The native levels of analytes present in the goat's blood serum sample (QC00) were previously determined in order to carry out an effective blank correction. In QC00, the following compounds were found: 7 α OHChol, 7 β OHChol, 24OHChol and 7KetoChol. The other Chol-related compounds were not detected (nd). The area values found for this sample QC00 were considered in all the subsequent calculations and were subtracted to the values achieved after spiking with the known concentration levels (QC00-05 (spiked with 0.5 $\mu\text{g/mL}$ of multi-standard solution), QC00-10 (spiked with 1.0 $\mu\text{g/mL}$ of multi-standard solution), QC00-25 (spiked with 2.5 $\mu\text{g/mL}$ of multi-standard solution) and QC00-50 (spiked with 5.0 $\mu\text{g/mL}$ of multi-standard solution)) to evaluate matrix effect.

Table 2. Analytical parameters achieved during the validation of the method.

Compound	RF ^(a)	LIN (%) ^(b)	LD (µg/mL) ^(c)	LQ (µg/mL) ^(d)	Recovery Factor (%) ^(e)	Matrix Effect (%) ^(f)	Repeatability ^(g)	
							Consecutive measurements	Overall process
7αOHChol	0.115	98.2	0.01	0.03	83.8	97.3	0.106	0.200
Desm	0.013	98.7	0.01	0.03	74.2	96.9	0.073	0.166
Latho	0.025	98.0	0.12	0.37	67.7	104.8	0.088	0.277
7βOHChol	0.100	98.8	0.01	0.03	67.0	100.1	0.159	0.271
Lano	0.040	98.6	0.02	0.07	72.7	98.0	0.148	0.170
24OHChol	0.0053	98.0	0.14	0.44	71.0	96.5	0.181	0.228
25OHChol	0.022	96.5	0.35	1.06	97.8	99.4	0.146	0.294
7KetoChol	0.020	97.9	0.07	0.20	73.2	95.5	0.114	0.197
27OHChol	0.023	98.5	0.02	0.07	65.8	97.7	0.059	0.203

^(a) RF: Response factor. The intercept was considered as zero in the calibrate curves. Thus, the areas of each compound were directly proportional to its concentration.

^(b) LIN: Linearity Index. This index gives an estimation of the linearity as follow: $LIN = 1 - (SD_{\text{residual}}/Slope)$.

^(c) LD: Detection Limit. Calculated from the standard deviation multiplied by 3.28.

^(d) LQ: Quantification Limit. Calculated by using the standard deviation multiplied by 10.

^(e) Recovery Factor. This percentage was calculated with the results of the global error divide the error factor to matrix effect.

^(f) Matrix Effect. This percentage was calculated with the results of the error factor due the matrix effect. 100% in the case in which the slopes of the external standard calibration and standard addition calibration curves are identical.

^(g) Repeatability. Standard deviation was calculated to the repeatability of the measuring system or the overall analytical process.

Compounds: 7α-hydroxycholesterol (7αOHChol), desmosterol (Desm), lathosterol (Latho), 7β-hydroxycholesterol (7βOHChol), lanosterol (Lano), 24(S)-hydroxycholesterol (24OHChol), 25-hydroxycholesterol (25OHChol), 7-ketocholesterol (7KetoChol), 27-hydroxycholesterol (27OHChol).

As can be seen in Table 2, matrix effect was found between 95.5 and 104.8%, for 7KetoChol and Latho, respectively (being 100% in the case in which the slopes of the external standard calibration and standard addition calibration curves are identical); in other words, matrix effect was not observed.

Moreover, to give an estimation of the overall repeatability (including sample preparation and measurement) goat's blood serum samples spiked with the Chol-related under study -before the extraction- were prepared at the same levels as before (samples codes: QC-05, QC-10, QC-25 and QC-50). Standard deviation (SD) of the determined values is given in Table 2, and was found within the range 0.059-0.181 for repeatability of the measurements and 0.166-0.294 for the overall process.

The theoretical added concentration did not exactly coincide with the concentration obtained after the extraction and measurement. Since in the previous step, it was determined that no matrix effect was observed, we concluded that the difference was due to the recovery (obviously lower than 100%). Recovery values were calculated and they are shown in Table 2, being found from 65.8 to 97.8%, for 27OHChol and 25OHChol, respectively. These recovery factors should be used to calculate the real concentration of each analyte in any sample.

3.3. Application of the method to the analysis of biological samples

Biological test samples were used for the application of the method. Each test sample was replicated twice. Figure 5 shows an example of the results achieved for a mouse's blood plasma sample (TA2) by using the SIM method with 9 optimised segments.

In this example, all the selected analytes were found, with the exception of Lano (segment 4) and 25OHChol (segment 6); both analytes were indicated as not detected. In addition, there are some Chol-related compounds which were not quantified but was detected, like Latho in the samples TA1, TB1, TC2 and TC3; and 7KetoChol in the samples TB1 and TC1. The rest of the analytes appeared with an appropriate signal to noise ratio. Segments 5 (24OHChol), 7 (7ketoChol) and 8 (27OHCHol) showed higher noise than the same segments when standard solutions were analysed, fact which is logical.

Quantitative results are summarised in Table 3, which includes the average concentration ($\mu\text{g/mL}$) of each analyte and the relative standard deviation (%RSD). The recovery factor was used to recalculate the real analyte concentration after extraction.

As stated above, there were three types of mice (A, B and C) and three specimens in each group (1, 2 and 3). Samples belonging to the same group did not always show very similar results, due to the fact that we studied different specimens. Even if different mice had the same treatments, they could exhibit different reactions.

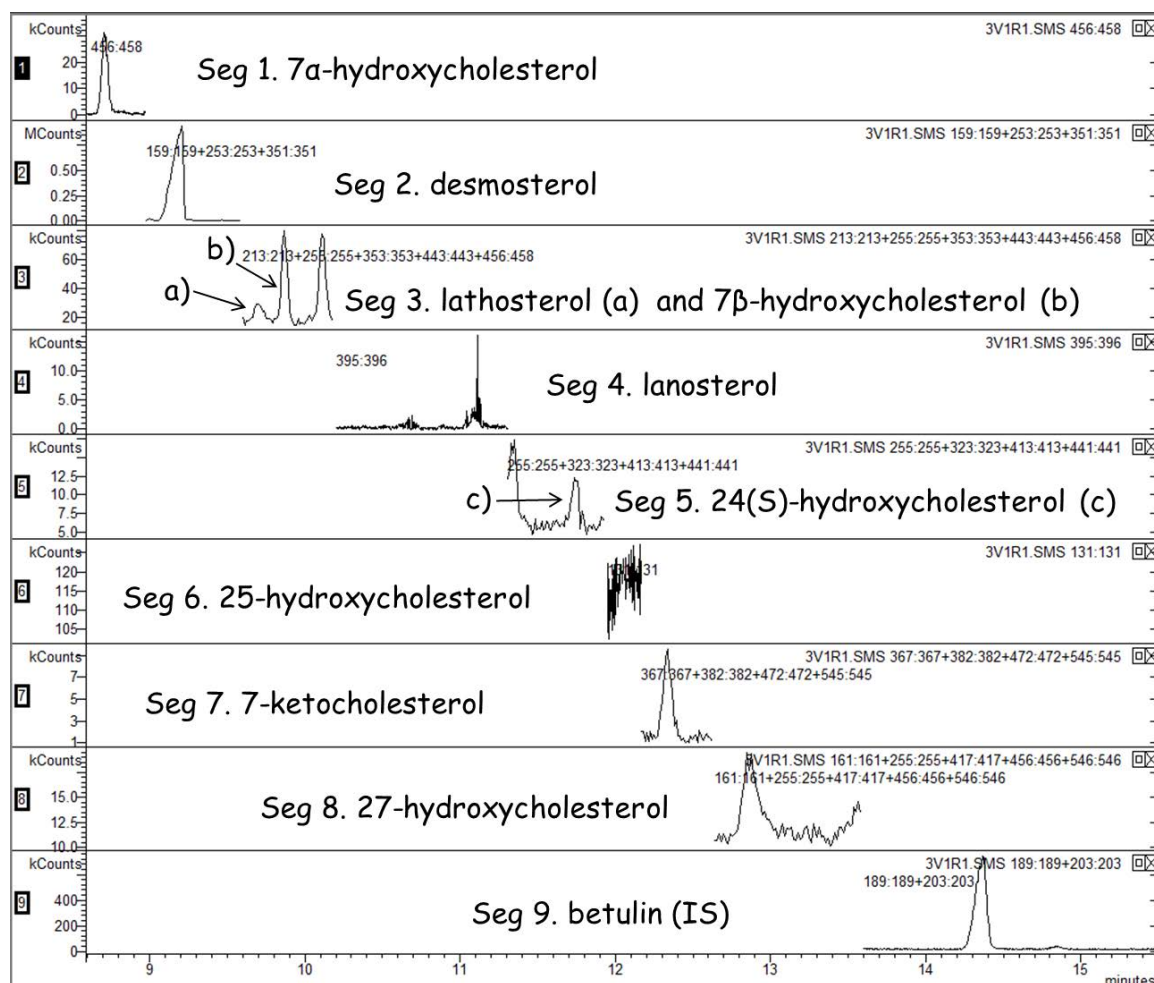


Figure 5. 9 segments (seg) which each Chol-related compounds in SIM mode of test sample (TA2) to show the seg in a real sample.

Lano and 25OHChol were not detected in any sample. Desm was only detected in TA samples (TA1, TA2 and TA3) with concentration values oscillating from 0.25 to 0.39 $\mu\text{g/mL}$, and in TB1, with 0.19 $\mu\text{g/mL}$. In the case of Latho, it was found in the three specimens of mice type A, with amounts 0.39 and 0.46 $\mu\text{g/mL}$ (TA3 and TA2), the other one was detected but was not quantified. It was also found in two specimens of mice type B (TB1 and TB2), with concentration values of 0.49 and 0.19 $\mu\text{g/mL}$ respectively, the last one was a rough estimate, and in two mice's

samples type C (0.20 and 0.34 µg/mL (samples TC2 and TC3)) was detected but was not quantified.

The rest of Chol-related compounds showed, in general, a decreasing trend in terms of concentration levels in mice A, B and C, apiece.

Table 3. Concentration of each Chol-related compounds from test samples. The concentration was recalculated with the appropriate recovery factor.

	Data	7αOH Chol	Desm	Latho	7βOH Chol	Lano	24OH Chol	25OH Chol	7Keto Chol	27OH Chol
TA1	Average (µg/mL)	0.76	0.25	d (0.13)	0.80	nd	1.28	nd	2.92	0.39
	RSD (%)	5.51	30.08	33.67	0.89	-	14.75	-	3.90	14.74
TA2	Average (µg/mL)	0.22	0.28	0.46	0.41	nd	1.06	nd	1.62	0.87
	RSD (%)	3.80	31.19	22.59	18.89	-	8.50	-	8.10	33.02
TA3	Average (µg/mL)	0.70	0.39	0.39	0.76	nd	0.99	nd	2.99	0.84
	RSD (%)	9.48	7.61	27.30	9.16	-	11.98	-	9.14	22.16
TB1	Average (µg/mL)	0.14	0.19	0.49	0.19	nd	0.76	nd	d (0.17)	0.51
	RSD (%)	19.53	30.19	21.59	21.76	-	17.01	-	27.23	26.05
TB2	Average (µg/mL)	0.42	nd	d (0.19)	0.35	nd	2.01	nd	1.70	0.42
	RSD (%)	0.24	-	30.19	10.11	-	13.94	-	11.17	10.57
TB3	Average (µg/mL)	0.11	nd	nd	0.11	nd	1.04	nd	0.53	0.35
	RSD (%)	5.24	-	-	11.97	-	1.36	-	9.00	28.14
TC1	Average (µg/mL)	0.061	nd	nd	0.04	nd	0.72	nd	d (0.16)	0.26
	RSD (%)	9.74	-	-	10.26	-	20.86	-	11.56	23.14
TC2	Average (µg/mL)	0.15	nd	d (0.20)	0.13	nd	0.85	nd	0.87	0.24
	RSD (%)	5.33	-	29.67	22.95	-	19.91	-	2.01	30.00
TC3	Average (µg/mL)	0.057	nd	d (0.34)	d (0.016)	nd	0.98	nd	0.56	0.31
	RSD (%)	1.99	-	24.09	21.72	-	19.98	-	21.57	22.01

Compounds: 7α-hydroxycholesterol (7αOHChol), desmosterol (Desm), lathosterol (Latho), 7β-hydroxycholesterol (7βOHChol), lanosterol (Lano), 24(S)-hydroxycholesterol (24OHChol), 25-hydroxycholesterol (25OHChol), 7-ketocholesterol (7KetoChol), 27-hydroxycholesterol (27OHChol). RSD: Relative standard deviation; nd: Not detected; d: detected (in parentheses, a rough estimate is stated).

As far as 7 α OHChol is concerned, it was found at higher concentrations in mice type A, with the following values: 0.76, 0.22 and 0.70 $\mu\text{g/mL}$. However, in samples belonging to the group type B, the concentrations values were 0.14, 0.42 and 0.11 $\mu\text{g/mL}$ in the three specimens respectively. Mice C showed the lowest values from 0.057 to 0.15 $\mu\text{g/mL}$. Situation was similar regarding 7 β OHChol, finding the highest values for mice belonging to group A (0.80, 0.41 and 0.76 $\mu\text{g/mL}$). 27OHChol, and 24OHChol were found in all the samples. Their most remarkable levels were determined in TA2 (0.87 $\mu\text{g/mL}$ of 27OHChol), and TB2 (2.01 $\mu\text{g/mL}$ of 24OHChol). 7KetoChol was found in all samples except for TB1 and TC1 which were detected but not quantified. In this case, his most remarkable levels were determined in TA3 (2.99 $\mu\text{g/mL}$ of 7KetoChol).

In general, mice C showed a lower concentration of every Chol-related compounds, followed by mice B and mice A, which, in general, exhibited the highest concentration values. These results are consistent with the fact that mice C were fed with statin. There are different studies which have pointed at a relationship between statins treatment and a reduction of the risk of suffering Alzheimer's disease [34,35,36,37]. This can be explained considering that the statins can help to reduce oxysterols levels, which affect amyloid beta production. Further studies are still necessary to clarify whether the use of statins can be a useful strategy for treating the onset of clinical dementia, but promising results can be already found in literature.

4. Conclusions

This study proves that *GC-(IT)MS*, in *SIM* mode, is an analytical technique suitable to detect and quantify cholesterol precursors and oxysterols in plasma blood samples. The high cholesterol concentration present in this biological fluid could mask the chromatographic peaks from such Chol-related compounds which are eluting close to cholesterol peak, but this drawback is avoided by using the *SIM* mode. As improvements, the background noise usually found in these biological samples was reduced by use of a *GC* pre-column. This strategy prolongs the analytical column's useful life.

In bibliography other chromatography-*MS* methods were reported showing better performance characteristics but it was proved that the proposed method is fitted for monitoring the contents of some AD biomarkers in blood plasma samples from mammals which are used as animals for experimental purposes.

Acknowledgments

The authors are grateful by collaboration to the biotechnology-based company Neuron Bio. The research project was founded by CEI BioTic / University of Granada.

Ethical conduct of research

The experimental protocol was approved by the Neuron Bio Ethics Committee for Animal Research. Animal care was carried out by qualified technicians supervised by veterinarians. Animals were treated in accordance with European laws (Directive 2010/63/EU) and the International guidelines for ethical conduct in the care and use of experimental animals were applied throughout the study.

Author contributions

All authors done bibliographic search for the introduction; L. Valverde-Som prepared the multi-standard stock solution which were designed by L.Valverde-Som, A. Carrasco-Pancorbo, N. Navas and L. Cuadros Rodríguez; S. Sierra, S. Santana and J.S. Burgos conceived and designed the extraction treatment; S. Sierra and S. Santana prepared the samples extract; L. Valverde-Som, A. Carrasco-Pancorbo and C. Ruiz-Samblás optimized the method GC-(IT)MS with SIM mode; L. Valverde-Som performed the experiments and she analyzed the data; A. Carrasco-Pancorbo and L. Cuadros-Rodríguez guided the project; L. Valverde-Som, A. Carrasco-Pancorbo and L. Cuadros-Rodríguez wrote the paper; all authors revised the paper.

Conflicts of interest

The authors declare no conflict of interest.

References

- [1] Garenc, C., Julien, P., Levy, E. (2010). Oxysterols in biological systems: the gastrointestinal tract, liver, vascular wall and central nervous system. *Free Radical Research*, 44, 47-73.
- [2] Bjökhem, I., Heverin, M., Leoni, V., Meaney, S., Diczfalusy, U. (2006). Oxysterols and Alzheimer's disease. *Acta Neurologica Scandinavica*, 114, 43-49.
- [3] Terao, J. (2014). Cholesterol hydroperoxides and their degradation mechanism. In: *Lipid hydroperoxide-derived modification of biomolecules*. Kato, Y. (Ed). Springer, London, 83-91.
- [4] Mulder, M. (2009). Sterols in the central nervous system. *Current Opinion in Clinical Nutrition and Metabolic Care*, 12, 152-158.
- [5] Poli, G., Biasi, F., Leonarduzzi, G. (2013). Oxysterol in the pathogenesis of major chronic diseases. *Redox Biology*, 1, 125-130.
- [6] Griffiths, W.J., Wang, Y. (2011). Analysis of oxysterol metabolomes. *Biochimica et Biophysica Acta*, 1811, 784-799.
- [7] Griffiths, W.J., Abdel-Khalik, J., Yutuc, E., Morgan, A.H., Gilmore, I., Hearn, T., Wang, Y. (2017). Cholesterolomics: An update. *Analytical Biochemistry*, 524, 56-67.
- [8] Trushina, E., Mielke, M.M. (2014). Recent advances in the application of metabolomics to Alzheimer's Disease. *Biochimica et Biophysica Acta*, 1842, 1232-1239.
- [9] Leoni, V., Caccia, C. (2011) Review. Oxysterols as biomarkers in neurodegenerative diseases. *Chemistry and Physics of Lipids*, 164, 515-524.
- [10] Baila-Rueda, L., Cenarro, A., Cofán, M., Orera, I., Barcelo-Batllori, S., Pocoví, M., Ros, E., Civeira, F., Domeño, C. (2013). Simultaneous determination of oxysterols, phytosterols and cholesterol precursors by high performance liquid

- chromatography tandem mass spectrometry in human serum. *Analytical Methods*, 5, 2249-2257.
- [11] Heverin, M., Bogdanovic, N., Lütjohann, D., Bayer, T., Pikuleva, I., Bretillon, L., Diczfalusy, U., Winblad, B., Bjökhem, I. (2004). Changes in the levels of cerebral and extracerebral sterols in the brain of patients with Alzheimer's disease. *Journal of Lipid Research*, 45, 186-193.
- [12] Lim, W.L.F., Martins, I.J., Martins, R.N. (2014). The involvement of lipids in Alzheimer's disease. *Journal of Genetics and Genomics*, 41, 261-274.
- [13] Kölsch, H., Heun, R., Kerksiek, A., Bergmann, K.V., Maier, W., Lütjohann, D. (2004). Altered levels of plasma 24S- and 27-hydroxycholesterol in demented patients. *Neuroscience Letters*, 368, 303-308.
- [14] De la Luz-Hdez, K. (2012) Metabolomics and mammalian cell culture. In: *Metabolomics*. Roessner, U. (Ed). InTech, Croatia, 3-18.
- [15] Nielsen, J. (2007). Metabolomics in functional genomics and systems biology. In: *Metabolome Analysis. An Introduction*. Villas-Bôas, S.G., Roessner, U., Hansen, M.A.E., Smedsgaard, J., Nielsen, J. (Eds). John Wiley & Sons, Hoboken, New York, 3-14.
- [16] Gouveia, M.J., Brindley, P.J., Santos, L.L., Correia da Costa, J.M., Gomes, P., Vale, N. (2013). Mass spectrometry techniques in the survey of steroid metabolites as potential disease biomarkers: A review. *Metabolism*, 62, 1206-1217.
- [17] Want, E.J., Cravatt, B.F., Siuzdak, G. (2005). The expanding role of mass spectrometry in metabolite profiling and characterization. *ChemBioChem*, 6, 1941-1951.
- [18] Dzeletovic, S., Breuer, O., Lund, E., Diczfalusy, U. (1995). Determination of cholesterol oxidation products in human plasma by isotope dilution-mass spectrometry. *Analytical Biochemistry*, 225, 73-80.

- [19] Ahmida, H.S.M., Bertucci, P., Franzò, L., Massoud, R., Cortese, C., Lala, A., Federici, G. (2006). Simultaneous determination of plasmatic phytosterols and cholesterol precursors using gas chromatography-mass spectrometry (GC-MS) with selective ion monitoring (SIM). *Journal of Chromatography B*, 842, 43-47.
- [20] McDonald, J.G., Smith, D.D., Stiles, A.R., Rusell, D.W. (2012). A comprehensive method for extraction and quantitative analysis of sterols and steroids from human plasma. *Journal of Lipid Research*, 53, 1399-1409.
- [21] Saraiva, D., Semedo, R., Castilho, M.C., Silva, J.M., Ramos, F. (2011). Selection of the derivatization reagent - The case of human blood cholesterol, its precursors and phytosterols GC-MS analyses. *Journal of Chromatography B*, 879, 3806-3811.
- [22] Andrade, I., Santos, L., Ramos, F. (2013). Advances in analytical methods to study cholesterol metabolism: the determination of serum noncholesterol sterol. *Biomedical Chromatography*, 27, 1234-1242.
- [23] Griffiths, W.J., Crick, P.J., Wang, Y. (2013). Methods for oxysterol analysis: Past, present and future. *Biochemical Pharmacology*, 86, 3-14.
- [24] Karuna, R., Christen, I., Sailer, A.W., Bitsch, F., Zhang, J. (2015). Detection of dihydroxycholesterol in human plasma using HPLC-ESI-MS/MS. *Steroids*, 99, 131-138.
- [25] Sugimoto, H., Kakehi, M., Satomi, Y., Kamiguchi, H., Jinno, F. (2015). Method development for the determination of 24S-hydroxycholesterol in human plasma without derivatization by high-performance liquid chromatography with tandem mass spectrometry in atmospheric pressure chemical ionization mode. *Journal of Separation Science*, 38, 3516-3524.
- [26] Pataj, Z., Liebisch, G., Schmitz, G., Matysik, S. (2016). Quantification of oxysterols in human plasma and red blood cells by liquid chromatography high-

- resolution tandem mass spectrometry. *Journal of Chromatography A*, 1439, 82-88.
- [27] Griffiths, W.J., Abdel-Khalik, J., Crick, P.J., Yutuc, E., Wang, Y. (2016). New methods for analysis of oxysterols and related compounds by LC-MS. *Journal of Steroid Biochemistry*, 162, 4-26.
- [28] Menéndez-Carreño, M., García-Herreros, C., Astiasarán, I., Ansorena, D. (2008). Validation of a gas chromatography-mass spectrometry method for the analysis of sterol oxidation products in serum. *Journal of Chromatography B*, 864, 61-68.
- [29] Matysik, S., Klünemann, H.H., Schmitz, G. (2012). Gas chromatography-tandem mass spectrometry method for the simultaneous determination of oxysterols, plant sterols, and cholesterol precursors. *Clinical Chemistry*, 58, 1557-1564.
- [30] Matysik, S., Schmitz, G. (2013). Application of gas chromatography-triple quadrupole mass spectrometry to the determination of sterol components in biological samples in consideration of the ionization mode. *Biochimie*, 95, 489-495.
- [31] Campoy, S., Sierra, S., Suarez, B., Ramos, M.C., Velasco, J., Burgos, J.S., Adrio, J.L. (2010). Semisynthesis of novel monacolin J derivatives: hypocholesterolemic and neuroprotective activities. *Journal of Antibiotics*, 63, 499-505.
- [32] Yamazaki, R., Nakagawa, S., Tanabe, A., Ikeuchi, T., Miida, T., Yamato, S. (2008). Determination of 24S-hydroxycholesterol in human cerebrospinal fluid by gas chromatography/mass spectrometry. *Bunseki Kagaku*, 57, 707-713.
- [33] Cuadros-Rodríguez, L., García-Campaña, A.M., Bosque-Sendra, J.M. (1996). Statistical Estimation of Linear Calibration Range. *Analytical Letters*, 29, 1231-1239.

- [34] Ostrowski, S.M., Johnson, K., Siefert, Shank, M.S., Sironi, L., Wolozin, B., Landreth, G.E., Ziady, A.G. (2016). Simvastatin inhibits protein isoprenylation in the brain. *Neuroscience*, 329, 264-274.
- [35] Sallustio, F., Studer, V. (2016). Targeting new pharmacological approaches for Alzheimer's disease: Potential for statins and phosphodiesterase inhibitors. *CNS and Neurological Disorders - Drug Targets*, 15, 647-659.
- [36] McGuinness, B., Craig, D., Bullock, R., Malouf, R., Passmore, P. (2004). Statins for the treatment of dementia. *Cochrane Database of Systematic Reviewers*, 7, 1-74.
- [37] Refolo, L.M., Pappolla, M.A., LaFrancois, J., Malester, B., Schmidt, S.D., Thomas-Bryant, T., Tint, G.S., Wang, R., Mercken, M., Pentanceska, S.S., Duff, K.E. (2001). A cholesterol-lowering drug reduces β -amyloid pathology in a transgenic mouse model of Alzheimer's disease. *Neurobiology Disease*, 8, 890-899.

1.6. CONCLUSIONES

A la vista de los resultados obtenidos, podemos concluir que se pueden **detectar y cuantificar los analitos de interés usando el método analítico optimizado propuesto en modo SIM con el equipo GC-(EI/IT)MS.**

Teniendo en cuenta el estudio del control de calidad, se comprueba que es necesario realizar un **cambio en el método analítico**, con lo que respecta a la selección de los iones en modo SIM del segmento seleccionado **para el 25-hidroxicolesterol**. Además de prestar especial atención a los metabolitos **latosterol y 7-ketocolesterol** que no están bajo control en este estudio, debido a que ha habido una falta de **reproducibilidad** en la obtención de los picos cromatográficos llegando a presentar cortes de los mismos en cada uno de sus segmentos.

Se ha propuesto una **modificación en el sistema de separación cromatográfica**, permitiendo la **eliminación del elevado ruido de fondo** de las muestras, que impedía la detección de algunos de los picos cromatográficos de los metabolitos de interés. Para ello se ha hecho uso de una pre-columna cromatográfica mejorando así el sistema de separación.

Las **muestras de plasma de ratón** son **buenas candidatas para el estudio**, y se ha comprobado que el pico del colesterol en las muestras de plasma de suero no solapaba con el del resto de metabolitos y es posible detectar los metabolitos objetivos del estudio.

A la vista de los resultados obtenidos, se puede concluir que los ratones con la dieta **Non-Tg**, ratones no transgénicos, son los que presentan **concentraciones mayores de metabolitos** del colesterol; las muestras que presentan **concentraciones menores** de metabolitos de colesterol son los ratones a los que se les ha suministrado la dieta codificada bajo el nombre **NST0037**, que son los ratones a los que se les ha suministrado un nuevo **fármaco con estatina**, desarrollado por Neuron Bio, en la dieta.

Por lo que, teniendo en cuenta la codificación dada del plasma de ratón, se puede concluir que el valor de concentración de los analitos desmosterol, 24(S)-

hidroxicolesterol y 27-hidroxicolesterol varía de mayor a menor de la siguiente manera:

Non-Tg > APPswe > NST0037

Se destacan esos analitos por ser los que hacen el intercambio homeostático entre el cerebro y la sangre (24(S)-hidroxicolesterol y 27-hidroxicolesterol), así como el desmosterol como precursor en la ruta del colesterol.

Por todo lo expuesto sería interesante **aumentar el número de muestras** proporcionadas de plasma de ratón, así como **trabajar con muestras de plasma humano** para llevar a cabo un estudio exhaustivo de la relación de cada uno de los metabolitos con la enfermedad de Alzheimer, **para el diagnóstico de la enfermedad mediante el análisis de sangre**. Así como hacer que se reduzca el ritmo al que avanza la enfermedad mediante el **fármaco** desarrollado por la empresa Neuron Bio.

1.7. COMUNICACIÓN A CONGRESO

- ❖ L. Valverde-Som, A. Carrasco-Pancorbo, A. González-Casado, L. Cuadros-Rodríguez. **Determinación de colesterol y sus metabolitos en suero mediante GC-MS con analizador de trampa de iones.** XIV Reunión del Grupo Regional Andaluz de la Sociedad Española de Química Analítica (GRASEQA). **Oral.**

ANEXOS I. ESTUDIOS COMPLEMENTARIOS NO PUBLICADOS

I.1. MUESTRAS INICIALES

La empresa Neuron Bio, previa a la investigación que se presenta en este capítulo, nos suministró el extracto de acetato de etilo obtenido sobre muestras de plasma de ratón de laboratorio, pertenecientes a tres ratones diferentes: transgénicos, no transgénicos y ratones alimentados con estatinas.

Estas muestras fueron cruciales para la optimización del método analítico, puesto que permitieron poner de manifiesto el **problema** que suponen estas **matrices** tan complejas, que presentan compuestos no volátiles que **saturan la columna cromatográfica** dando lugar a un elevado nivel de ruido cromatográfico. Por lo tanto, con estas pruebas preliminares fue posible optimizar el método analítico de separación cromatográfica con la introducción de una **pre-columna** en el cromatógrafo de gases.

La figura I.1.1 muestra dos ejemplos de una mezcla de patrones que contiene: lanosterol, latosterol, desmosterol, 7-ketocolesterol, 7 α -hidroxicolesterol, 7 β -hidroxicolesterol, 24(S)-hidroxicolesterol, 25-hidroxicolesterol, y 27-hidroxicolesterol, a una concentración de 0.5 $\mu\text{g}/\text{mL}$. En color negro y rojo aparecen dos las réplicas de la misma mezcla de patrones inyectados dentro de la misma secuencia cromatográfica, pero entre los que se han inyectado las muestras de plasma de ratón (un total de 18 inyecciones de muestra). En este caso, se muestran únicamente los segmentos del espectrómetro de masas en modo SIM, en los que aparecen: (a) el 24(S)-hidroxicolesterol (segmento comprendido entre 8.52 – 10.29 minutos, que corresponden con las imágenes "a" y "b" de la figura I.1.1); (b) el 27-hidroxicolesterol (segmento comprendido entre 10.29 – 12.74 minutos, que corresponde con las imágenes "c" y "d" de la figura I.1.1). En esta figura se ve claramente, como le afecta a la señal cromatográfica el problema de las matrices tan complejas que se están utilizando sin el uso de la pre-columna.

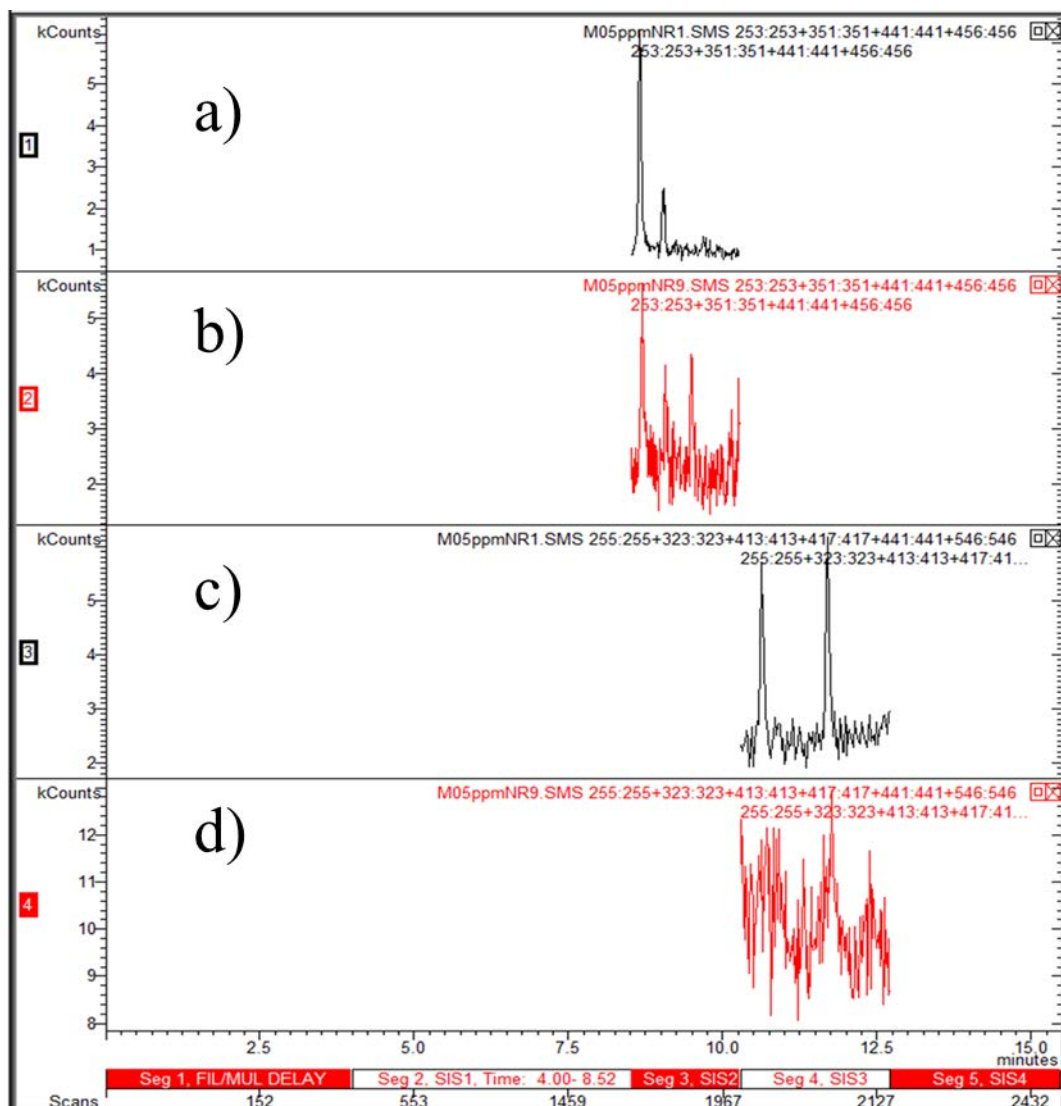


Figura I.1.1. Señal cromatográfica de los segmentos correspondientes al 24(S)-hidroxicolesterol (a y b) y al 27-hidroxicolesterol (c y d) de una mezcla de patrones de 0.5 $\mu\text{g}/\text{mL}$ de concentración perteneciendo a dos réplicas diferentes (que se muestran de color negro y azul).

Tras implementar el uso de la pre-columna, los problemas del aumento de la relación señal/ruido que se veían en la figura I.1.1 quedan eliminados. Así, en la figura I.1.2 se muestran dos ejemplos de la misma mezcla de patrones de 0.5 $\mu\text{g}/\text{mL}$ de concentración, pero esta vez con el uso de la pre-columna en el equipo GC-(IT)MS. En este caso, entre las réplicas que aparecen en color negro y rojo se han inyectado las muestras de plasma de ratón (un total de 28 inyecciones de muestra). Como se observa en la imagen, con el uso de la pre-

columna el ruido de fondo no va aumentando y la señal analítica aparece a la misma altura de pico cromatográfico.

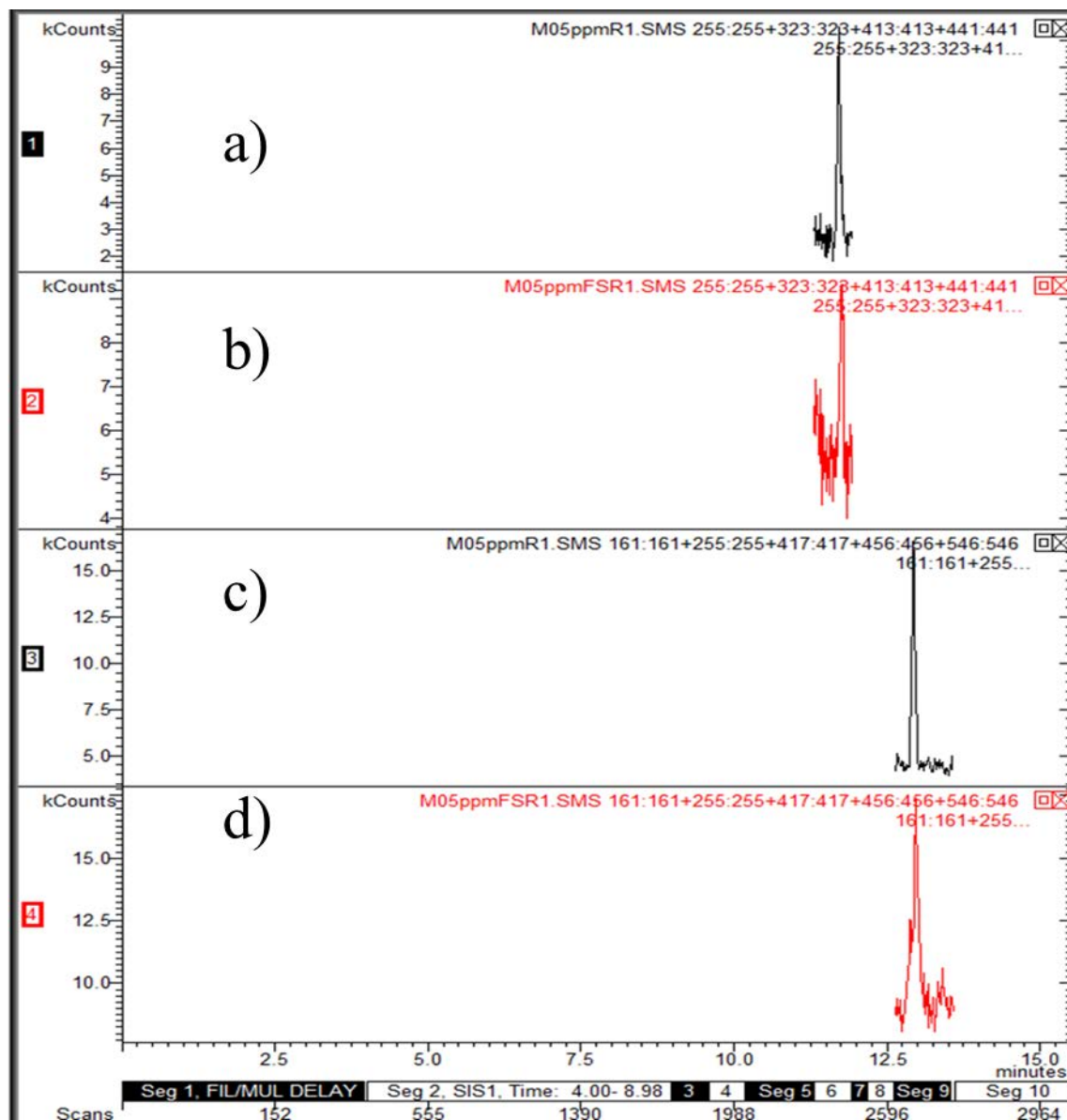


Figura I.1.2. Señal cromatográfica de los segmentos correspondientes al 24(S)-hidroxicolesterol (a y b) y al 27-hidroxicolesterol (c y d) de una mezcla de patrones de 0.5 $\mu\text{g}/\text{mL}$ de concentración perteneciendo a dos réplicas diferentes (que se muestran de color negro y azul), con el uso de pre-columna en el GC.

Además, esta primera prueba, sirvió para explicar a la empresa una serie de directrices que tenían que llevar a cabo a la hora de preparar las muestras y que sirvieron para dar los resultados del artículo científico. Entre ellas se destaca la

concentración de betulina, como patrón interno, que es necesario añadirlo en una concentración no demasiado elevada. También se le pide a la empresa una matriz de plasma que no contenga los analitos de interés para hacer ensayos de recuperación.

Por último, indicar que esta prueba nos sirvió para **confirmar la posibilidad de detectar y cuantificar los analitos de interés usando el modo SIM con el equipo GC-(IT)MS**, así como confirmar que dichas muestras son buenas candidatas para seguir con el estudio y poder comprobar si existe relación entre la concentración de determinados oxisteroles con la enfermedad de Alzheimer.

I.2. ESPECTROS DE MASAS

El modo SIM de adquisición en el espectrómetro de masas fue optimizado a partir de la relación m/z de los diferentes iones característico de cada analito de interés. Para ello, se inyectó individualmente cada patrón de cada analito de interés en modo *full scan* (con una concentración de 20 $\mu\text{g/mL}$). Posteriormente se inyectó una mezcla de todos ellos, con el objetivo de verificar los iones característicos seleccionados cuando se encuentran mezclados, por si existe algún tipo de interrelación entre ellos, y de esta forma seleccionar el tiempo de retención al que aparecen.

El método analítico usado es el que se ha descrito en la introducción, en este caso se seleccionó el rango de masas 50-600 m/z en modo *full scan*.

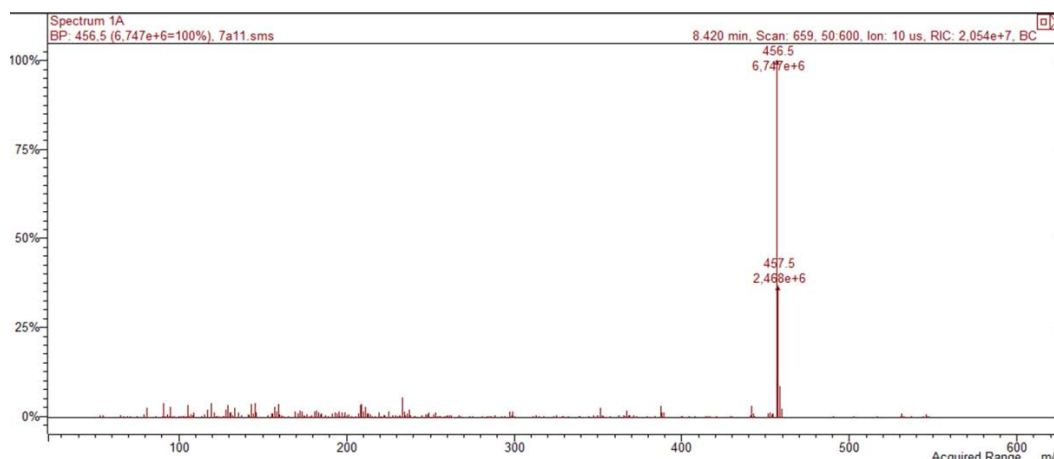
Una vez seleccionados los iones de interés, se pasó a la selección de diferentes segmentos, así como a incluir o eliminar alguno o varios de los iones seleccionados a priori. El método final consta de nueve segmentos, que tienen como objetivo identificar los analitos de interés con la máxima resolución posible para facilitar así su cuantificación.

En la figura I.1.3 se muestran los **espectros de masas** para cada uno de los analitos inyectados individualmente.

El **primer** segmento seleccionado corresponde con el metabolito 7α -hidroxicolesterol, el espectro de masas coincide con el del 7β -hidroxicolesterol puesto que presentan la misma fórmula molecular y lo único que cambia es la orientación del grupo hidróxido que se encuentra en la posición 7 en la molécula del colesterol. El 7β -hidroxicolesterol aparece en el tercer segmento, junto con el latosterol, ya que ambos metabolitos aparecían muy juntos en cuanto a tiempo de retención y no fue posible su separación en dos segmentos individuales. En el modo SIM se pueden seleccionar un rango de iones, por lo que se seleccionó para la búsqueda de ambos metabolitos el rango de iones 456-458 ya en ese rango se encuentra prácticamente los únicos iones que aparecen en el espectro de masas.

7 α -hidroxicolesterol

456 – 458



Latosterol (a) y 7 β -hidroxicolesterol (b)

213, 255, 353, 443, 456 – 458

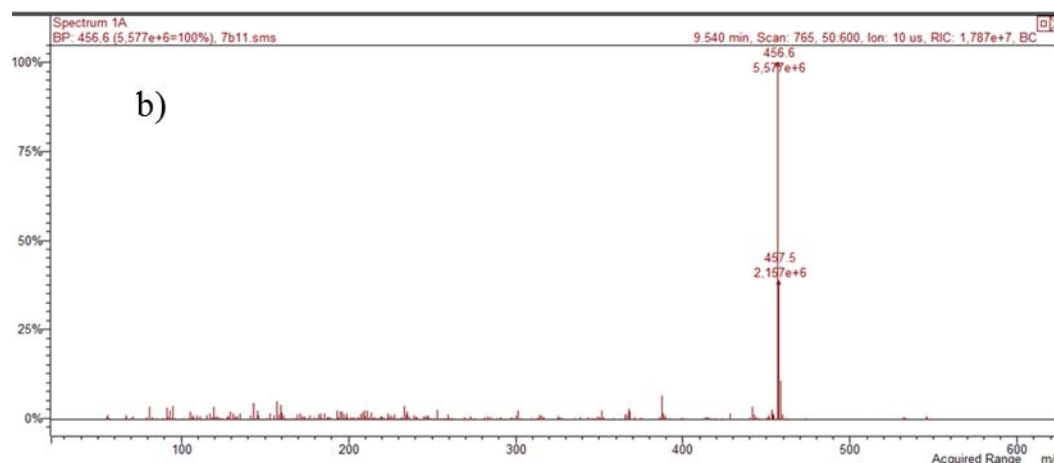
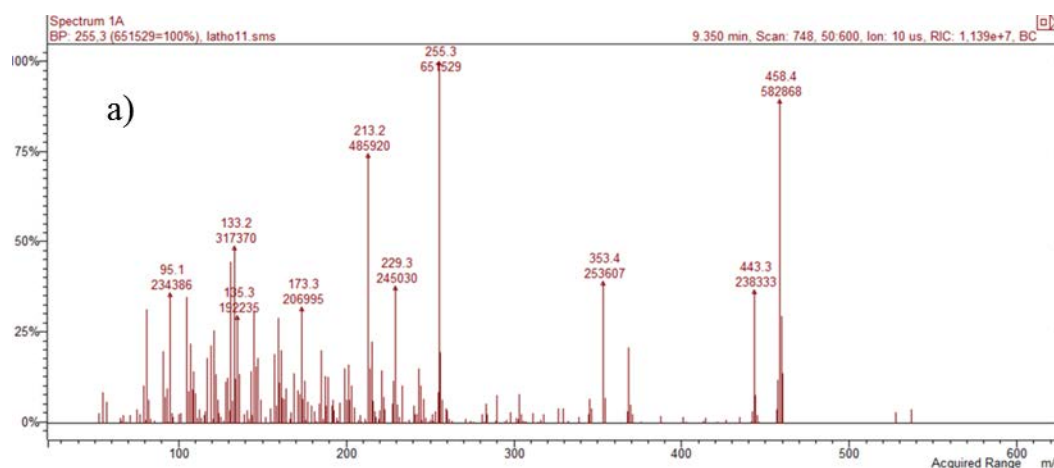
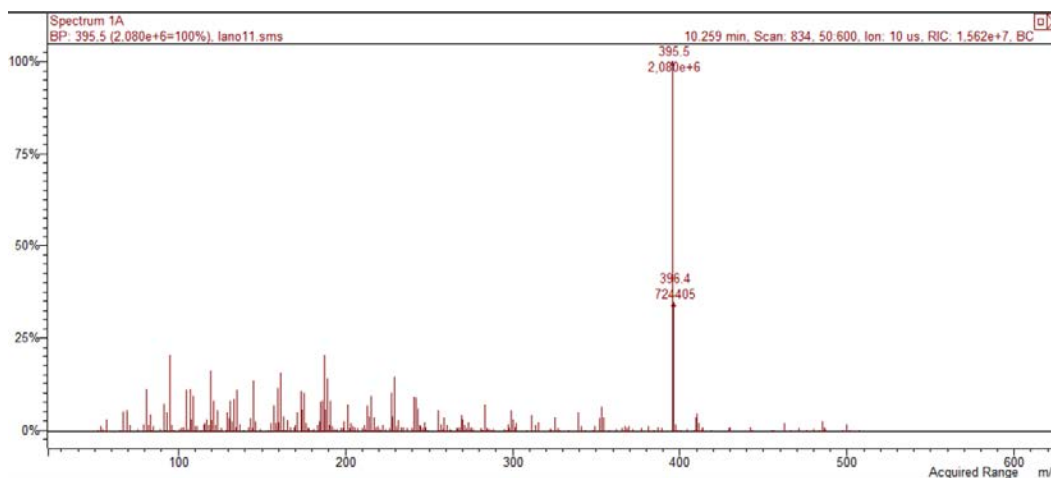


Figura I.1.3. Espectro de masas de cada uno de los analitos de interés por segmentos, a la izquierda aparece el nombre del analito y a la derecha la relación m/z seleccionada para ese segmento a posteriori en modo SIM.

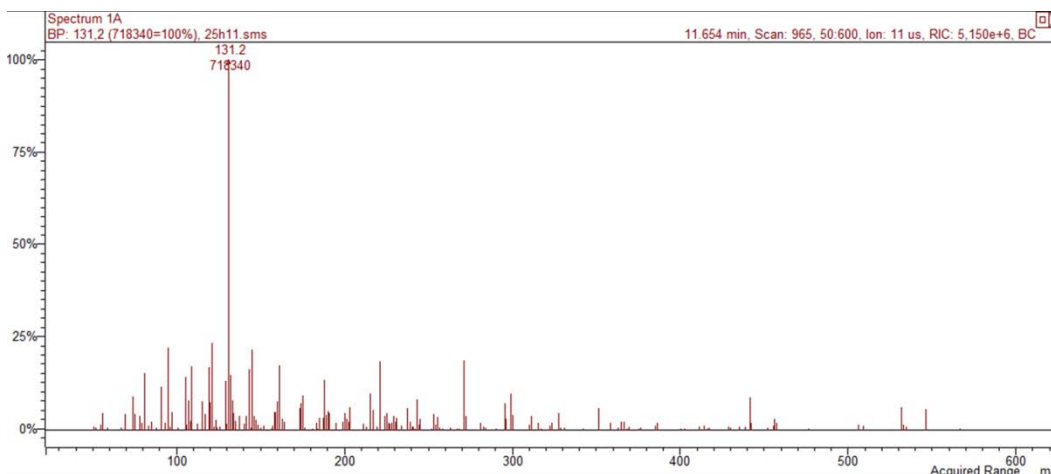
Lanosterol

395 – 396



25-hidroxicolesterol

131



7-ketocolesterol

367, 382, 472, 545

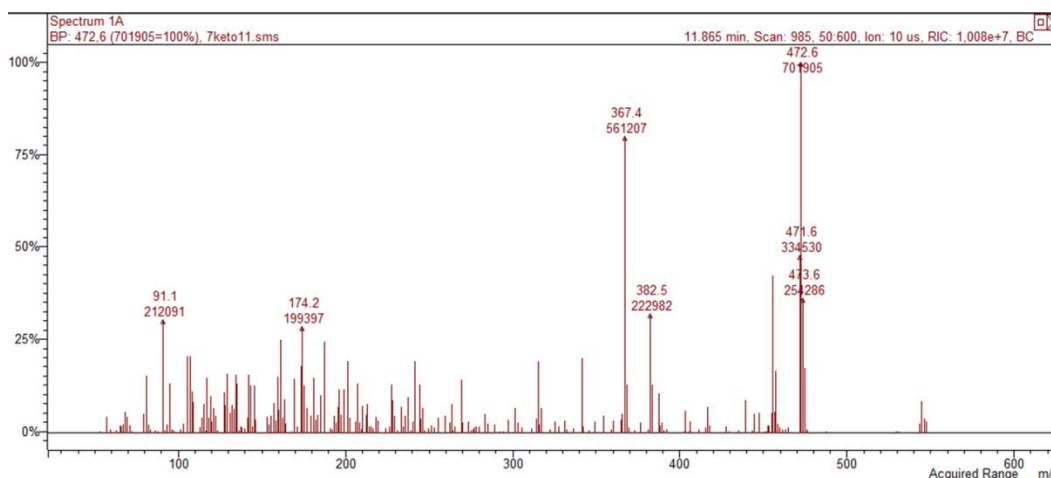
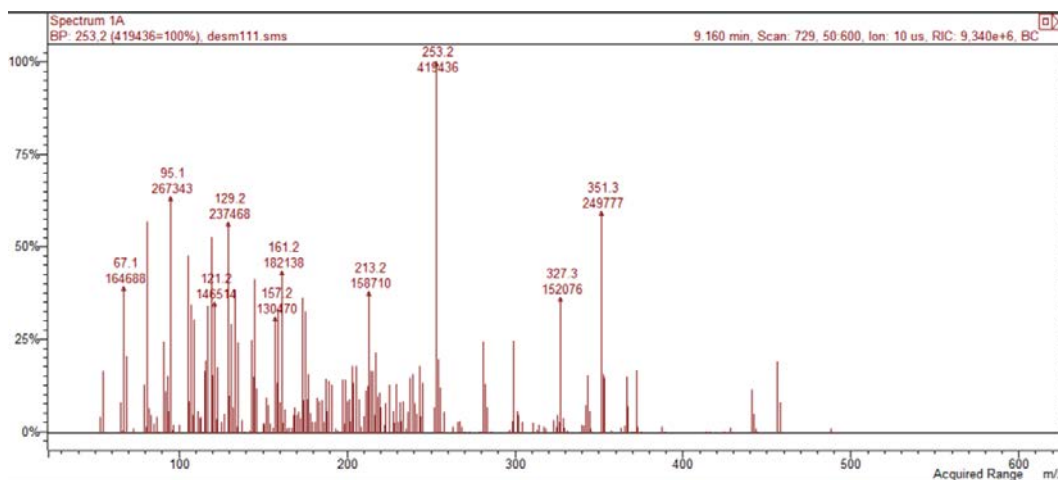


Figura I.1.3. (continuación).

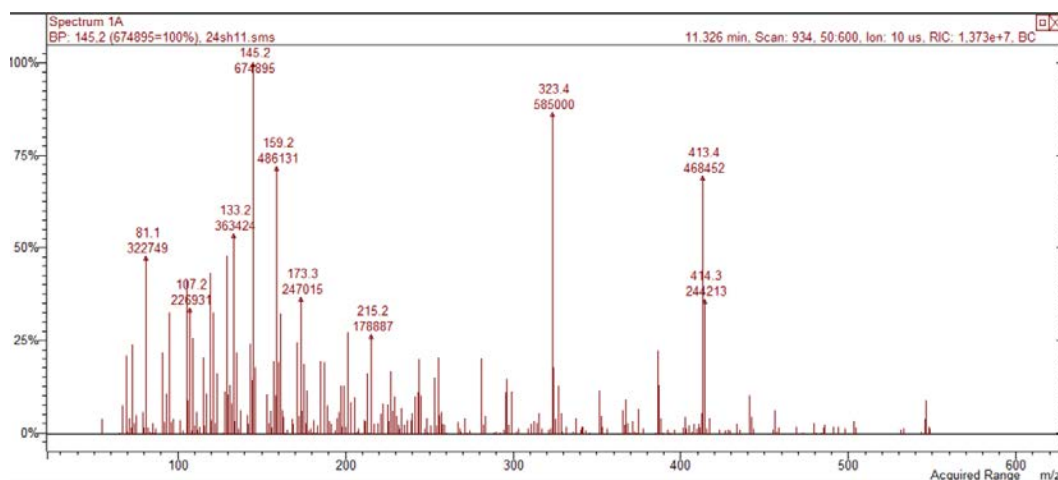
Desmosterol

159, 253, 351



24(S)-hidroxicolesterol

255, 323, 413, 441



27-hidroxicolesterol

161, 255, 417, 456, 546

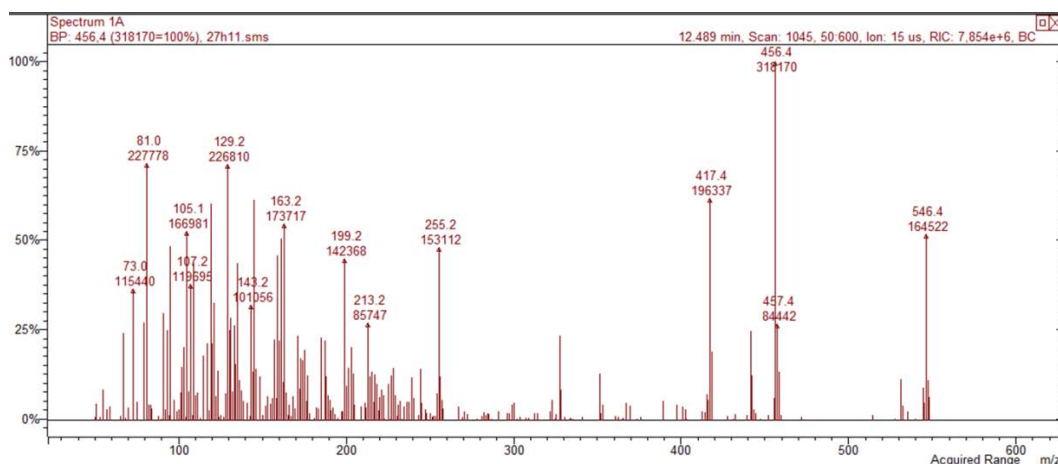


Figura I.1.3. (continuación).

Betulina (patrón interno)

189, 203

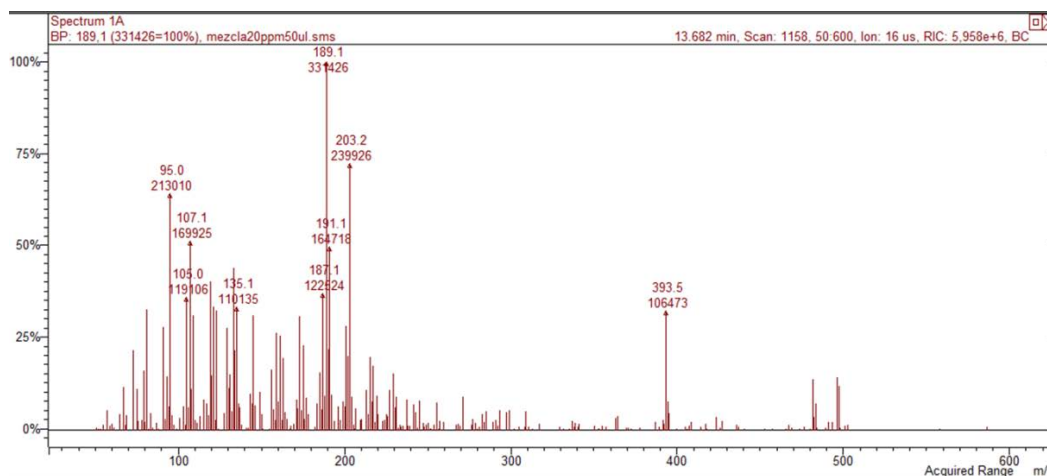


Figura I.1.3. (continuación).

El latosterol, también presenta el ion 458 y algunos a menor intensidad próximos a él dentro de ese rango. Por lo que es coherente que ambos metabolitos se busquen en el mismo segmento. Además, aparecen otros iones seleccionados entre los que destacan el 255 y 213 por ser los más intensos, aparte del 458. Así como los iones 353 y 443, que se seleccionan por aparecer con una abundancia de alrededor del 50% y con un rango de masas más elevado, puesto que las masas más bajas aparecen también en el ruido de fondo del cromatograma, lo que puede llevar a una búsqueda errónea del metabolito de interés.

El **cuarto segmento**, corresponde con la selección de los iones correspondientes con el lanosterol. En este caso, ocurre como con el 7-hidroxicolesterol, puesto que aparece prácticamente solo un ión característico en torno al 395, por lo que se selecciona el rango 395-396.

Otro metabolito que se comporta de forma similar es el 25-hidroxicolesterol cuyo ión más abundante es el 131, que corresponde con el **sexto segmento**. En este caso únicamente se selecciona ese ión, porque al poner un rango de iones, se incluían otros iones más pequeños que podrían confundirse con iones responsables del ruido de la señal y no era factible para cuantificar dicho metabolito.

En el caso del 7-ketocolesterol, **segmento siete**, aparecen dos iones claros y más abundantes: 367 y 472, y luego aparecen los iones 382 y 545 que no son muy abundantes, pero están en un rango de m/z alto y dan mejores resultados que si se eliminan o se seleccionan otros iones de alrededor.

El resto de segmentos presentan muchos iones en el rango m/z de 70 a 250, que corresponden con el desmosterol, 24(S)-hidroxicolesterol y 27-hidroxicolesterol, correspondientes con los **segmentos 2, 5 y 8**, respectivamente. En ellos, los iones 253 y 351 para el desmosterol, 323 y 413 para el 24(S)-hidroxicolesterol y para el 27-hidroxicolesterol 417, 456 y 546, fueron fácilmente escogidos por ser los iones más abundantes, el resto fueron elegidos a base de hacer diferentes pruebas en modo SIM con los patrones inyectados individualmente, así como en una mezcla en la que estaban presentes todos ellos.

Como **último segmento**, se seleccionó la betulina, que es utilizada como patrón interno y que tiene dos iones muy abundantes correspondientes con el 189 y 203.

Es importante saber que siempre, antes de inyectar una tanda de muestras o de patrones, es necesario hacer un **chequeo** para comprobar si se han movido los **tiempos de retención**, puesto que tenemos muchos **segmentos** en el **modo SIM** que pueden verse alterados si existe algún tipo de modificación en el equipo cromatográfico o si se inyectan las muestras en diferentes días.

I.3. RECTAS DE CALIBRADO

Para llevar a cabo la cuantificación de la concentración de cada uno de los metabolitos presentes en muestras de plasma de ratón, se obtienen las rectas de calibrado para cada metabolito en una mezcla de patrones a diferentes concentraciones (0.1, 0.2, 0.5, 1.0, 2.5 y 5.0 $\mu\text{g/mL}$, con 40 $\mu\text{g/mL}$ de betulina como patrón interno). La cuantificación se hace a partir de la obtención del área bajo la curva de cada metabolito obtenida del **cromatograma TIC** (*Total Ion Current*), que representa la suma de todas las intensidades presentes en el rango de masas seleccionado y detectado. En este caso, al hacer modo SIM, son pocas las masas presentes en cada segmento y todas características del mismo metabolito de interés, por lo que se prefiere trabajar con el TIC.

En esta investigación se definen **tres tandas analíticas** para cada uno de los estudios que se quieren llevar a cabo. De manera que, en cada una de las tandas analíticas se introducen patrones externos de calibración al principio y al final de la secuencia para obtener con ellos la recta de calibración externa correspondiente.

Las **rectas de calibrado** externo se hacen a partir de la **señal relativa** ($\text{Área_analito}/\text{Área_betulina}$) frente a la **concentración de analito**. Para cada tanda analítica se obtiene una recta de calibrado para cada analito, con la unión de los datos correspondiente con los patrones de calibración externos que se introducen al principio y al final de la secuencia en cada tanda, y se tratan los datos de las muestras de esta tanda.

En todos los casos se comprobó estadísticamente que para todos los analitos, el término independiente u ordenada en el origen, no era significativamente diferente de cero, por lo que se pudo aplicar un modelo del tipo $y_{\text{relativa}} = b \cdot x$. Lo que implica que el cociente de áreas (y_{relativa}) es directamente proporcional a la concentración de cada analito (x), o lo que es lo mismo, que para cada analito el **factor de respuesta** (FR), que corresponde con la pendiente de la recta (b) y es constante e independiente de la concentración. Además, se comprobó que las rectas de calibrado introducidas al principio y al final de la secuencia se pueden considerar iguales, con un índice de linealidad superior al 95%.

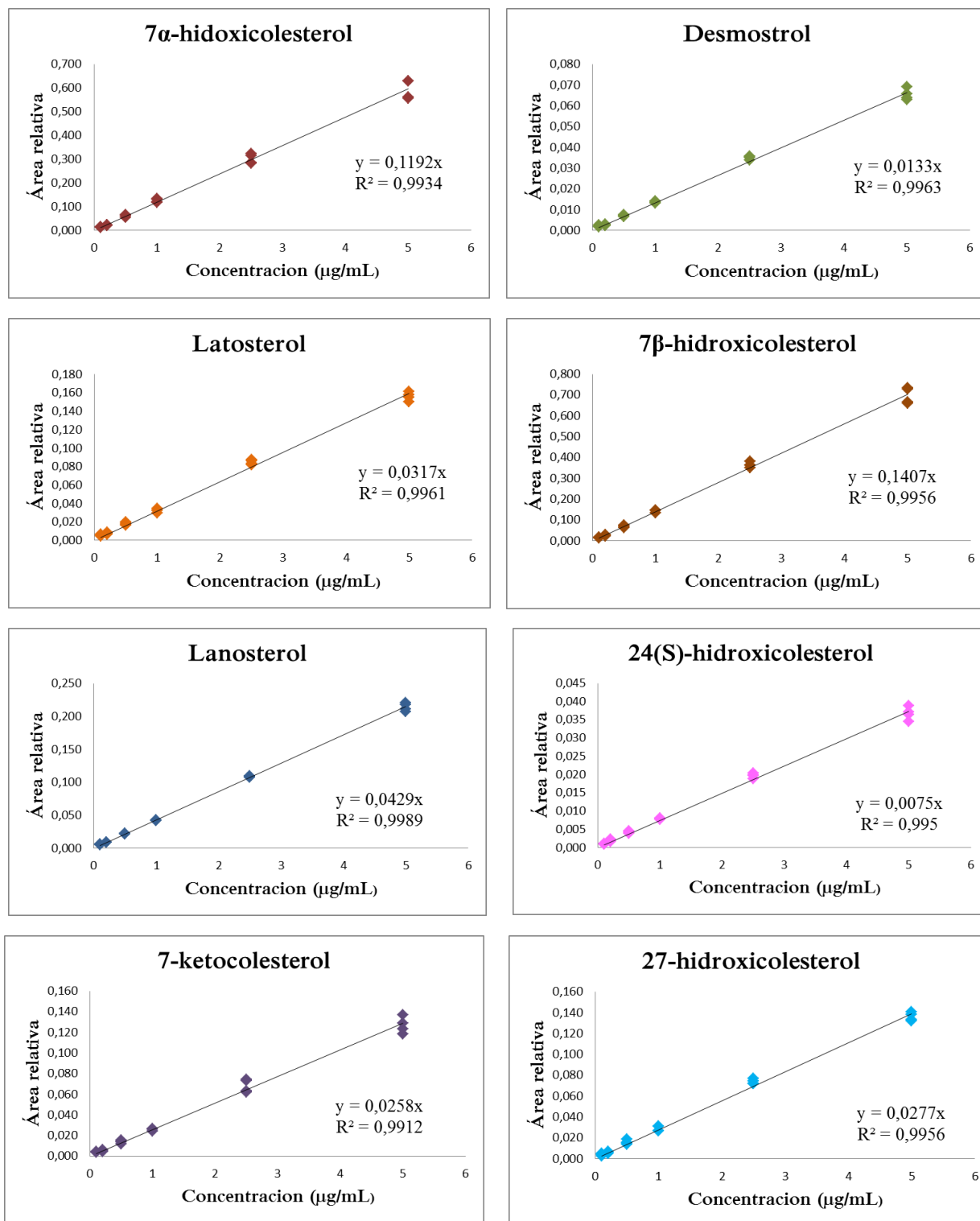


Figura I.1.4. Recta de calibrados de cada uno de los metabolitos de interés.

En la figura I.1.4 se muestran las rectas de calibrado obtenidas para cada uno de los metabolitos mediante los patrones de calibrado externos introducidos al principio y al final de la secuencia correspondiente con la tanda analítica en las que se inyectaron las muestras de plasma de ratón, por lo que corresponde con las rectas de calibrado usadas para el cálculo de la concentración de cada uno de los precursores y derivados del colesterol en muestras de plasma de ratón.

A partir de estas rectas de calibrado, o del correspondiente factor de respuesta, se calcularon todos los datos de concentración de las muestras de ratón que aparecen en la tabla 3 del artículo científico.

I.4. HOJA DE CONTROL DE PATRONES Y CONTROL DE CALIDAD

En este ANEXO se muestra la hoja de control de peso que se hizo para el control de los patrones utilizados en esta investigación.

En ella se toma nota del nombre de la disolución patrón que va a ser usada, así como de la fecha en la que se prepara y a continuación se toma nota del peso final (P.F.) teniendo en cuenta que corresponde con el peso de la disolución, en el recipiente en el que va a ser conservado, incluyendo el tapón.

La siguiente vez en la que se va a usar dicho patrón, se toma registro de la fecha de uso, el peso inicial del recipiente con tapón (P.I.), y se calcula la diferencia que existe entre el peso final de la última vez en que se usó la disolución patrón y el peso inicial que tiene ahora (P.F. - P.I.). Esta diferencia indicará el volumen de disolvente que se ha evaporado, y por tanto, la cantidad de metanol en la que tenemos que reconstituir el mismo.

A continuación, se lleva a cabo el tarado de la balanza analítica y se indica el peso de disolvente añadido (P.D.). A continuación, se suma esa cantidad (P.D.) al peso inicial que teníamos (P.I.), dando así el peso total (P.T.).

Una vez anotados todos los pesos, se toma el recipiente, se agita y se extrae la cantidad de la disolución patrón con la que se desee trabajar. Una vez finalizado todo este proceso, se vuelve a hacer constancia del peso final que queda para volver a operar de la misma forma la siguiente vez que se vaya a trabajar con dicha disolución.

AnAMA

Análisis en Alimentación
y Medioambiente

Control de peso PATRONES	Versión: 1.0
Grupo Investigación FMQ-232	Hoja N°:

• **Datos identificativos:**

PATRÓN:

Fecha de preparación:

Control de peso:

Fecha	P.I.	P.F. – P.I.	P.D.	P.T.	P.F.

P.I: Peso Inicial.

P.D: Peso del Disolvente añadido.

P.F: Peso Final del vial después de tomar la alicuota.

P.T: Peso Total (P.I. + P.D.)

Observaciones:

Para llevar a cabo el control de calidad, se recurre a mezcla de **patrones externos** de los analitos a diferentes concentraciones, que fueron **inyectados a lo largo de la secuencia cromatográfica**. Lo que se hace es calcular el factor de respuesta para cada analito y representarlo frente al orden de inyección del mismo a lo largo de la secuencia. La figura I.1.5 muestra dichos gráficos de control en la tanda en la que se inyectan muestras de ratón. Los **huecos** que aparecen a lo largo del orden de inyección, indican que ese dato **supera el 10% de aceptación** fijada.

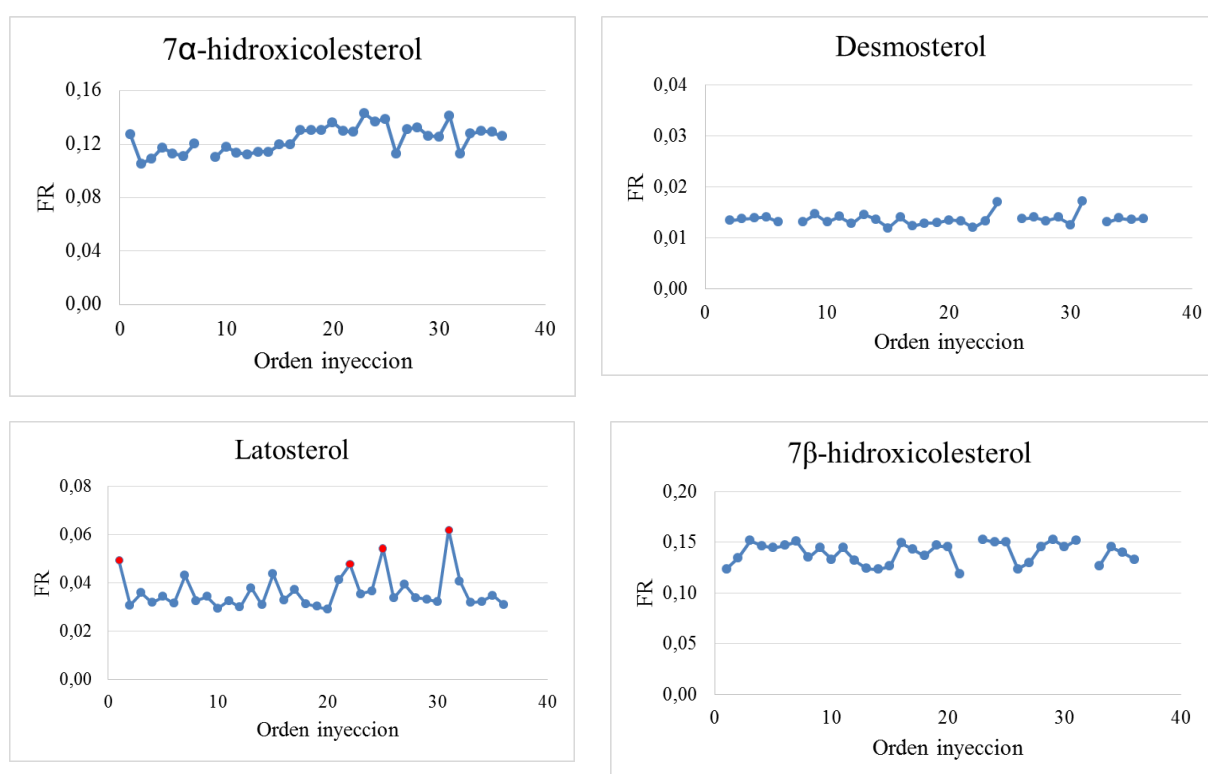


Figura I.1.5. Gráficos de control de calidad de cada uno de los analitos mediante patrones externos con el cálculo del factor de respuesta (FR) a lo largo de la secuencia de inyección de la tanda de muestra de plasma de ratón.

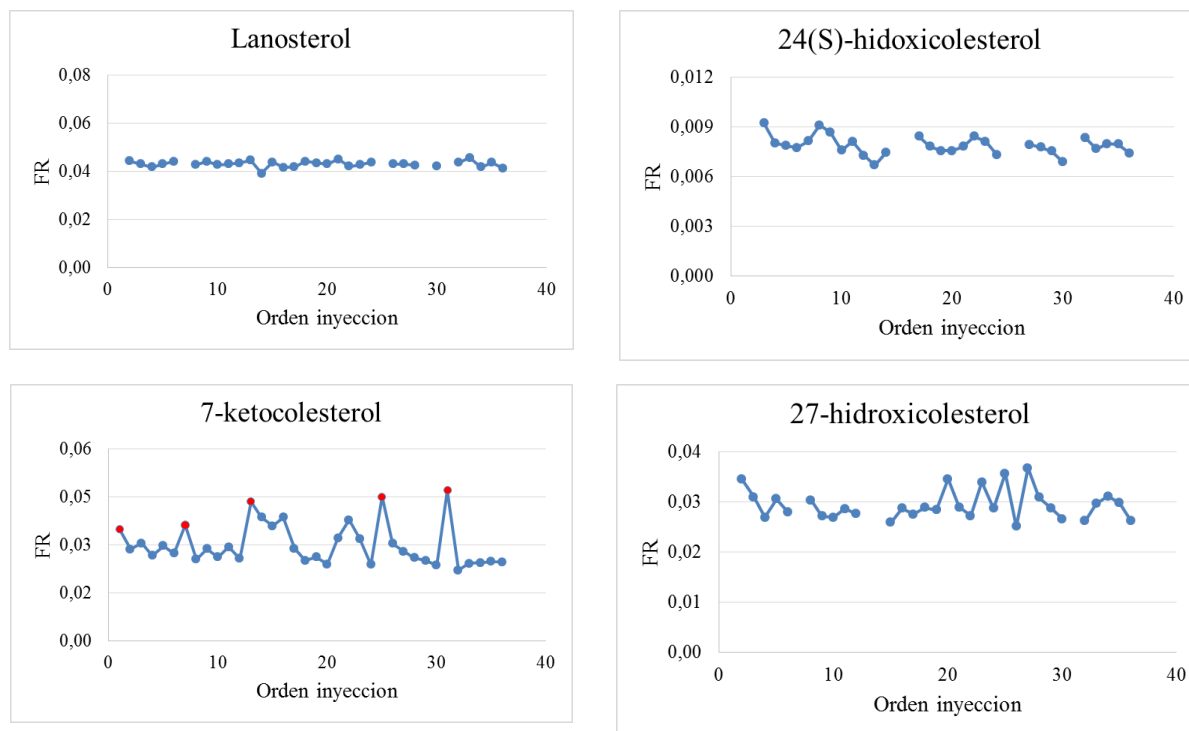


Figura I.1.5. (continuación).

A la vista de los resultados obtenidos de factor de calibración se puede concluir que el método que se ha llevado a cabo es adecuado para los analitos de interés.

El **lanosterol** y **7-ketocolesterol** se encuentran **fuera de control** en este estudio, puesto que la mayoría de los datos se salen fuera del 10% de aceptación. Cabe destacar que dichos analitos presentaron cierta **dificultad** a la hora de cuantificar y obtener el área bajo la curva, puesto que había **irregularidades** en la obtención del pico cromatográfico en cada uno de sus segmentos, llegando a aparecer el **pico cortado**.

CAPÍTULO 2

"Huellas dactilares" cromatográficas



CAPÍTULO 2

2.1. PRESENTACIÓN

Las "huellas dactilares" instrumentales son señales que contienen la información de interés y que es característica de una determinada muestra en relación a los compuestos o metabolitos presentes en ella. Ocurre lo mismo que la huella dactilar humana, es característica y biunívoca de cada individuo.

El **tratamiento de datos quimiométrico** es capaz de extraer la información presente en la "huella dactilar" instrumental por medio de diferentes modelos con los que se identifican relaciones y vínculos entre las muestras o grupos de muestras previamente definidas, que permiten su agrupamiento, clasificación, cuantificación, etc.

Este capítulo es el **punto de inflexión entre las dos líneas de investigación** que se han llevado a cabo a lo largo de la Tesis Doctoral. Así, se pasa de estudios en los que se aplica la cromatografía clásica de detección y cuantificación de los analitos, donde se han utilizado muestras de sangre en una aplicación en el campo de la salud; a estudios donde se emplea la "huella dactilar" cromatográfica con dos aplicaciones distintas en el campo del aceite de oliva virgen.

Por tanto, el **objetivo** de este capítulo es compilar la información necesaria para entender en qué se basa la metodología de "huellas dactilares" instrumentales, así como el uso de algunas metodologías de tratamiento de datos quimiométricos y las aplicaciones que derivan de ellas.

2.2. INTRODUCCIÓN

Señal analítica

La **medida analítica** consta de dos componentes [1]: (i) la **señal**, que contiene la información de interés; y (ii) el **ruido**, que es información ajena indeseada, que degrada la exactitud y precisión del análisis y además establece el límite inferior de la cantidad de analito que se puede detectar y/o cuantificar.

La **señal analítica** es la *señal instrumental, que puede ser diferenciada de la señal de fondo y/o de la señal del blanco, y que es utilizada para obtener información analítica. A partir de la señal se pueden obtener diferentes datos* [2]. Esta señal puede ser específica o inespecífica, con la premisa de que ambas señales han de contener información útil.

Señal específica

Para poder obtener información de **componentes individuales o grupos de compuestos** formados por la misma **familia** es necesario obtener una **señal instrumental con información explícita**, que se conoce como **señal específica**.

Este tipo de señales son con las que se trabaja normalmente para la detección, identificación y cuantificación de metabolitos o compuestos de interés. En estos casos y por ejemplo, trabajando con una técnica instrumental separativa, se puede obtener la señal con una buena separación de los picos cromatográficos; o, en el caso en el que exista solapamiento de picos, con un detector que sea capaz de identificar los compuestos presentes en la muestra. En el caso en el que se obtenga una buena separación de los picos

[1] D. A. Skoog, F.J. Holler, T.A. Nieman (2000). "Principios de análisis instrumental", quinta edición, Mc Graw Hill, Capítulo 5, pp. 103-119.

[2] L. Cuadros-Rodríguez, L. Gámiz-Gracia, A. Carrasco-Pancorbo, C. Ruiz-Samblás (2013). "Glosario de términos analíticos". Primera edición, Grupo Regional Andaluz de la Sociedad Española de Química Analítica (GRASEQA).

cromatográficos, éstos han de estar bien resueltos y para ello la columna cromatográfica ha de ser eficaz. La eficacia de la columna se mide a través de los platos teóricos e indica la capacidad de la misma para proporcionar bandas estrechas. Para obtener más información de los compuestos presente en la muestra se recurre a la espectrometría de masas, usada como detector.

Para la obtención de este tipo de señales es necesario llegar a una situación de compromiso entre los picos bien resueltos y el tiempo que dura el análisis. Aun así, suelen presentar tiempos de análisis largos, sobre todo cuando se analizan muestras complejas con muchos analitos.

Señal inespecífica

Para poder trabajar con tiempos de análisis menores, se puede recurrir a la obtención de **señales instrumentales inespecíficas**. En este tipo de señales no es necesario obtener una señal clara y unívoca que diferencie la información de cada uno de los componentes de forma individual. La información está implícita en las muestras. Volviendo al ejemplo de una técnica cromatográfica, estas señales serían picos cromatográficos que no necesitan estar bien resueltos, pudiendo aparecer dos picos cromatográficos solapados, o incluso con "joroba". La gran ventaja es que se puede reducir mucho el tiempo de análisis, la única premisa es que la señal obtenida contenga información útil y representativa de la muestra, puesto que de no ser así, no se podrá obtener ninguna información relevante. Para asegurar esto se trabaja como mínimo con dos grupos de muestras: (1) ambos grupos de muestras han de tener características diferentes entre sí; (2) dentro de un mismo grupo de muestras las características han de ser similares. Esto es, por ejemplo, en el ámbito de ciencias de la salud, dos grupos de muestras pertenecientes a pacientes sanos y enfermos; en el ámbito de ciencias de los alimentos, dos grupos de muestras pertenecientes a aceites de diferentes variedades de aceituna. Lo que se busque en cada situación dependerá del problema analítico a resolver.

Datos analíticos

Un dato es un *valor numérico suministrado por un instrumento o sistema de medida (u obtenido de una señal instrumental previamente*

adquirida) que contiene la información relativa a la magnitud a medir [2]. A partir de la señal analítica se puede extraer un único dato o datos de diferente dimensionalidad dependiendo del tensor seleccionado (escalar, vector o matriz). De manera que, **cada instrumento analítico puede ser clasificado dependiendo del tipo de dato que proporciona**, tomando la nomenclatura de los tensores en matemática [3,4]. Un instrumento que genera **un dato numérico individual**, un solo dato por muestra, es un instrumento de orden cero (por ejemplo, pH-metro, balanza analítica, etc.), y corresponde con un **tensor de orden cero (escalar)**. Un instrumento que genera un conjunto de **múltiples medidas al mismo tiempo o para la misma muestra** es un instrumento de orden uno (por ejemplo, espectrómetros, cromatógrafos, etc.), el conjunto de datos constituye un **tensor de orden uno (vector)**. Un instrumento que **genera una tabla de datos con respecto a dos magnitudes** es un instrumento de orden dos (por ejemplo, la unión entre instrumentos como un cromatógrafo acoplado a un espectrómetro de masas, resonancia magnética nuclear de dos dimensiones, etc.), ya que se genera una **matriz** de datos bidimensional siendo su **tensor de orden dos**. No existe un límite para el orden de los instrumentos, pero es raro encontrar instrumentos por encima del orden tres, que son aquellos que generan un **cubo** de datos por cada muestra, dicho cubo constituye una matriz tridimensional que es un **tensor de orden tres**.

Los instrumentos de orden uno o superiores generan señales con un elevado número de datos, es importante no perder la información contenida en ellos, por lo que se recurre a *software* que permitan su obtención y tratamiento a través de diferentes **herramientas quimiométricas**.

"Huella dactilar" instrumental y técnicas instrumentales

La señal inespecífica se usa para trabajar con la metodología de "**huella dactilar**" instrumental, se obtiene información de la muestra sin

[3] R.G. Brereton (2009). "Chemometrics for pattern recognition", John Wiley and Sons, Ltd., Capítulo 2, pp. 15-46.

[4] K.S. Booksh, B.R. Kowalski (1994). "Theory of analytical chemistry", Analytical Chemistry, 66, 782-791.

necesidad de identificar ni cuantificar los compuestos presentes en ella. Siguiendo con el ejemplo del uso de la cromatografía como técnica instrumental, esta metodología de "huellas" se aleja de los esquemas convencionales de la cromatografía clásica, puesto que trabaja con las señales que no presentan una buena resolución pero que contienen toda la información de interés. Para ello es necesario recurrir a técnicas avanzadas de tratamiento de datos, como es el uso de la quimiometría, que permite extraer toda la información útil contenida en la señal analítica (o una parte seleccionada de ella).

También existe la opción de trabajar con "huellas dactilares" a partir de señales específicas, en estos casos sería para ampliar la información que se está buscando. Por ejemplo, en el caso de trabajar con muestras de pacientes sanos o enfermos, al aplicar la "huella dactilar" y las técnicas avanzadas de tratamiento de datos se podrían seleccionar uno o varios marcadores o picos característicos, e incluso parte de una "joroba" en un pico con uno o varios compuestos concretos al aplicar alguna técnica de detección específica.

A continuación se enumeran las **técnicas analíticas instrumentales** más usadas para la obtención de la "huella dactilar" [5,6,7,8]:

(a) **Espectroscópicas.** Estas técnicas se basan en la medida de la radiación electromagnética emitida o absorbida por la materia. A través de la radiación sobre el material y su energía radiada se obtiene información de los compuestos de la muestra. El espectro da la respuesta de interés en función de la longitud de onda o de la frecuencia de la radiación utilizada. Pueden obtenerse diferentes tipos de espectros en función de la naturaleza de la energía de radiación (infrarrojo, visible, ultravioleta, rayos X) y la naturaleza

-
- [5] J.L. Wolfender, G. Marti, A. Thomas, S. Bertrand (2015). "Current approaches and challenges for the metabolite profiling of complex natural extract", *Journal of Chromatography A*, 1382, 136-164.
 - [6] E. Capuano, S.M. van Ruth (2012). "QA: Fraud control for foods and other biomaterials by product fingerprinting", En: I. Akyar (Ed.), *Latest research into quality control*, Intech, pp. 111-143.
 - [7] D.I. Ellis, V.L. Brewster, W.B. Dunn, J.W. Allwood, A.P. Golovanov, R. Goodacre (2012). "Fingerprinting food: current technologies for the detection of food adulteration and contamination", *Chemical Society Reviews*, 41, 5569-5868.
 - [8] S. Moco, R.J. Bino, R.C.H. de Vos, J. Vervoort (2007). "Metabolomics technologies and metabolite identification", *Trends in Analytical Chemistry*, 26, 855-866.

de la interacción entre energía y materia (absorción, emisión, dispersión, resonancia). Las técnicas espectroscópicas más usadas para la obtención de "huella dactilar" son la **resonancia magnética nuclear e infrarroja**. La primera de ellas basada en la radiación de las frecuencias de radio o radiofrecuencias. Esta técnica se usa para identificar **determinados átomos** de los compuestos químicos presentes en una muestra (principalmente ^1H o ^{13}C), dando información de la estructura molecular. Por ejemplo, se ha usado en el ámbito de alimentos para la clasificación del origen geográfico en muestras de lentejas por la asignación de diferentes grupos metilos [9]. La segunda, espectroscopía de infrarrojo, utiliza la parte de la radiación infrarroja del espectro electromagnético. Esta técnica se usa para identificar los **grupos funcionales** de los compuestos químicos presentes en muestra. El espectro infrarrojo se obtiene midiendo la absorbancia de la luz infrarroja por muestra en función de la longitud de onda. Por ejemplo, se ha usado en la autenticación de aceite de oliva virgen extra italiano para la clasificación de una denominación de origen protegida [10].

(b) **Cromatográficas**: se trata de la etapa separativa previa a la detección con el objetivo de identificar **un compuesto o una familia de compuestos** presentes en una muestra. Por ejemplo, para la detección de compuestos de interés biológicos en plantas propias de la medicina China tradicional [11].

(c) **Espectrometría de masas**: esta técnica se puede usar con infusión directa de la muestra o acoplado a otro tipo de técnicas, como detector. Esta técnica determina la relación m/z de los compuestos químicos presentes en una muestra. En el espectro de masas se representa la intensidad relativa de la señal frente a la relación m/z de los compuestos o componentes presentes en la muestra. Por ejemplo, el uso de la infusión directa en MS para descubrir

-
- [9] F. Longobardi, V. Innamorato, A. Di Gioia, A. Ventrella, V. Lippolis, A.F. Logrieco, L. Catucci, A. Agostiano (2017). "Geographical origin discrimination of lentils (*Lens culinaris* Medik.) using ^1H NMR fingerprinting and multivariate statistical analyses", *Food Chemistry*, 237, 743-748.
- [10] M. Casale, P. Oliveri, C. Casolino, N. Sinelli, P. Zunin, C. Armanino, M. Forina, S. Lanteri (2012). "Characterisation of PDO olive oil Chianti Clasico by non-selective (UV-visible, NIR and MIR spectroscopy) and selective (fatty acid composition) analytical techniques", *Analytica Chimica Acta*, 712, 56-63.
- [11] Y. Tao, X. Gu, W. Li, H. Wu, B. Cai, B. Zhang (2017). "Techniques for biological fingerprinting of traditional Chinese medicine", *Trends in Analytical Chemistry*, 97, 272-282.

potenciales marcadores del diagnóstico precoz de enfermedades a partir de sus metabolitos [12].

Existen otras técnicas analíticas para la obtención de la "huella dactilar" como es la **calorimetría diferencial de barrido**, que mide la cantidad de calor requerida para aumentar la temperatura de una muestra con respecto a un material de referencia. El espectro térmico de la muestra viene dado por la cantidad de calor absorbido/liberado en función de la temperatura. Estas propiedades físicas dependen de las propiedades estructurales y de composición de las muestras. Por ejemplo, en el ámbito de los alimentos, se han usado en muestras de jamón ibérico curado para la detección de fraude en la certificación de dichas muestras usando la técnica calorimétrica diferencial de barrido para diferenciar las categorías comerciales de jamón [13].

Tras la obtención de las "huellas dactilares" instrumentales se pasa al tratamiento quimiométrico de los datos.

Quimiometría

La quimiometría es la ciencia que usa técnicas estadísticas para reducir la dimensionalidad de los datos multivariados y extraer de ellos la información química relevante [8].

Existe un gran número de **herramientas quimiométricas**, entre ellas las técnicas de reconocimiento de pautas, con las que es posible identificar relaciones y vínculos entre objetos (o muestras) que permitan su agrupamiento y clasificación según la similitud de su "huella dactilar" instrumental [14].

[12] R. González-Domínguez, A. Sayago, A. Fernández-Recamales (2017), "Direct infusión mass spectrometry for metabolomic phenotyping of diseases", *Bioanalysis*, 9, 131-148..

[13] L. Bayés-García, A. Tres, S. Vichi, T. Calvet, M.A. Cuevas-Diarte, R. Codony, J. Boatella, J. Caixach, S. Ueno, F. Guardiola (2016). "Authentication of Iberian dry-cured ham: new approaches by polymorphic fingerprint and ultrahigh resolution mass spectrometry", *Food Control*, 60, 370-377.

[14] P. Oliveri, R. Simonetti (2016). "Chemometrics for food authenticity applications", En: G. Downey (Ed.), *Advances in food authenticity testing*, Elsevier, pp. 701-728.

Los datos con los que se trabajan en esta disciplina son datos multivariantes, que pueden ser vectores o matrices procedentes de instrumentos de orden uno o dos. En el caso de trabajar con vectores de datos, éstos se unen en una matriz de dimensiones ($m \times n$), donde "m" corresponde con el número de filas, el número de objetos o muestras analizadas; y "n" es el número de columnas o variables, que corresponde con el número total de datos procedentes de la señal [15]. En el caso de trabajar con matrices de datos, estos pueden tener diferentes direcciones. En el caso de trabajar con tres vías, tendríamos una matriz de datos ($m \times n$), donde cada fila "m" corresponde con lo que se denomina "capa o losa" y que es otra matriz de datos en lugar de un vector [16].

Para obtener la información de interés o relevante, se puede trabajar con toda la señal completa o con una parte (o varias) de ella.

Cualquier tratamiento quimiométrico requiere de una etapa previa, denominada **pre-procesado**, donde se modifican ligeramente las señales analíticas obtenidas para que no conduzca a resultados erróneos. No existe una regla general que permita preparar la información de los datos pero normalmente, e independientemente del tipo de instrumento, se pueden realizar los siguientes [14,17,18]: (i) **corrección de línea base**, para los desplazamientos o desviaciones que pudieran darse en la línea base (por ejemplo la corrección de la deriva) entre las señales de las distintas muestras o entre réplicas de una misma muestra; (ii) **suavizado**, se centra en la eliminación de ruido en la señal, es importante no poner demasiado suavizado porque se podría perder información de la intensidad de la señal analítica obtenida a priori; (iii) **centrado en la media**, corrige las diferencias entre variables, variando los datos en torno a la media para fijar las variaciones

-
- [15] R. Leardi (2017). "Data handling", En: F. Toldrá, L.M.L. Nollet (Eds.), *Advances in food diagnostics*, segunda edición, Wiley Blackwell, pp. 401-430.
- [16] J.M. Amigo, F. Marini (2013). "Multiway methods", En: F. Marini (Ed.), *Chemometrics in food chemistry*, Elsevier, pp. 265-313.
- [17] R.G. Brereton (2009). "Chemometrics for pattern recognition", John Wiley and Sons, Ltd., Capítulo 4, pp. 107-176.
- [18] J.M. Andrade-Garda, A. Carlosena-Zubieta, R. Bosqué-Martí, J. Ferré-Baldrich (2013). "Partial Least-Squares Regression". En: J.M. Andrade-Garda, "Basic chemometric techniques in atomic spectroscopy", segunda edición, RSC Analytical Spectroscopy Monograph, pp. 280-347.

entre muestras, se calcula un vector medio de los datos en cada una de las columnas restando a cada variable el valor medio de la columna correspondiente a dicha variable; (iv) **autoescalado**, cada columna se divide por su desviación estándar, suele usarse cuando las variables se miden en diferentes escalas y corrige cambios en las variaciones de intensidad.

Además de estos pre-procesados generales existen otros específicos en función del tipo de señal instrumental con la que se esté trabajando. Por ejemplo, en el caso de señales cromatográficas se requiere de un pre-procesado de **alineamiento de los picos**, donde se corrigen variaciones en los tiempos de retención de los picos, que pueden variar fácilmente por unos segundos de un análisis a otro [19,20].

En las técnicas de reconocimiento de pautas, se pueden diferenciar diferentes niveles entre los métodos supervisados y no supervisados, en la figura 2.1 se muestra un esquema de las técnicas quimiométricas divididas en ambas técnicas [15,21].

Las técnicas de **reconocimiento de pautas no supervisadas** no requieren ningún conocimiento previo de los datos de la muestra. Por lo que se puede trabajar con este tipo de métodos cuando no se dispone de información previa o en caso de que se disponga de ella pero sea interesante estudiar el agrupamiento para observar las relaciones naturales de los datos con este tipo de métodos, previo a las técnicas de reconocimiento de pautas supervisadas.

I. Análisis de componentes principales (PCA)

Una de las técnicas principales de **análisis exploratorio de datos** es el **PCA**, es uno de los métodos más utilizados en el análisis multivariable y suele realizarse siempre antes de comenzar con cualquier análisis quimiométrico.

-
- [19] G. Tomasi, F. Savorani, S.B. Engelsen (2011). "Icoshift: an effective tool for the alignment of chromatographic data", *Journal of Chromatography A*, 1218, 7832-7840.
- [20] G. Tomasi, F. Van der Berg, C. Andersson (2004), "Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data", *Journal of Chemometrics*, 18, 231-241.
- [21] J.N. Miller, J.C. Miller (2010). "Statistics and chemometrics for analytical chemistry", Sexta edición, Pearson, Capítulo 8, pp. 221-250.

PCA reduce la dimensionalidad de los datos y extrae la parte más relevante de la información [14,22]. Si "X" es una matriz de datos con "m" filas (objetos o muestras) y "n" columnas (variables), PCA descompone "X" como suma de términos en vectores, de menor o igual dimensión que "X":

- ✓ Los vectores "scores" (puntuaciones) contienen la información de cómo las muestras se relacionan unas con otras.
- ✓ Los vectores "loadings" (pesos o ponderales) contienen información de cómo las variables se relacionan unas con otras.

Cada **componente principal** (PC) consiste en un vector "scores" y uno "loadings". PC1 tiene la varianza más alta posible. PC2 es ortogonal (en scores y loadings) y tiene la varianza siguiente más alta posible, y así sucesivamente [23].

II. Análisis factorial (FA)

El objetivo de este método es caracterizar las variables en la matriz "X" en un número pequeño de "p" factores comunes y un conjunto de errores o factores específicos "e". Se asume que los factores comunes y los factores específicos son independientes unos de otros [22]. La matriz de factores se interpreta de forma similar al análisis de componentes principales.

III. Análisis de clúster (CA)

El **análisis exploratorio de clúster** incluye una familia de métodos capaces de **identificar agrupaciones entre muestras y/o variables por medio de una medida de similitud**. La similitud entre las muestras se define como la función inversa de la distancia entre ellas y la de las variables se puede estimar mediante el coeficiente de correlación [14,22]. El análisis de clúster se representa mediante un **dendrograma**, un diagrama en forma de árbol que

[22] R.G. Brereton (2007). "Applied chemometrics for scientists", John Wiley and Sons, Ltd., Capítulo 5, pp. 145-191.

[23] R. Leardi (2003). "Chemometrics in data analysis", En: M. Lees (Ed.), Food authenticity and traceability, CRC Press, pp. 315-336.

organiza los objetos en función de sus similitudes, desgranándolo hasta llegar a un nivel deseado.

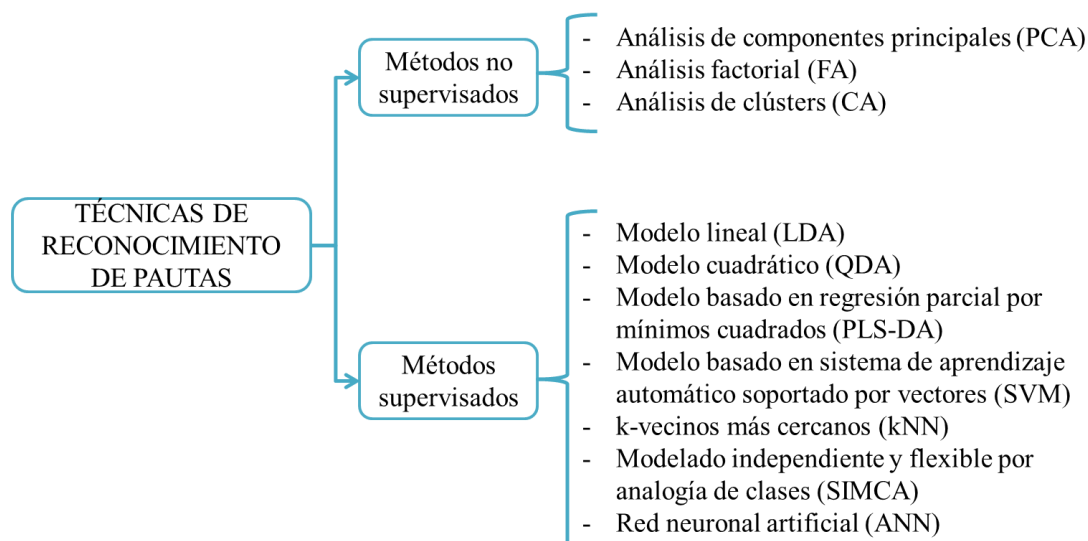


Figura 2.1. Esquema de diferentes técnicas quimiométricas.

Las **técnicas de reconocimiento de pautas supervisadas** están centradas en clasificar o cuantificar una propiedad determinada de la muestra. En este tipo de modelos, la información previa sobre la pertenencia de las muestras a una determinada clase o grupo se utiliza para clasificar muestras nuevas desconocidas. Los modelos de regresión/calibración establecen una relación funcional entre una propiedad cuantitativa de la muestra (variable dependiente) y una variable independiente multivariable.

Para caracterizar una muestra respecto a una propiedad se recurre a técnicas de clasificación y modelado de clase. Las clases (o categorías) pueden definirse como un conjunto de muestras que tienen en común al menos una característica, atributo, cualidad y/o propiedad [24].

La construcción de un modelo de clasificación consta de cuatro pasos esenciales [14]: (1) selección de un conjunto de entrenamiento, que consiste en objetos pertenecientes a una o varias clases conocidas (o valores

[24] L. Cuadros-Rodríguez, E. Pérez-Castaño, C. Ruiz-Samblás (2016), "Quality performance metrics in multivariate classification methods for qualitative analysis", Trends in Analytical Chemistry, 80, 612-624.

cuantitativos conocidos de una propiedad de la muestra) para los que se miden las variables; (2) selección de variables, se mantienen las variables que contienen información, las que dan ruido o no tienen poder discriminante, se eliminan; (3) construcción del modelo utilizando el conjunto de entrenamiento y sus categorías conocidas (o valores cuantitativos de la propiedad de la muestra); (4) validación del modelo, con el fin de evaluar la fiabilidad de la clasificación obtenida mediante un grupo del conjunto de entrenamiento (validación interna cruzada) o un conjunto independiente de muestras no utilizadas en el entrenamiento (validación externa).

Las técnicas de modelado de clase se caracterizan por definir una región cerrada en una sola clase, de manera que una nueva muestra entrará o no dentro de ese espacio en función de si las características con esa clase de interés coinciden o no. Las siguientes constituyen algunas de ellas:

I. Análisis discriminante de regresión por mínimos cuadrados parciales (PLS-DA)

PLS es un método de regresión basado en combinaciones lineales de las variables latentes de predicción. Para establecer los límites de la clase se realiza un análisis discriminante con el fin de desarrollar y optimizar el número de variables latentes en el modelo PLS. Se seleccionan combinaciones lineales de las variables de predicción, de manera que estén altamente correlacionadas con las variables de respuesta [21].

La clasificación se basa en construir un modelo de calibración con diferentes clases a las que se les asocia un valor numérico distinto, por ejemplo 0 y 1 en caso de una clasificación binaria. La asignación de clase se realiza prediciendo el valor numérico de la muestra [25,26].

[25] M. Bevilacqua, R. Nescatelli, R. Bucci, A.D. Magri, A.L. Magri, F. Marini (2014). "Chemometric classification techniques as a tool for solving problems in analytical chemistry", *Journal of AOAC International*, 97, 19-28.

[26] D. Ballabio, V. Consonni (2013). "Classification tools in chemistry. Part 1: linear models. PLS-DA", *Analytical Methods*, 5, 3790-3798.

II. Sistema de aprendizaje automático soportado por vectores (SVM-C)

SVM es un método de aprendizaje. Funciona mediante un modelo de regresión para la construcción de hiperplanos en un espacio multidimensional que separa las diferentes clases de objetos [27].

La clasificación la realiza de forma similar al método PLS-DA antes descrito [28].

III. k vecinos cercanos (kNN)

kNN es un método de clasificación basado en la similitud entre objetos mediante el cálculo de la distancia entre ellos. Se usa cuando los grupos de muestras no pueden ser separados en un plano [21].

La muestra desconocida será asignada a una de las clases del conjunto de entrenamiento en el que el número de k-vecinos sea mayor, siendo k un número entero impar [29].

IV. Modelado independiente y flexible por analogía de clases (SIMCA)

Este método está basado en la construcción de componentes principales que encajan alrededor de cada una de las clases de muestras en el conjunto de datos [30], de manera que realiza tantos modelos de componentes principales como clases de entrada en el estudio.

-
- [27] R.G. Brereton (2009). "Chemometrics for pattern recognition", John Wiley and Sons, Ltd., Capítulo 5, pp.145-191.
- [28] J. Luts, F. Ojeda, R. Van de Plas, B. de Moor, S. Van Huffel, J.A.K. Suykens (2010). "A tutorial on support vector machine-based methods for classification problems in chemometrics", *Analytica Chimica Acta*, 665, 129-145.
- [29] B.K. Alsberg, R. Goodacre, J.J. Rowland, D.B. Kell (1997) "Classification of pyrolysis mass spectra by fuzzy multivariate rule induction-comparison with regression, K-nearest neighbor, neutral and decision-tree methods", *Analytica Chimica Acta*, 348, 389-407.
- [30] D. Livingstone (2009). "A practical guide to scientific data analysis", John Wiley and Sons, Ltd., capítulo 7, pp. 187-218.

La clasificación se lleva a cabo desde la menor distancia del objeto al centro de cada espacio del componente principal, asignando así la muestra desconocida a una de las clases [25].

V. Red neuronal artificial (ANN)

Las redes neuronales intentan imitar la operación de las neuronas en el cerebro, por lo que tiene una serie de capas enlazadas denominadas neuronas artificiales. Las variables medidas se llevan a una capa de entrada y se procesan mediante una serie de operaciones matemáticas intermedias (ocultas) para producir una o más salidas. Se usa el conjunto de entrenamiento para observar las discrepancias a la salida de la red y ajustar la red neuronal hasta alcanzar el grado de precisión requerido [21].

Esta técnica quimiométrica es útil puesto que no asume ninguna relación lineal inicial, como ocurre con el PLS, por ejemplo.

Aplicaciones

En este apartado se dan unas pinceladas acerca de las aplicaciones que existen en el ámbito de ciencias de la salud y de los alimentos, los dos enfoques abordados en esta Tesis Doctoral. Las aplicaciones están centradas en el uso de diferentes herramientas quimiométricas sobre señales instrumentales específicas o inespecíficas.

Las herramientas quimiométricas pueden usarse como método de cribado (indicando: si/no; presencia/ausencia). A partir de aquí se pueden seleccionar uno o varios marcadores, picos de la señal instrumental específicos o características concretas de la señal que permitan dicha clasificación (presencia/ausencia). Por lo que, con el uso de las "huellas dactilares" instrumentales permite un análisis rápido de las muestras, y a partir de ahí se pueden usar métodos instrumentales de detección y/o cuantificación de los compuestos de interés.

Por ejemplo, en el ámbito de la salud, con este tipo de metodología se podría diagnosticar una enfermedad con grupos de pacientes sanos o enfermos. Por otro lado, en el ámbito de ciencias de los alimentos, y más concretamente, en la autenticación, se pueden clasificar las muestras conociendo su origen

geográfico, así como si presentan algún tipo de adulteración o fraude. Esta última aplicación ha sido la seleccionada para la elaboración de los artículos científicos publicados y que se presentan en este capítulo.

A continuación se hace una breve introducción de diferentes tipos de aplicaciones en los dos ámbitos de trabajo de esta Tesis:

- **Diagnóstico de una enfermedad**

El diagnóstico de una enfermedad mediante la metodología de "huella dactilar" depende de muchos factores como: la persona, su edad, el sexo, la genética, el grupo étnico, etc. Además los signos químicos de la enfermedad en una etapa temprana pueden ser muy sutiles. Por lo que a veces resulta difícil obtener muestras representativas y controles adecuados [31].

Una de las técnicas analíticas que pueden usarse son las técnicas espectroscópicas en combinación con diferentes métodos quimiométricos, por ejemplo en la detección de leucemia con la técnica de espectroscopía infrarroja de transformada de Fourier [32,33].

Además existen otras enfermedades en las que se estudia su diagnóstico mediante técnicas como la espectroscopía infrarroja y de Raman para cáncer cervical, leucemia, cáncer de próstata, artritis, diabetes [34].

También las técnicas analíticas de espectrometría de masas [35,36] y de cromatografía [37], que son de amplia aplicación en el diagnóstico de enfermedades de cáncer.

-
- [31] R.G. Brereton (2007). "Applied chemometrics for scientists", John Wiley and Sons, Ltd., Capítulo 10, pp.288-318.
- [32] J. Babrah, K. McCarthy, R.J. Lush, A.D. Rye, C. Bessant, N. Stone (2009), "Fourier transform infrared spectroscopic studies of T-cell lymphoma, B-cell lymphoid and myeloid leukaemia cell lines", *Analyst*, 134, pp. 763-768.
- [33] C.P. Schultz, K.Z. Liu, J.B. Johnston, H.H. Mantsch (1996). "Study of chronic lymphocytic leukemia cells by FT-IR spectroscopy and cluster analysis", *Leukemia Research*, 20, 649-655.
- [34] D.I. Ellis, R. Goodacre (2006), "Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and Raman spectroscopy", *Analyst*, 131, 875-885.
- [35] A.G. Ramos, A.P. Antón, M.D. Sánchez, J.L.P. Pavón, B.M. Cordero (2017), "Urinary volatile fingerprint based on mass spectrometry for the discrimination of patients with lung cancer and controls", *Talanta*, 174, 158-164.
- [36] K. Awedat, I. Abdel-Qader, J.R. Springstead (2017), "Prostate cancer recognition based on mass spectrometry sensing data and data fingerprint recovery", *Biomedical Signal Processing and Control*, 33, 392-399.

Este tipo de metodología ha cobrado especial importancia en los últimos años en el diagnóstico de diferentes enfermedades de cáncer con diferentes técnicas analíticas y quimiométricas para su resolución. Aunque el problema que presentan es la necesidad de obtener un número elevado de muestras representativo de la población.

- Alimentos

En el ámbito de los alimentos existen un gran número de aplicaciones, como son estudios de adulteración, clasificación de ingredientes, origen geográfico de las muestras, etc. [38]. Además existe un gran número de alimentos que se pueden estudiar. Por ejemplo, el uso de la técnica analítica de espectroscopía del infrarrojo cercano para clasificar muestras de diferentes tipos de aceites vegetales [3,39]. Siguiendo con el aceite como muestra, y centrandolo en el ejemplo en el aceite de oliva virgen, existen gran variedad de estudios, por ejemplo, la cuantificación de aceite de oliva y de otros aceites vegetales mediante la "huella dactilar" usando técnicas de cromatografía de líquidos y de gases [40,41]. Otra de las aplicaciones recientes que se están llevando a cabo en los últimos años, es la clasificación del aceite de oliva en una de sus tres categorías (aceite de oliva virgen extra, virgen o lampante) mediante técnicas analíticas y los compuestos orgánicos volátiles como complemento al análisis sensorial. Por ejemplo: el uso de la fuente de ionización química a presión atmosférica en combinación con la cromatografía de gases acoplada a

-
- [37] R. Taware, K. Taunk, J.A.M. Pereira, R. Dhakne, N. Kannan, D. Soneji, J.S. Câmara, H.A. Nagarajaram, S. Rapole (2017), "Investigation of urinary volatome alterations in head and neck cancer: a non-invasive approach towards diagnosis and prognosis", *Metabolomics*, 13:111, 1-11.
- [38] R.G. Brereton (2007). "Applied chemometrics for scientists", John Wiley and Sons, Ltd., Capítulo 13, pp.351-368.
- [39] A.M. Jiménez-Carvelo, M.T. Osorio, A. Koidis, A. González-Casado, L. Cuadros-Rodríguez (2017), "Chemometric classification and quantification of olive oil blends with any edible vegetable oils using FTIR-ATR and Raman spectroscopy", *LWT- Food Science and Technology*, 86, 174-184.
- [40] A.M. Jiménez-Carvelo, A. González-Casado, L. Cuadros-Rodríguez (2017), "A new analytical method for quantification of olive and palm oil in blends with other vegetable edible oils based on the chromatographic fingerprints from the methyl-transesterified fraction", *Talanta*, 164, 540-547.
- [41] C. Ruiz-Samblás, F. Marini, L. Cuadros-Rodríguez, A. González-Casado (2012), "Quantification of blending of olive oils and edible vegetable oils by triacylglycerol fingerprint gas chromatography and chemometric tools", *Journal of Chromatography B*, 910, 71-77.

un detector de espectrometría de masas [42], el uso de la espectroscopía infrarroja con transformada de Fourier [43], así como la cromatografía de gases acoplada a un espectrómetro de masas de movilidad iónica [44].

Existe un gran número de alimentos a los que se les puede aplicar esta metodología. En los artículos que se presentan en los apartados 2.3 y 2.4, se centran en la autenticación de alimentos y en la diferencia que existe entre perfil y "huella dactilar" instrumental.

2.3. ARTÍCULO 1

Chromatographic fingerprinting: An innovative approach for food 'identification' and food authentication – A tutorial

(Artículo publicado en Enero 2016 en la revista: Analytica Chimica Acta)

Analytica Chimica Acta 909 (2016) 9–23



Contents lists available at [ScienceDirect](#)

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca



Tutorial

Chromatographic fingerprinting: An innovative approach for food 'identification' and food authentication – A tutorial



Luis Cuadros-Rodríguez, Cristina Ruiz-Samblás, Lucia Valverde-Som, Estefanía Pérez-Castaño, Antonio González-Casado*

Department of Analytical Chemistry, University of Granada, c/ Fuentenueva, s.n., E-18071 Granada, Spain

- [42] C. Sales, M.I. Cervera, R. Gil, T. Portolés, E. Pitarch, J. Beltran (2017). "Quality classification of Spanish olive oils by untargeted gas chromatography coupled to hybrid quadrupole-time of flight mass spectrometry with atmospheric pressure chemical ionization and metabolomics-based statistical approach", *Food Chemistry*, 216, 365-373.
- [43] A. Hirri, M. Bassbasi, S. Platikanov, R. Tauler, A. Oussama (2016). "FTIR spectroscopy and PLS-DA classification and prediction of four commercial grade virgin olive oils from Morocco", *Food Analytical Methods*, 9, 974-981.
- [44] R. Garrido-Delgado, M.M. Dobao-Prieto, L. Arce, M. Valcárcel (2015), "Determination of volatile compounds by GC-IMS to assign the quality of virgin olive oil", *Food Chemistry*, 187, 572-579.

Abstract

Fingerprinting methods describe a variety of analytical methods that provide analytical signals related to the composition of foodstuffs in a non-selective way such as by collecting a spectrum or a chromatogram. Mathematical processing of the information in such fingerprints may allow the characterisation and/or authentication of foodstuffs. In this context, the particular meaning of 'fingerprinting', in conjunction with 'profiling', is different from the original meanings used in metabolomics. This fact has produced some confusion with the use of these terms in analytical papers. Researchers coming from the metabolomic field could use 'profiling' or 'fingerprinting' on a different way to researchers who are devoted to food science. The arrival of an eclectic discipline, named 'foodomics' has not been enough to allay this terminological problem, since the authors keep on using the terms with both meanings. Thus, a first goal of this tutorial is to clarify the difference between both terms. In addition, the chemical approaches for food authentication, i.e., chemical markers, component profiling and instrumental fingerprinting, have been described. A new term, designated as 'food identification', has been introduced in order to complete the life cycle of the chemical-based food authentication process. Chromatographic fingerprinting has been explained in detail and some strategies which could be applied has been clarified and discussed. Particularly, the strategies for chromatographic signals acquisition and chromatographic data handling are unified in a single framework. Finally, an overview about the applications of

chromatographic (GC and LC) fingerprints in food authentication using different chemometric techniques has been included.

Keywords

Chromatographic fingerprinting; food analysis; food 'identification'; food authentication.

1. Introduction

Fingerprinting is described as a variety of analytical techniques or methods which can inform about the composition of some foods in a non-selective way with the main aim of characterizing or authenticating the food. However this term, together with the term 'profiling', comes from the metabolomic terminology. The concept of 'metabolite profiling' was firstly proposed by Horning and Horning, in 1971, in relation to the determination of metabolites using mass spectrometry (MS) [1,2]. The uses of the term 'metabolomic fingerprinting' is more recent and it may have been introduced by Fiehn at the end of the twentieth century [3] from the idea of infrared fingerprint.

The metabolome of a biological system refers to the complement of all low molecular weight metabolites in that system (typically <1500 daltons). Metabolomics refers to the quantitative analysis of the metabolome, or a fraction of this, by the use of advanced analytical techniques. However, it also involves sampling, sample preparation, chemical analysis and data analysis. Whilst the measurement and

quantification of individual or small numbers of metabolites is well established in biochemistry, metabolomics differs from more analyses in the number of classes of metabolites being detected, the range of analytical techniques being employed and the need for advanced signal processing and bioinformatics tools [4,5].

Generally, the metabolomics approaches can be generally grouped as 'profiling (targeted)' or 'fingerprinting (untargeted)' strategies [6]. While 'profiling' involves the analysis of a group of related metabolites, which are in most cases identified and quantified, 'fingerprinting' is based on the determination of as many metabolites as possible without necessarily identifying or quantifying the compounds present. The 'metabolomics (targeted) profiling' provides direct functional information, and the data can be integrated into metabolic models, whereas the 'metabolomics (untargeted) fingerprinting' provides a pattern that is used to classify samples based on provenance of either their biological relevance or origin by using a fingerprinting technology that is rapid but does not necessarily give specific metabolite information. According to these conceptions, 'metabolite profiling' is the analysis of a given set of known metabolites, for instance, a set of amino and organic acids, whereas 'metabolic fingerprinting' is an unspecific analysis of a test material, e.g., a set of mass peaks obtained by mass spectrometry.

However, a double distinction can also be applied for each case. The 'metabolite profiling' does not need to be quantitative, but often is at least semiquantitative, hence it is not possible to optimize the method

for all the known metabolites. However, the 'metabolite targeted analysis' is a quantitative analysis of metabolites participating in a specific part of the metabolism in a sample associated to specific pathways, for example, a particular enzyme system that would be directly affected by abiotic or biotic perturbation [7,8]. In a similar way, 'metabolite untargeted analysis' is currently differentiated from fingerprinting analysis. The 'metabolite untargeted analysis' is a first approximation to the complexity of the unknown metabolites in a test material and it should be seen as a first step before more concrete approaches as profiling analysis. This strategy, known as top-down strategy, avoids the need for a prior specific hypothesis on a particular set of metabolites and, instead, analyses the global metabolite profile [9]. The term 'fingerprinting' remain faithful to the initial mean, and in this approach the intention is not to identify each observed metabolite, but to compare patterns or 'fingerprints' of metabolites that change in response to disease, environmental or genetic alterations.

Due to all this and since there is no single analytical method, different terms are often used in the field of metabolomics, different authors have given different definitions along the bibliography and there are a lot of definitions in bibliography, as well as misuses. For this, these four terms are often used interchangeably; also some of them have been mixed up such us 'untargeted metabolomic profiling' as synonym of 'global profiling' [10], or 'targeted metabolomic fingerprinting' which could lead to misunderstanding. Indeed, in reviewing the literature over the past 40 years, it is evident that these various disciplines of

metabolite analysis are related via an evolution of methods and technology.

The amount of data provided by the different metabolomics approaches, mainly from untargeted and fingerprinting strategies, is of great complexity, and correct treatment of those data is of the utmost importance [11,12]. There are also classifications of metabolomics based on the specific objective of the analysis and data manipulation. For instance, Cevallos *et al.* [13] classified metabolomics as targeted and untargeted but also propose an interesting sorting as discriminative, informative, and/or predictive. Discriminative analyses aim to find differences between sample populations without necessarily creating statistical models or evaluating possible pathways that may elucidate such differences. Discrimination is usually achieved by the use of multivariate data analysis. In contrast, informative metabolomic analyses focus on the identification and quantification of targeted or untargeted metabolites to obtain sample intrinsic information. Informative metabolomics is used in the development and continuous update of metabolite databases. Finally, some metabolomics reports have been predictive. In this case, statistical models based on metabolite profile and abundance are created to predict a variable that is difficult to quantify by other means.

In general, in the strategy of fingerprinting, a chemical picture is made by direct analysis of crude sample extracts, typically by pyrolysis mass spectrometry (Py-MS) [14], direct infusion mass spectrometry (DIMS) [15], ambient mass spectrometry (Ambient MS), also designated as

direct ionisation mass spectrometry [16], or nuclear magnetic resonance spectrometry (NMR) [17]. Both techniques, MS and NMR, either generates a metabolite profile, in which the signals are assigned to specific metabolites, or a metabolite fingerprint, in which the analysis is based on the distribution of intensities in the corresponding spectra rather than the assignment of the signals. However, the metabolites detected by profiling must be recognized consistently and should be also quantified. For this, profiling requires, in general, a separation of the compounds and it is typically done by gas or liquid chromatography (GC, LC) or by capillary electrophoresis (CE) prior to the spectrometric detection by, e.g., UV, Fluorescence, NMR, or MS [2,18,19]. Therefore, this coupling tremendously expands the capability of the chemical analysis of highly complex biological samples.

Planning an efficient strategy for metabolome analyses requires consideration of what kind of information is needed, what kind of chemistry is expected the concentration range and what are the analytical facilities available. The selection of different analytical approaches (including the analytical techniques) will depend on the topic of the study and choosing a suitable analytical strategy requires a clear formulation of the problem to which we want some answers.

This tutorial aims at presenting the use of the fingerprinting methodology, applied on chromatography as a technique able of yielding proper fingerprints, independently of the coupled detector, and as suitable technique for food authentication. This tutorial involves several topics, which are: (i) the discussion about the meanings, in the

metabolomics and food authentication fields, of the related terms 'profiling', 'fingerprinting', and both targeted and untargeted analysis; (ii) a proposal of new terms related to food authentication as 'food identification'; and (iii) a critical overview about the applications of chromatographic (GC and LC) fingerprints analytical methods in food authentication using different chemometric techniques. These applications have been tabled according to several items as chromatographic method, chemical fraction, analytical aim, and chemometric target: similarity analysis, exploratory and clustering analysis, signal resolution, classification and regression on food properties.

2. Food authentication: targets and approaches. Current perspectives

Food authentication involves the confirmation of the stated specifications as true. It could include many aspects as the identification and/or quantification of characteristic components, adulterants and/or contaminants, and the verification of the differentiated quality requirements, among them the botanical or geographical origin and the manufacturing or processing procedure [20,21,22]. The authenticity is bound to the truthfulness and, therefore, food is considered authentic (or genuine) when it is not affected by any fraud [23].

Food authentication and food fraud detection are two sides of the same coin. They are similar problems and their solutions require parallel analytical strategies. In principle, three analytical approaches could be distinguished depending on the level in the structure of the matter for obtaining the target information: in tier 1, the information is obtained from the food chemical composition (chemical components approach, or just chemical approach); tier 2 works at biomolecular level, fundamentally on DNA or proteins (biomolecular approach); and tier 3 is focused on the stable isotopic composition of certain atoms (isotopic approach).

The chemical approach (tier 1) could be on three separate directions which lead to three different methodologies: (a) contents of chemical species of diverse nature, obtained by different techniques or analytical methods. A peculiar case of this methodology is constituted by the use of chemical markers, defined as particular components which are representative of a quality feature related to the authenticity; (b) component profile defined as a set of values which describe the contents of related constituents which share some physical (for example, volatile compounds, soluble fraction in a certain solvent) or chemical properties (for example, polar compounds), or even those compounds which belong to one chemical family (for example, phenolic compound, fatty acids, phytosterols, etc.). The component profiles are generally obtained by application of a single analytical method. If the constituents are expressed as a relative proportion, then it is named compositional profile; and (c) instrumental fingerprints that can be

defined as *a characteristic instrumental outline reflecting the complex chemical composition of the analysed sample and which can be related to its authenticity*. A similar scheme could be found in reference [23].

In addition, these methodological approaches could be applied with two goals: to demonstrate the fraud absence or to verify the own authenticity. Traditionally, the value of one or more analytical features of non-authenticity, characteristic of the fraud, are examined and if the found value is significant, food is stated as adulterated whereas, if it is negative, food is considered authentic; this is the "negative attitude" for food authentication. The (a) and (b) methodological approaches are mainly focused on this way and they require the application of high sensible and selective analytical techniques to detect, identify and quantify (if it is the case) the involved compounds or chemical species. Therefore, it is essential to have appropriate standards of each compound. The authentication is carried out by applying rules or previously developed decision trees which are based on the knowledge of the genuine food composition.

However, currently there is a growing importance in the application of a "positive attitude" directly concerned about showing the authenticity of food in study. The fingerprint methodology is rightly a positive-oriented approach because it already considers the acquired analytical signal in a global way, i.e., a spectrum or a chromatogram. This record, which could be obtained directly from the entire foodstuffs or from one fraction, constitutes an instrumental fingerprint of the food since it is characteristic and it, implicitly, reflects its chemical composition.

Fingerprinting analysis is formally incorporated to the proper vocabulary of food authentication in the EU Project "Tracing Food Commodities in Europe" - TRACE (FP6-2003-FOOD-2-A), in which food analysis by 'fingerprinting' is described as part of Work Package 2: "Fingerprint and profiling methods". However, this approach has been previously mentioned, even with a similar designation, as a non-selective way for food authentication [24]. In general, in the strategy of fingerprinting, a chemical picture is made by direct analysis of crude sample extracts, typically by (optical o mass) spectrometries [25,26], although sensor technologies as both electronic nose/tongue have to be also considered [27]. The application of this methodology does not require selective analytical methods since the test material characterization is carried out in a non-selective or untargeted way, but a high sensibility is advisable [28,29]. However, the practical application of the fingerprint methodological approach requires to have an adequate number of samples of previously stated as authentic foodstuffs in order to establish a representative data base of the genuine food population.

The Figure 1 shows a scheme of the described approaches.

The methodologies, based on profiles and fingerprints, require the use of chemometrics tools [30,31,32,33], which are essential in finding behaviours, trends or significant information from the data. In this way, the development of pattern recognition models, as exploration, classification / discrimination and prediction, allow to make visible the significant and non-evident information. There are a wide range of

chemometric methods to be applied, and all they could be categorised in two groups: (i) mathematical-statistical methods of the classical chemometrics (latent variables-based linear methods), and (ii) data mining methods (automatic or machine learning-based non-linear methods) which result a combination of computational intelligence, statistical methods and data bases systems.

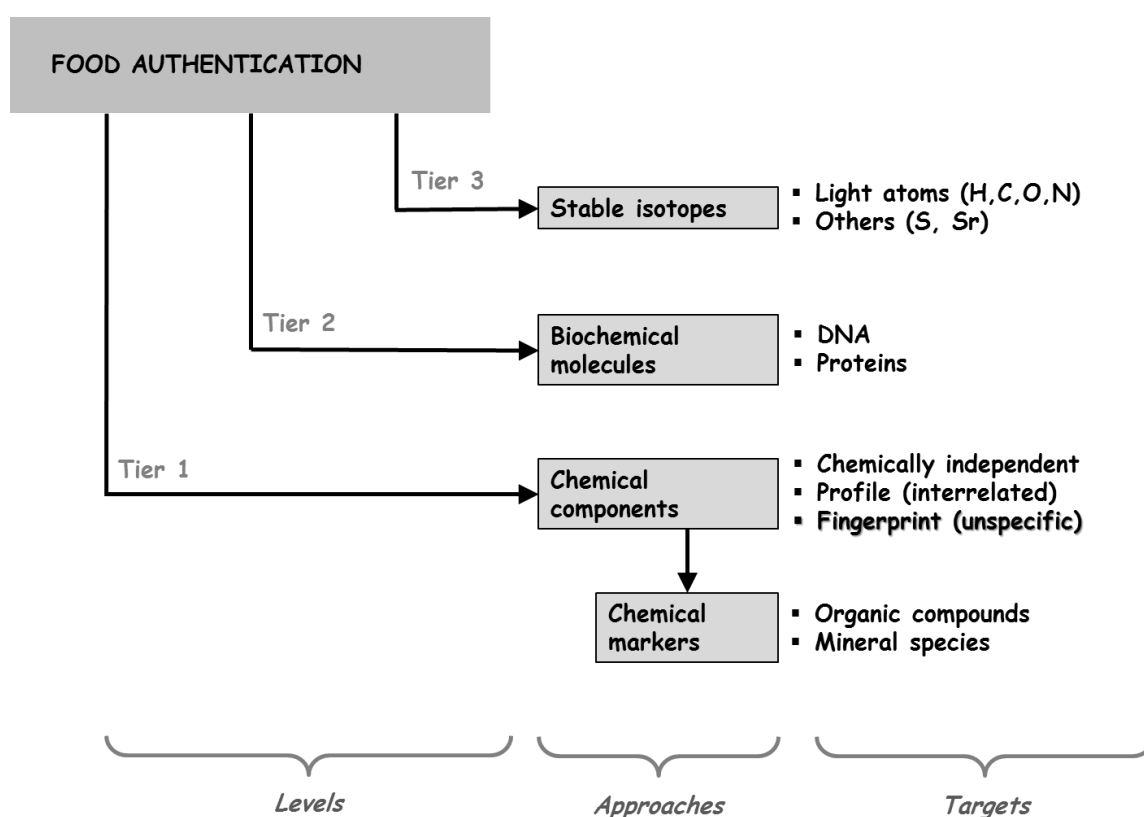


Figure 1. Approaches for food authentication.

Among the first ones it could be quoted the non-supervised exploratory methods, as principal components analysis (PCA), factorial analysis (FA) and cluster analysis (CA); the classification methods, based on both principal component models (SIMCA) and discriminant models from

latent variables (PLS-DA); and regression methods (PLS), in which food quality features are modelled to predict such values in unknown samples. Representative methods from the second group are artificial neuronal networks (ANN), support vector machine (SVM), decision trees, (M5, CART) as well as the ensemble methods that use groups of trees forming a "forest", as 'bagging' or 'random forest' (RF).

A further step in the use of Chemometrics to improve the quality and reliability of the conclusions from experimental data (profile or fingerprints) is the application of data fusion methods which involve the generation of models from the combination (fusion) of outputs of different instrumental techniques or analytical methods. The main goal of the data fusion, as it has been stated in a recent review where their application in food and beverage authentication and quality assessment are described [34], is to enhance the synergy between the fused data set by using complementary inputs, so to finally obtain better performance in the food authentication process.

And finally, an upper level is focused on the use of analytical signals with a multidimensional structure, equal or higher than 3 (3D signals or larger) which come generally from molecular fluorescence spectrometry or from coupled instruments, for example chromatography-mass spectrometry). The data handling requires the use of multiway (or N-way) chemometrics methods, as parallel factor analysis (PARAFAC) or Tucker3. Recently, the applications of such methods in food science have been reviewed [35,36].

2.1. Food metabolomics, foodomics and food authentication

The comprehensive characterization of food composition and the study of the modifications induced by the technological processes are basic issues of food science. Foods are, in general, mixtures of very different chemical compounds and they reflect the complexity of the animal or vegetable organism from which they derive. Just like any other biological system, in the 'omic' sciences, applications to foods are also found [37]. The food metabolome comprises a great variety of components that are very different from a chemical point of view, indeed belonging to a variety of chemical classes (sugars, organic acids, amines, amino acids, peptides, nucleic acids, and many other health-related compounds such as steroids or phenolic compounds, among others). Such molecules are known to play important roles in safety and quality of foods [38].

The aims and strategies of the metabolomics applied on foods have resulted in the proposal of a new eclectic 'omics' discipline, formally baptised as foodomics, and an innovative approach on food authentication. Figure 2 shows the transitions and the targets of these three items.

Foodomics was for the first time defined in 2009 as *a new discipline that studies the food and nutrition domains through the application of advanced omics technologies to improve consumer's well-being, health, and confidence* [39].

In the perspective of the global knowledge of foods, the concept of foodomics, more than a simple term covering the panel of the 'omic' sciences, also includes the nutritional aspects [40].

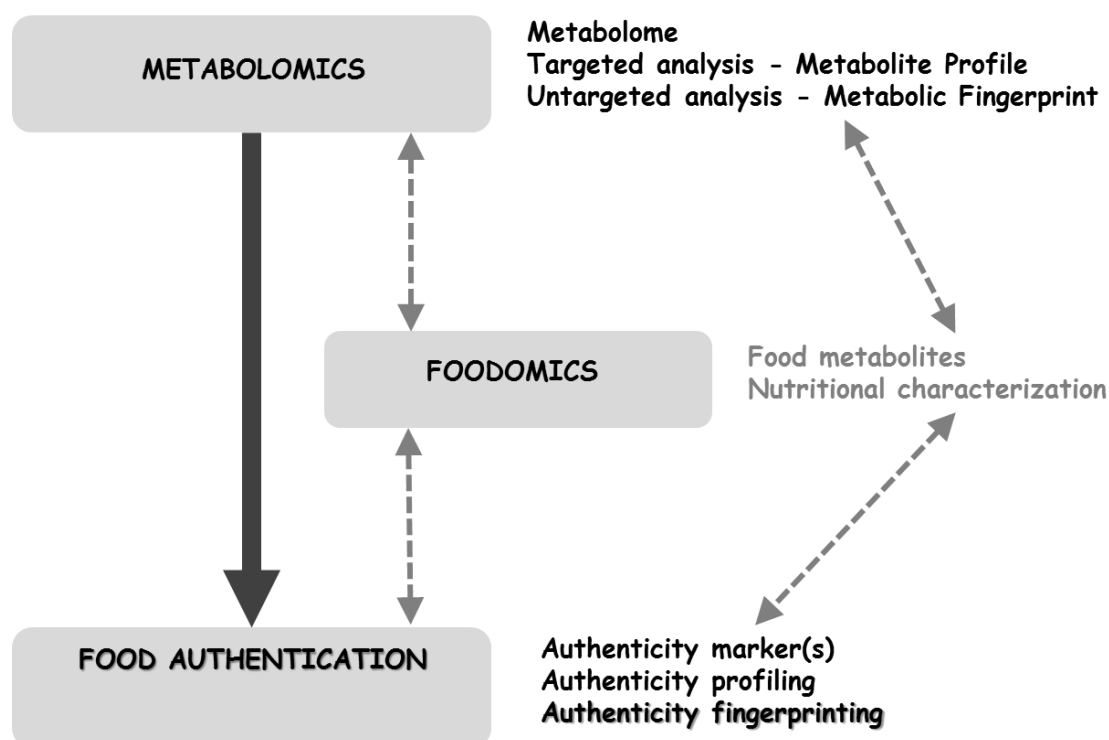


Figure 2. Food metabolomics, foodomics and food authentication.

Because metabolomics allows the simultaneous characterization of large numbers of chemicals in biological matrices, it actually offers food and nutrition scientists a superb opportunity to acquire a far more detailed and comprehensive molecular picture of food composition. In this line, Wishart [41] has reviewed the trends and potential applications of metabolomics in four areas of food science and technology: (1) food component analysis; (2) food quality/authenticity assessment; (3) food

consumption monitoring; and (4) physiological monitoring in food intervention or diet challenge studies.

The main advantage of metabolomics applications in food authentication makes use of its untargeted nature, which can enable the detection of emerging frauds. Cubero *et al.* [42] have developed a recent and very complete review on metabolomics for food authentication which describes the main steps, technical platforms and chemometrics tools used as well as the application, advantages, challenges and future trends of metabolomics fingerprinting approaches for the authentication of different food commodities [43]. Furthermore, applications devoted to food quality, safety and traceability, involving the use of MS-based metabolomics approaches from 2010 up to 2013, are described by Castro and Herrero [44]. On the other hand, an interesting review where applications about metabolomics to evaluate quality and safety of food production mainly plant-derived food related has been written by Oms Oliu *et al.* [45]. The authors stated that, significant improvements in terms of instrumentation are needed to make metabolite detection and quantification technology more robust, automated and comprehensive. In addition, several factors related to the optimization and validation of the analytical techniques, need to be assessed. Particular attention must be paid to the reproducibility and repeatability of the techniques and also to the preparation and sampling of biological material to minimize variability and in order to get more reliable results [45].

Moreover, apart from the fact of the considerable progress in this field, some important aspects need to be improved. For instance, there are thousands of applications in bibliography, however the number of samples used in them remain small. This implies that the results are not representative at all, the conclusion to be extended need to be supported by well-designed experimental studies which cover all the variability in the studied group of metabolites. Finally, it is necessary the development of appropriate standards analytical method in conjunction with the libraries of mass spectra of GC-MS and LC-MS for standard analytical methods.

2.2. Food authentication and food 'identitation' by fingerprinting

In a similar way to the use of the human fingerprint in people recognition, the food authentication by instrumental fingerprinting is based on verifying that the signal is consistent with a reference signal, within the sample and instrumental variability. Therefore, the practical application is not immediate but it necessarily requires the attainment of a previous stage that involves the establishment of an instrumental fingerprint which is characteristic of the authenticity.

To designate this previous process, the term 'food identitation' is proposed in this tutorial. The use of the word 'identification' with this aim is not right since this term implies an action which is different to the meaning of 'identitation'. Indeed, the Collins (British) English Dictionary defines identification as *the act of considering as the same*

or equivalent, or also as the act of proving or recognising as being a certain person or thing. So, identification involves a comparison, in a similar way to the authentication. On the contrary, the term 'food identitation' is related to the previous fact of defining the identity of a particular food based on the characteristic features which make it singular or unique. It could also be used, with the same meaning the expression food 'identity singularisation'.

This means that food 'identitation' and food authentication involve two independent and consecutive actions, in such a way that the second one it is not possible without having the first one finished. Also, reliability and confidence on food authentication depend on the 'identitation' goodness. Surprising, besides its importance, this step is generally obviated in the scientific publications on food authentication. The 'identitation' is applicable either to a specific food / group of foods or to specific food commodity with the same goals already described in section 2 for food authentication.

3. Chromatographic fingerprinting for food authentication: framework

The essential aim of the chromatography is the separation of chemical species and it is not the achieving of instrumental signals related to the chemical composition of the material test in study. This last key role is played by the detector which is the true analytical measurement instrument and is defined as *a device that measures the change in the*

composition of the eluent by measuring physical or chemical properties [46]. So, only the combination of both chromatographic and detection processes results in the chromatographic signals.

3.1. Chromatographic signals and analytical data

Before continuing, let us clarify some issues which will be used later on. The analytical signal refers to the quantitative record which is acquired by mean of an analytical instrument. A physical or chemistry property that, in the case of the chromatographic detectors, it is always related to a change in the composition of the eluent is measured continuously. Each analytical signal can be mathematically represented as a data tensor of different dimensionality, depending on the number of variables defining each one of the elements (numbers) which constitute the tensor [47,48].

For example, a conventional chromatogram is a two-dimensional record that can be described by a data vector where each element represents the signal intensity and it is characterized by a single retention time value; therefore a chromatogram is a first-order signal, this is a first-order tensor (a vector). An (UV-Vis or mass) spectrum-chromatogram, obtained respectively by a diode array detector or a mass spectrometer constitutes a three-dimensional signal resulting in a second-order tensor (a matrix) since each element is the intensity for each couple of values: a retention time and a wavelength of absorption or a mass/charge ratio, respectively. Finally, if a total fluorescence detector

is used, the acquired spectrum-chromatogram constitutes a four-dimensional signal resulting in a third-order tensor (a cube), since each element (fluorescence intensity) is defined by a retention time and two wavelengths corresponding to the excitation and emission process.

In a similar way, the 2D chromatography developed in equipment that include two orthogonal chromatographic stages, usually named comprehensive 2D chromatography, ($GC \times GC$, $LC \times LC$, or other combinations) lead to three-dimensional chromatograms which are a second-order signal. These ones are graphically displayed by contour maps or colour maps, since each intensity value is defined by two retention times corresponding to each chromatographic stages. Obviously, if the second stage is also coupled to a mass spectrometer or to diode array absorption spectrophotometer, a third-order signal could be acquired.

The analytical data are the set of figures obtained from the analytical signals. They are also represented as data tensors which are processed to obtain the corresponding results (data handling or data treatment). This conversion from signal to data could be carried out by several strategies that results in data tensor of different characteristics: (A) a data tensor with the same order and number of elements than the original signal; (B) a data tensor with the same order but lesser elements than the original signal (data reduction); (C) a data tensor of inferior order but with the same number of elements than the original signal; and (D) a data tensor of inferior order than the original signal

and with a reduced number of data (it is like merging the strategies (B) and (C)).

The data reduction resulting from strategies (B) and (D) could be performed in several ways and, among them, the most easy to apply is the signal sampling, for example, carrying out a proper decimation process. However, there are three specific methodologies which will be now considered: (B1) selection of a chromatogram zone delimited by two certain values of retention time; (B2) compilation of a single representative value from each chromatographic peak; and (B3) selection, by applying an appropriate algorithm, of a reduced number of elements from the original data tensor which have the relevant information. The strategy (A) does not suppose any transformation of the raw signal while the remaining strategies involve a handling on such signal.

As example, a straightforward case such as a conventional chromatographic system GC/FID which yields a chromatogram which is graphed as a 2D signal, will be considered. The application of the strategy (A) generates a data vector (first-order tensor 1) with many elements as measurements have been acquired by the detector, usually several thousands. The strategy (B1) extracts from the previous vector the elements that define one (or more) chromatographic peak previously selected. The strategy (B2), possibly the most applied, happens when the data vector is the result of joining the values of the areas (or heights) of the chromatographic peaks, or the values of the analyte calculated after a quantitation stage. On the other hand, the strategy

(B3) is a vector of few elements gathering the significant information from the original chromatogram. And finally, the strategies (C) and (D) lead, in this example, to the same result which happens when an only chromatographic peak is considered and an only value (peak area or analyte content); this is then a zeroth-order tensor (a scalar). The Figure 3 summarizes the described strategies.

The 3D signals, for instance chromatography-mass spectrometry (GC-MS or LC-MS) signals or 2D chromatograms acquired by a conventional detector in a comprehensive $GC \times GC$ chromatographic system, offer a diversity of options to apply the different strategies, between these, those devoted to reduce the dimensionality of the tensor (matricizing). There are different terms in the literature used for the matricizing: unfolding, juxtaposition, concatenation and stringing-out. The term unfolding it is the most employed and is performed by concatenating two-way array (matrices) extracted from a three-way data array [35].

3.2. Chromatographic profiles and chromatographic fingerprints

In the Introduction section, the analytical use of the terms profile and fingerprint has been described in the metabolomics context. In this section, both terms are redefined and the corresponding meanings are justified regarding to chromatographic analysis.

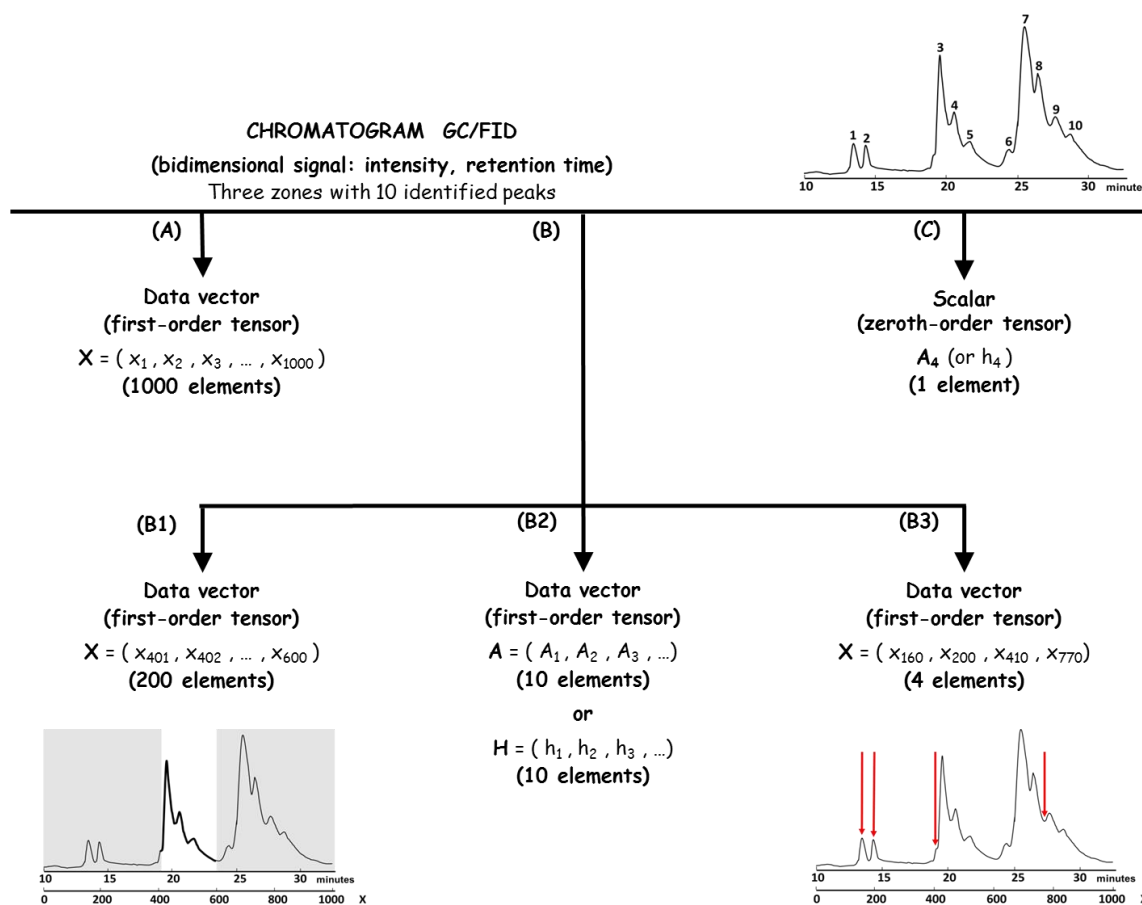


Figure 3. Types of analytical data (tensors) resulting from a conventional GC/FID chromatogram (Chromatograms are adapted for a better understanding from references 75 and 87).

It is possible to find at least two meanings for the term profile that could be applied in the field of the chromatography. The Collins (British) English Dictionary depicts the following two definitions: (P1) *a view or representation of an object, especially a building, in contour or outline*; (P2) *a graph, table, or list of scores representing the extent to which a person, field, or object exhibits various tested characteristics or tendencies*. The meaning (P1) is related to the shape of an object, and the meaning (P2) is focused in the distinctive features. Therefore, the term chromatographic profile may be understood as (P1) the

silhouette of the chromatogram and (P2) the set of characteristic features defining the chromatographic peaks. (P1) relates to the signal concept while (P2) to the data. This double meaning of the word profile leads to confusion, since diverse authors could use the term with one or another meaning and, therefore, it is always necessary to analyse the context in that it is being used to find the right meaning. In this tutorial, the second meaning (P2) will be always used and so chromatographic profile will refer to the collection of chromatographic peak areas (or heights) which characterise the chromatogram, or its respective analyte-derived values.

The meaning of the term fingerprint in analytical chemistry is an extending use from the forensic one in people identification. In this context, a fingerprint constitutes a record characteristic of the identity of certain person. It joints together two determining properties: (i) it does not describe the person's anthropomorphic attributes, as height or weight, eyes or hair colour, etc., and they cannot be expected from the fingerprint analysis; and (ii), it cannot be used for identification by itself since this is based on a comparison with the fingerprint from the same person, previously recorded and kept in a database. In the same way, the term chromatographic fingerprint should always be referred to the entire chromatogram from a certain test material which is distinctive of its composition. However, it does not infer explicit information about composition or other physical or chemical properties. The most immediate application of such chromatogram is to establish a specific and differentiating tool, as an

identity card, which could be used in order to 'identitate' or identify a certain material. With this aim, the chromatographic fingerprints are currently applied in differentiating botanical material [49,50,51], particularly traditional Chinese herbal medicines [52,53,54]. In this context, an herbal medicine fingerprint has been properly defined as *a chromatographic pattern of the extract of some common chemical components of pharmacologically active and or chemically characteristics* [55].

In some occasions, for example in phytochemical analysis, both terms, profile and fingerprint, are used conjointly in a right way, as 'fingerprint profile' [53], to refer to a chromatogram where the term profile is used now with the meaning (P1). Nevertheless, and in order to avoid confusions, both terms should always be used regarding the chromatographic signal-derived analytical data (see Figure 4).

In conclusion, the most important conceptual difference between both terms, chromatographic profile and chromatographic fingerprint, comes from the type of analytical information that it is involved. A chromatographic profile implies specific information on particular chemical components while a chromatographic fingerprint contains unspecific and non-evident information which should be extracted by chemometric tools.

3.3. *Chromatographic fingerprinting methodology*

The most common modalities of chromatography are gas chromatography (GC) and liquid chromatography (LC). The tandem GC and FID have been available for decades providing an analytical technology among the highest robust, although it presents the drawbacks of needing to derivatise before analysis. In contrast, many classes of food solutions, into both polar and non-polar solvents, can be analysed directly and easy by LC without derivatisation. Currently, the LC systems are shifting from standard HPLC to UHPLC, which can significantly increase resolution sensitivity and peak capacity due to the reduced particle size, while significantly decreasing sample volumes and mobile phases.

The data acquired from a chromatographic method include values for retention time, peak intensity (area or height) and, if MS is coupled, peak m/z . In GC-MS-based approaches, the spectra comprise several m/z fragments from a single analyte whereas in LC-MS there is less fragmentation and isotopic peaks can be detected. In addition, the use of comprehensive two-dimensional chromatography opens the possibility of having an additional chromatographic dimension, as there are two retention times instead of one, increasing the potential of profiling and identifies markers making the data handling even more challenging [56].

The versatility of the chromatographic methods allows the analyst to interact on both the separation and measurement steps in order to acquire an analytical signal with the maximum useful information, as well

as to select the data to be treated. As consequence, these methods are excellent candidates for getting proper fingerprints for food 'identitation' as well as for food authentication. The framework about the application of chromatography in this subject was already suggested in a previous paper published in this same journal [57]. An updated diagram is shown in the Figure 4.

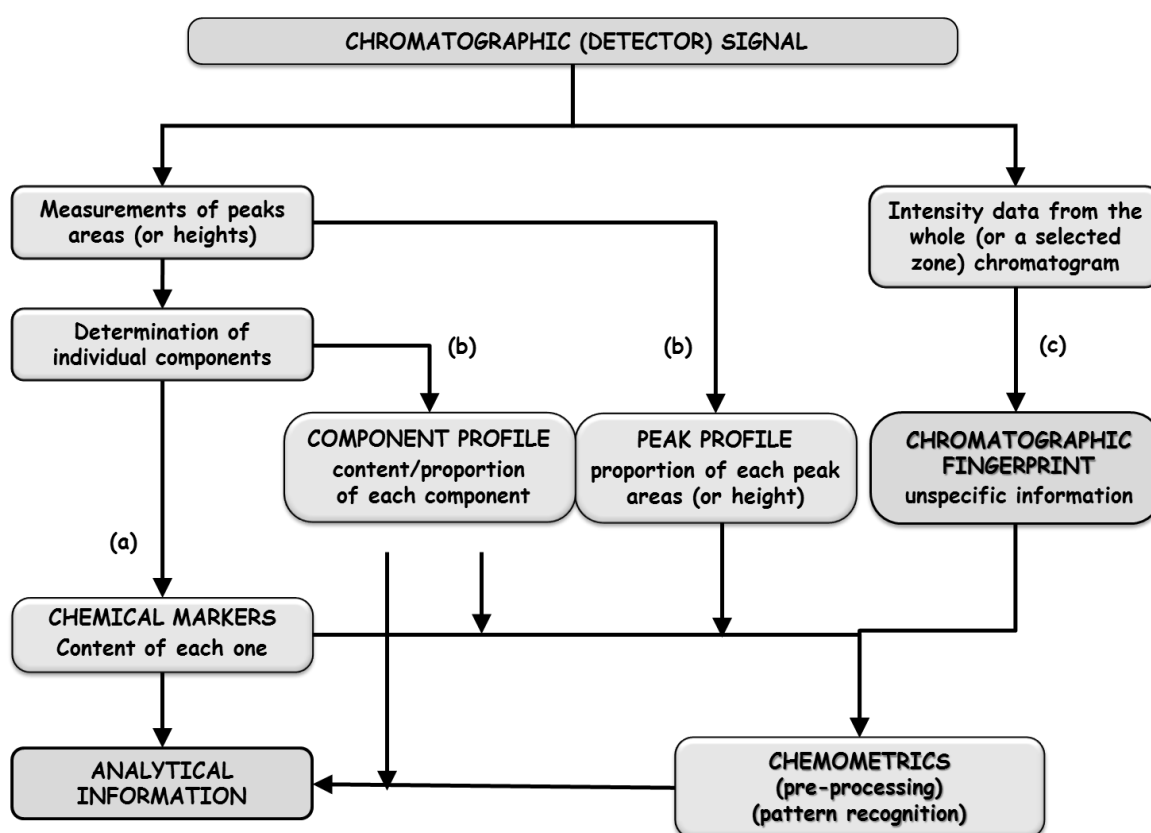


Figure 4. Different alternatives of the use of chromatographic signals for food authentication.

As it could be deduced from the diagram, the chromatographic methods are suitable in applying any methodological approach on food authentication which have been previously described in section 2.

However, and this is also the problem, the aim of the chromatographic method is different depending on the methodology that will apply. Indeed, the profiling methodology requires traditional chromatography, so, the chromatographic method has to be optimized reaching the maximum resolution in the separation, so the signals are the most specific possible from the test material composition. In other words, ideally each peak should inform about only one component. When this is not possible, coupled high resolution mass spectrometers allow to get specific signals of different components which have co-eluted in the same chromatographic peak.

However, the use of chromatographic fingerprint implies a methodological change on the major paradigm of chromatography. Indeed, the speed of the analysis, besides the achieved separation, is now the main priority in order to compete with the more recognised spectrometric methods. Ideally, if the only aim is to obtain a proper chromatographic fingerprint, the run time should not be longer than to 5-10 minutes. However, an analysis time around 15-20 minutes would give the opportunity to obtain compositional information in order to carry out a more proper "identification".

But, obviously, the acquired chromatogram in this time interval has to comprehend the enough characteristic information which allows the 'identification' or the authentication of the food under study. A feasible solution is optimizing the method to obtain a chromatographic hump with an irregular shape with jagged edges resulting of the overlap of the potential and unresolved chromatographic peaks. Although in

principle, the universal detectors seem to be the most suitable for this end, a more specific detector could provide a valid and suitable signal for complex test materials. The coupling of mass spectrometry is indifferent in this methodology unless the additional signal dimension, defined on the axis m/z , is used. Therefore, the same chromatographic method might be fitted for profiling and fingerprinting.

The application of fingerprinting methodology from a ideally well-resolved chromatogram, when no hidden component is co-eluting, is equivalent to use the corresponding chromatographic profile. Indeed, the only elements of the data tensor from a well-resolved chromatogram which have useful information are the elements corresponding to the peaks, since the rest are on the baseline. And all the elements of the data tensor belonging to the same chromatographic peak are strongly correlated and, in the practice, they could be substituted by an only element as the maximum intensity (or peak height) or the peak area because they gathers the global information an only element. In this way, the built chemometric model from the reduced tensor (profile) is equivalent to the derivate of whole tensor collecting all the points of the chromatogram, and the overfitting risk is also avoided. However, it is possible that a chromatogram seems "well-resolved" but it could have some overlapped peaks with other components that render the information retrieved by the chemometrics analysis. In this case, the fingerprinting methodology could also be an advantage over the analysis of only the maximum intensity values for each "peak".

4. Applications of chromatographic fingerprinting on food 'identification' and food authentication

In this section, it is shown an overview of the published articles about right applications of the fingerprinting methodology for both food 'identification' and food authentication. Therefore, the articles which state the use of chromatographic fingerprints but effectively they are applying profiling methodology, are not included.

The number of published applications is even limited; only 32 and 26 articles have been found using *GC* and *LC*, respectively. Specific chemometrics tools are used with three main objectives: (i) verification of the similarity between chromatographic signals for food identification; (ii) resolution of overlapped chromatographic signals; (iii) classification of foods regarding diverse criteria; and (iv) prediction of specific properties related to food quality. In certain cases, a previous data examination is also applied by exploratory and/or cluster analysis.

There are several key aspects which are worth mentioning:

- The most cases, the run times are excessive for a competitive application of the fingerprinting methodology. There is only three articles with the run time lesser than 10 minutes, while a lot of cases show run times upper than 1 h. This fact possibly happens because the methods have been optimized searching the best resolution, as it would be carried out in profiling methodology.

- Except for two articles, one for each kind of chromatography in which chemometrics tool of data mining are applied properly (SVM, RF, M5), the remainder use classical chemometric methods.
- There is not any example using 3D fingerprints, obtained by comprehensive 2D chromatography where 3-way chemometric methods would be required for the data handling.
- There is not either any antecedent using data fusion methods [34] applied to more than one chromatographic fingerprint.

These aspects indicate that the fingerprinting methodology is not still settled in the analytical community, and many researchers are reluctant to develop *ad-hoc* methods and applications. This fact demonstrates that there is a lot of work to do on this topic.

4.1. Gas chromatography

Gas chromatography has been applied to a wide range of food for 'identification', authentication or quality features prediction. Table 1 summarises the published applications.

Different conventional methods of multivariate data treatment have been used to extract useful information for similarity analysis and pattern recognition. The most widespread ones are PCA, CA, PLS, and PLS-DA. The chromatographic problems, such as baseline/background contribution or retention time shifts were corrected, by pre-processing data techniques, before building the models.

In addition, other tools such as MCR or SFA were also achieved in order to resolve co-elutions (overlapped and embedded peaks), and to get a second better resolved fingerprint. MCR method was proposed to improve the pattern of GC-MS fingerprints for a subsequent profiling of secondary metabolites in citrus fruits peels [81]. The qualitative and quantitative profiling of coffee flavour compounds was developed with the help of SFA, which was used as an auxiliary tool to resolve the overlapped peaks of the different volatile constituents [70].

The application of multi-way models to GC fingerprint for authentication of food is not very extended. These methods help to improve the visualization and interpretation of complex multidimensional data defined by more than 2 variables. Durante *et al.* [67] propose the application of a multi-way regression method on the whole chromatogram for monitoring both the sensory and compositional profiles during the ageing process of Aceto Balsamico Tradizionale di Modena. In addition, Salvatore *et al.* [35] propose the assessing of the geographical origin of table wines and extra virgin olive oils by three-way data collected from GC-MS; different multi-way classification models were built.

Table 1. Applications related to the use of gas chromatographic fingerprinting in food 'identification' and food authentication (in chronological order).

Food	Chromatography, run time	Chemical fraction	Analytical aim	Chemometrics	Ref.
Balsamic vinegars	GC-MS, 53 min	Volatiles	Classification of balsamic vinegars	Exploratory analysis: PCA Classification: SIMCA	[66]
Balsamic vinegars	GC-FID, 63 min	Volatiles	Prediction of sensory attributes	Regression: NPLS, unfold-PLS	[67]
Green tea	GC-(TOF)MS, 25 min	Hydrophilic primary metabolites	Classification of Japanese green tea by qualities	Exploratory analysis: PCA Regression: PLS	[68]
Rice brand wax	GC-MS, 41 min	Policosanols	Prediction of product quality	Similarity analysis: Euclidean distance	[69]
Coffee	GC-MS, 45 min	Volatiles	Identification of coffee flavor	Curve resolution: SFA Similarity analysis: correlation coefficient	[70]
Balsamic vinegars	GC-FID, 59 min	Volatiles	Evolution of the volatile compounds during ageing	Exploratory analysis: PCA	[71]
Green tea	GC-MS, 47 min	Pyrolyzed metabolites	Prediction of product quality	Exploratory analysis: unfolding PCA Regression: PLS	[72]

Table 1. (continued).

Food	Chromatography, run time	Chemical fraction	Analytical aim	Chemometrics	Ref.
Green tea	Fast GC-FID, 14 min GC-(TOF)MS, 28 min	Organics extracts (methanol/water/chloroform)	Prediction of product quality	Regression: PLS	[73]
Apple	GC-(TOF)MS, 46 min	Volatiles	Classification of apple varieties	Classification: PLS-DA Exploratory analysis: PCA	[74]
Extra virgin olive oil	GC-(IT)MS, 63 min	Volatiles	Authentication of geographical origin	Exploratory analysis: PCA Classification: LDA, SLDA	[75]
Beer	GC-FID, 18 min	Volatiles	Classification by types, breweries and tastes	Exploratory analysis: PCA Clustering: HCA	[76]
Olive oil	GC-(IT)MS, 35 min	Triacylglycerol	Classification of olive oil varieties	Exploratory analysis: PCA Clustering: HCA	[77]
Green tea	GC-MS, 70 min	Volatiles	Classification of Japanese (Sen-cha) green tea by qualities	Classification: PLS-DA Regression: PLS	[78]
Apple juice	GC-MS, 40 min	Volatiles	Discrimination of juices by apple varieties and geographical origin	Exploratory analysis: PCA Classification: LDA, SLDA	[79]

Table 1. (continued).

Food	Chromatography, run time	Chemical fraction	Analytical aim	Chemometrics	Ref.
Hard and semi-hard natural cheese	GC-FID, 18 min	Hydrophilic low molecular weight components	Modelling of sensory prediction	Regression: PLS	[80]
Citrus fruits peels	GC-MS, 91 min	Secondary metabolites	Classification by fruit types	Resolution: MCR	[81]
Coffee	GC-FID, 26 min	Volatiles	Modelling of sensory prediction	Regression: PLS	[82]
Edible vegetable oils	GC-(IT)MS, 17 min	Triacylglycerols	Quantification of olive oil in oil blends	Classification: SIMCA Regression: PLS	[83]
Edible vegetable oils	GC-MS, 45 min	Fatty acids	Classification of vegetable oils. Detection of adulteration of canola oil	Exploratory analysis: PCA Classification: PLS-DA, OPLS-DA	[84]
Port wine	GC-FID, 121 min	Volatiles	Classification of Port wines by aging	Regression: PLS Exploratory analysis: PCA	[85]
Mushrooms	GC-(IT)MS, 121 min	Volatiles	Discrimination mushrooms species	Regression: PLS Exploratory analysis: PCA	[86]
Palm oil	GC-MS, 17 min	Triacylglycerols	Authentication of geographical origin	Classification: PLS-DA	[87]

Table 1. (continued).

Food	Chromatography, run time	Chemical fraction	Analytical aim	Chemometrics	Ref.
Table wines	GC-MS, 45 min	Volatiles	Authentication of geographical origin	Multiway classification: unfold PLS-DA, unfold SIMCA, NPLS-DA, NSIMCA	[35]
Extra virgin olive oil	GC-MS, 67 min	Volatiles	Authentication of geographical origin	Multiway classification: unfold PLS-DA, unfold SIMCA, NPLS-DA, NSIMCA	[35]
Oxidized tallow	GC-MS, 26 min	Volatiles	Identification of oxidized tallow	Similarity analysis: correlation coefficient and Euclidean distance	[88]
White wine	GC-MS, 48 min	Volatiles	Authentication of grape varieties	Exploratory analysis: PCA	[89]
White apricot almond	GC-MS, 31 min	Fatty acids	Quality control and detection of frauds	Exploratory analysis: PCA	[90]
Coffee	GC-MS, 70 min	Volatiles	Detection and quantification of coffee adulteration	Classification: OPLS-DA	[91]
Cherry wines	GC-MS, 49 min	Volatiles	Discrimination of Sherry wines	Clustering: CA	[92]
				Regression: PLS	
				Exploratory analysis: PCA	

Table 1. (continued).

Food	Chromatography, run time	Chemical fraction	Analytical aim	Chemometrics	Ref.
Edible vegetable oils	GC-(IT)MS, 17 min	Triacylglycerols	Quantification of olive oil in oil blends	Classification: RF, M5 rules	[93]
Olive oil	GC-FID, 60 min	Volatiles	Characterization of sensory attributes	Exploratory analysis: PCA	[27]
Green tea	GC-MS, 63 min	Volatiles	Identification and quality control of Pu-Erh green tea	Similarity analysis: CASE Exploratory analysis: PCA Clustering: CA	[94]

Finally, the use of data mining methods (RF and M5-based rules) [93] in authenticating olive oil should be highlighted. The modelling techniques are not comparable to traditional chemometric ones. Their advantages, as it is stated by authors, include interpretable models with much more information than the traditional chemometric methods. In addition, those methods provide valuable information for detecting which vegetable oil is mixed with olive oil and the percentage of oil used, with a single chromatogram.

4.2. Liquid chromatography

Liquid chromatography has also been applied to 'identify', classify and assess the quality of several kinds of foods. Table 2 reviews the main applications.

HPLC fingerprints are used to evaluate the similarity of pomegranate, raspberry, soybean and tea samples. The similarity analysis is applied to calculate some correlation parameters (correlation (r) and congruence (c) coefficients), and distance measurements. Noteworthy is the comparison of dissimilarity of chromatographic fingerprints only by means of a visual way, the overlapped chromatograms [102].

The conventional applications, which are the most numerous, imply the use of classical chemometric tools as exploratory/clustering and classification/discrimination methods. Between all the studies devoted to this issue, a paper should be emphasised because of their current interest. Pérez *et al.* [120] use two sterolic chromatographic

fingerprints and two classification methods to discern the geographical origin of edible palm oil. However, the main interest of this paper is showing how the results obtained from the four classification scenarios (combination of classification chemometric methods and measured analytical dataset) are ranked. To this end, several quality performance features are introduced.

In addition, a paper deserves also a special mention here for the nature of the particular chemometric tools that applies. Indeed, Zhen *et al.* [101] use SVMs and RF machine learning methods for assessing comparatively the ability to classify unknown tea samples using models constructed from HPLC fingerprints. Finally, in four papers multivariate regression methods are used to verify the potential use of fast chromatograms in prediction of the antioxidant capacity of green tea from HPLC fingerprints. It happens that the four studies apply different regression methods on the same data set which was firstly describe by Nederkassel *et al.* [95].

4.3. Comprehensive 2D chromatography

About this topic, there are a few of recent reviews which are aimed to describe the developments and applications of comprehensive 2D chromatography on food analysis using a target approach [58,59], and focused on the chemometrics techniques to be applied [60,61].

Table 2. Applications related to the use of liquid chromatographic fingerprinting in food 'identification' and food authentication (in chronological order).

Food	Chromatography, run time	Chemical fraction	Analytical aim	Chemometrics	Ref.
Green tea	HPLC-UV, 2 and 11 min	Aqueous extracts	Prediction of the total antioxidant capacity	Regression: PLS	[95]
Red wine	HPLC-DAD, 55 min	Phenols	Classification of the Chilean wines varieties	Classification: QDA; LDA; KNN; PNN	[96]
Vanillin	HPLC-UV, 60 min	<i>No information is available</i>	Discrimination of vanillin from different origins	Exploratory analysis: PCA	[97]
Green tea	HPLC, 2 min	Aqueous extracts	Prediction of the total antioxidant capacity	Regression: PLS; PRM	[98]
Green tea	HPLC-UV, 2 and 11 min	Aqueous extracts	Prediction of the total antioxidant capacity	Regression: MLR; PCR; PLS; OPLS	[99]
Raspberry	HPLC-DAD, 30 min	Flavonoids	Identification of raspberry	Similarity analysis: SESCOF	[100]
Tea	LC-(Orbitrap)MS, 60 min	Aqueous extracts	Distinction between different commercial teas	Exploratory analysis: PCA Classification: SIMCA, SVM, RF	[101]
Green tea	HPLC-UV, 11 and 13 min	Aqueous extracts	Prediction of the total antioxidant capacity	Regression: MLR; OPLS	[102]
Trappist beer	UPLC-(Q-TOF)MS, 10 min	Entire beer without dilution	Classification of the beers from different brands	Exploratory analysis: PCA Classification: O2PLS-DA	[103]

Table 2. (continued).

Food	Chromatography, run time	Chemical fraction	Analytical aim	Chemometrics	Ref.
Edible vegetable oils	HPLC-CAD, 45 min	Triacylglycerols	Quantification of olive oil in blends with vegetable oils	Regression: PLS; iPLS	[104]
Edible vegetable oils	HPLC-CAD, 25 min	Triacylglycerols	Discrimination of olive oil from vegetable oils	Exploratory analysis: PCA	[105]
Green and white tea	UPLC-DAD/(ESI)MS, 14 min	Alkolooids, phenols, catechins, flavonoids, and proanthocyanidins	Differentiation between tea categories	Exploratory analysis: PCA	[106]
Kiwi and pomelo	HPLC-(ESI)MS, 30 min	Catechins, phenols and flavonoids	Classification according to the fruit species and subspecies	Exploratory analysis: PCA Clustering: CA Classification: LDA	[107]
Tofu-type soybean	HPLC-UV, 40 min HPLC-(ESI)MS, 40 min	Isoflavones	Differentiation of Tofu-type soybeans	Similarity analysis: CASE	[108]
Green tea	HPLC-UV, 11 and 13 min	Aqueous extracts	Evaluation of (dis)similarity	Similarity analysis: CASE; heat maps; Euclidean and Mahalanobis distances	[109]
Red wine	HPLC-UV, 20 min HPLC-FLD, 20 min	Phenols	Discrimination of Spanish wines from PDO, grape varieties, ageing periods and vintage	Exploratory analysis: PCA	[110]

Table 2. (continued).

Food	Chromatography, run time	Chemical fraction	Analytical aim	Chemometrics	Ref.
Extra virgin olive oil	HPLC-DAD, 30 min	Phenols	Discrimination of the PDO Sabina olive oil	Classification: PLS-DA Regression: biPLS-GA	[111]
Honeybush tea	HPLC-DAD, 40 min	Aqueous extracts	Differentiation between wild-harvested and cultivated seedling plants Evaluation of similarity	Exploratory analysis: PCA Similarity analysis: correlation coefficient	[112]
Sugarcane	HPLC-DAD, 60 min	Flavonoids	Discrimination of genotypes	Exploratory analysis: PCA	[113]
Honey	HPLC-DAD, 50 min	Flavonoids	Authentication and classification of floral origin	Similarity analysis: SESCOF Exploratory analysis: PCA Classification: PLS-DA, SIMCA	[114]
Edible vegetable oils	HPLC-UV, 12 min	Sterols (from unsaponifiable fraction)	Discrimination of the virgin olive oils from edible vegetable oils	Regression: PLS Clustering: HCA Exploratory analysis: PCA Classification: PLS-DA	[115]
Honey	HPLC-DAD, 40 min	Aqueous solution	Verification of the honey floral origin	Exploratory analysis: PCA Classification: KNN	[116]

Table 2. (continued).

Food	Chromatography, run time	Chemical fraction	Analytical aim	Chemometrics	Ref.
Pomegranate peel	HPLC-UV, 80 min	Phenols	Identitation of Chinese pomegranate from different orchards	Similarity analysis: SESCOF	[117]
Green tea	HPLC-DAD, 60 min	Phenols	Identitation of Chinese Ziyang green tea from different plantations	Similarity analysis: SESCOF	[118]
Red wine	LC-(ESI)(Q-TOF)MS, 45 min	Non-volatiles	Discrimination between grape varieties	Exploratory analysis: PCA	[119]
Palm oil	HPLC-UV, 15 min HPLC-CAD, 18 min	Methyl transesterified fraction	Differentiation of geographical origin	Exploratory analysis: PCA Classification and prediction: PLS-DA, SIMCA	[120]

However, as it has been stated before to the best of knowledge there is not any published study in which multidimensional chromatographic fingerprints are properly used in food authentication.

It is true that several papers report analytical methods for acquiring comprehensive 2D chromatograms (in fact, 3D signals), which would be rightly designated as 2D chromatographic fingerprint. However, these signals are habitually used to apply profiling following the B2 strategy (section 3.1). An example is the paper of Vaz Freire *et al.* [62] where PCA was performed using the peak areas obtained directly by image treatment from the contour plots, to discriminate three different olive oils varieties by the analysis of the aroma compounds. Another example of this confusing use is the recent proposal of the term 'blueprint' to denote the two-dimensional chemical profile of the volatile compounds characteristic of the virgin olive oil, by comprehensive *GC*×*GC* [63].

Accordingly, we would like to encourage the food analytical researchers to work on the challenge of applying N-way chemometrics techniques to comprehensive 2D chromatographic fingerprints [64] in order to develop analytical methods for both food 'identification' and authentication as it is already applying other fields [65] although keeping in mind that the requirements of section 3.3 should be fulfilled. This field is far from being exploited and it could be of great interest to the food science community. The results could be surprising.

5. Concluding remarks

The framework of food authentication, particularly by chromatographic methods, has been described and clarified. As a result, we have detected that, there is not a proper term that describes the step of obtaining the reference fingerprint from genuine foods. For this reason the term 'food identification' is proposed in order to define this stage and to close the life cycle of the chemical-based food authentication process.

Some reviews have been published in recent years devoted to fingerprinting methodology and food authentication. However, they are only focused on the vibrational spectroscopies (NIR, MIR, Raman), NMR spectroscopy, as well as a range of MS techniques but applications of chromatographic methods are usually not included. For this, an exhaustive bibliographical search focused on chromatographic fingerprinting methods in food authentication has been carried out, and the found results have been reviewed in order to discover the right analytical applications. It has to be highlighted that this revision has been complex for two reasons. One of them because in a lot of papers appear the terms 'fingerprint' or 'fingerprinting' in the title and/or the keywords but they are really applying profiling methods; in such cases, these papers have not been included. And the other one because, in some papers which the terms profiling or untargeted are stated in

the title and/or the keywords but they are related to fingerprinting methodology.

The authors trust that the use of fingerprinting chromatographic in food authentication (or "identitation") is perfectly suitable and valid for working in-house, when authentication purposes are required. However when transferability of the methods is needed, for acceptance by the authorities, producers and consumers, it should be compulsory the use of food reference material for instance, then the fingerprinting signal could be normalized between equipment. In addition, compositional information on significant chemical markers could be used as a second step for confirmation of the authenticity as well as for transferability. Finally, the chromatographic fingerprinting combined with chemometric techniques are a powerful tool for detection of food fraud. Therefore, the laboratories are encouraged to implant the use of this methodology.

Acknowledgment

Authors acknowledge to the programme of R&D strengthening from the University of Granada, which is supported by the Andalusia Regional Government (Consejería de Innovación, Ciencia y Empresa) and, partially, by European Regional Development Funds (ERDF).

Appendix A. List of acronyms

➔ *Chemometrics methods:*

Similarity analysis:

- CASE: Computer-aided similarity evaluation system
- SESCOF: Similarity evaluation system for chromatographic fingerprinting

Exploratory analysis:

- PCA: Principal component analysis
- unfolding PCA: PCA applied on the unfolded data set

Clustering:

- CA: Cluster analysis
- HCA: Hierarchical cluster analysis

Curve resolution:

- MCR: Multiple curve resolution
- SFA: Sub-window factor analysis

Classification:

- kNN: k-Nearest neighbour
- LDA: Lineal discriminant analysis
- M5: M5 trees regression rules
- M5P (or M5'): M5 prime trees regression rules
- MLR: Multiple linear regression
- NPLS-DA: N-way PLS-DA
- NSIMCA: N-way SIMCA

- OPLS-DA: Orthogonal PLS-DA
- O2PLS-DA: A modification of the original PLS-DA method
- PLS-DA: Partial least squares - discriminant analysis
- QDA: Quadratic discriminant analysis
- PNN: Probabilistic neural networks
- RF: Random forest
- SIMCA: Soft independent modelling class analogy
- SLDA: Stepwise LDA
- SVM: Support vector machine
- Unfold PLS-DA: PLS-DA applied on unfolded data set
- Unfold SIMCA: SIMCA applied on unfolded data set

Regression:

- biPLS-GA: Backward-interval PLS - genetic algorithms
- iPLS: Interval PLS
- NPLS: N-way PLS
- OPLS: Orthogonal PLS (or orthogonal projections to latent structures)
- PCR: Principal component regression
- PLS: Partial least squares regression
- PRM: Partial robust M-regression
- Unfold-PLS: PLS applied on unfolded data set

➔ *Chromatographic techniques:*

Gas chromatography:

- GC-FID: Gas chromatography with flame ionization detection
- GC-MS: Gas chromatography coupled to mass spectrometry
- GC-(IT)MS: Gas chromatography coupled to mass spectrometry with a trap ionic ioniser
- GC-(TOF)MS: Gas chromatography coupled to mass spectrometry with a time-of-flight analyser

Liquid chromatography:

- HPLC-CAD: High performance liquid chromatography coupled to a charged aerosol detector
- HPLC-DAD: High performance liquid chromatographic coupled to a diode array detector
- HPLC-(ESI)MS: High performance liquid chromatographic coupled to a electrospray ionization mass spectrometry
- HPLC-UV: High performance liquid chromatographic coupled to an ultraviolet detector
- UPLC-(Q-TOF)MS: Ultrahigh performance liquid chromatography coupled to high-resolution quadrupole-time-of-flight mass spectrometry

References

- [1] E.C. Horning, M.G. Horning, Human metabolic profiles obtained by GC and GC/MS, *J. Chromatogr. Sci.* 9 (1971) 129-140.
- [2] J-L. Wolfendera, G. Martia, A. Thomas, S. Bertrand, Current approaches and challenges for the metabolite profiling of complex natural extracts, *J. Chromatogr. A*, 1382 (2015) 136-164.
- [3] O. Fiehn, Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks, *Comp Funct Genom* 2 (2001) 155-168.
- [4] J. Gummer, M. Banazis, G. Maker, P. Solomon, R. Oliver, R. Trengove, Use of mass spectrometry for metabolite profiling and metabolomics, *Australian Biochemist* 40 (2009) 5-8.
- [5] E. Trujillo, C. Davis, J. Milner, Nutrigenomics, proteomics, metabolomics, and the practice of dietetics, *J. Am. Diet. Assoc.* 106 (2006) 403-413.
- [6] F. Moritz, S. Forcisi, M. Harir, B. Kanawati, M. Lucio, D. Tziotis, P. Schmitt-Kopplin, The Potential of Ultrahigh Resolution MS (FTICR-MS) in Metabolomics, in: M. Lämmerhofer, W. Weckwerth (Eds.), *Metabolomics in Practice*, Wiley-VCH, Weinheim (Germany), 2013, pp. 118-119.
- [7] J. Nielsen, Metabolomics in functional genomics and system biology, in: S. G. Villas-Boas et al. (Eds), *Metabolome Analysis. An Introduction*, John Wiley & Sons, Hoboken (NJ), 2007, pp. 9-10.
- [8] A.B. Robinson, N.E. Robinson. Origin of metabolomics profiling, in: T.O. Metz (Ed.) *Metabolomic Profiling. Methods and Protocols*, Humana Press / Springer, New York, 2011, pp. 1-24.
- [9] A. Alonso, S. Marsal, A. Julià, Analytical methods in untargeted metabolomics: state of the art in 2015, *Front. Bioeng. Biotechnol.* 3:23 (2015) 3-16, doi: 10.3389/fbioe.2015.00023.

- [10] S.K. Wiedmer T. Hyötyläinen, Selection of analytical methodology for metabolomics, in: T. Hyötyläinen and S. Wiedmer (Eds), *Chromatographic Methods in Metabolomics*, Royal Society of Chemistry, Cambridge, 2013, pp. 1-10.
- [11] M.M.W.B. Hendriks, F.A. van Eeuwijk, R.H. Jellema, J.A. Westerhuis, T.H. Reijmers, H. C.J. Hoefsloot, A.K. Smilde, Data-processing strategies for metabolomics studies, *Trends Anal. Chem.* 30 (2011) 1685-1698.
- [12] B. Worley, R. Powers, *Multivariate Analysis in Metabolomics*, *Curr. Metabolomics*, 1 (2013) 92-107.
- [13] J.M. Cevallos Cevallos, J.I. Reyes de Corcuera, E. Etxeberria, M.D. Danyluk, G.E. Rodrick, Metabolomic analysis in food science: a review. *Trends Food Sci. Tech.* 20 (2009) 557-566.
- [14] K. Dettmer, P.A. Aronov, B.D. Hammock, Mass spectrometry-based metabolomics, *Mass Spectrom. Rev.*, 26 (2007), 51-78.
- [15] J.A. Kirwan, R. J.M. Weber, D.I. Broadhurst. M.R. Viant, Direct infusion mass spectrometry metabolomics dataset: a benchmark for data processing and quality control, *Sci. Data* 1:140012 (2014), doi: 10.1038/sdata.2014.12.
- [16] C.W. Klampfl, M. Himmelsbach, Direct ionization methods in mass spectrometry: An overview, *Anal. Chim. Acta* (in press 2015), doi: 10.1016/j.aca.2015.07.012.
- [17] P. Krishnan¹, N.J. Kruger, R.G. Ratcliffe, Metabolite fingerprinting and profiling in plants using NMR, *J. Exp. Bot.* 56 (2005) 255-265.
- [18] S. Moco, R.J. Bino, R.C.H. De Vos, J. Vervoort, Metabolomics technologies and metabolite identification, *Trends Anal. Chem.* 26 (2007) 855-866.
- [19] C. Cavaliere, E. Corradini, P. Foglia, R. Gubbiotti, R. Samperi, A. Laganà, From metabolomic profiling to metabolomics: fifty years of instrumental and methodological improvements, in: J.S. Knapp, W.L. Cabrera (Eds.),

- Metabolomics: Matabolites, Metabonomics, and Analytical Technologies, Nova Science Publishers, New York, 2011, pp. 121-161.
- [20] L.M. Reid, C.P. O'Donnell, G. Downey, Recent technological advances for the determination of food authenticity, *Trends Food Sci. Tech.* 17 (2006) 344-353.
- [21] D.M.A.M. Luykx, S.M. van Ruth, An overview of analytical methods for determining the geographical origin of food products, *Food Chem.* 107 (2008) 897-911.
- [22] E. Capuano, R. Boerrigter-Eenling, G. van der Veer, S.M. van Ruth, Analytical authentication of organic products: an overview of markers, *J. Sci. Food Agric.* 93 (2013) 12-28.
- [23] E.C. Wilhelmsen, Adulteration determination, in: R.A. Meyers (Ed.), *Encyclopedia of Analytical Chemistry*, John Wiley & Sons, Online © 2006.
- [24] D.A. Hammond, Authenticity of fruit juices, jams and preserves, in: P.R. Ashurst, M.J. Dennis (Eds), *Food authentication*, Chapman & Hall, London, 1996, pp 38-42.
- [25] J. Zhang, X. Zhang, L. Dediu, C. Victor, Review of the current application of fingerprinting allowing detection of food adulteration and fraud in China, *Food Control* 22 (2011) 1126-1135.
- [26] D.I. Ellis, V.L. Brewster, W.B. Dunn, J.W. Allwood, A.P. Golovanovc, R. Goodacre, Fingerprinting food: current technologies for the detection of food adulteration and contamination, *Chem. Soc. Rev.* 41 (2012) 5706-5727.
- [27] M. Bonnefille, Sensory and chemical fingerprinting aids quality and authentication of ingredients and raw materials of vegetal origin, in: K.A. Reynertson K. Mahmood (Eds.), *Botanicals: Methods and Techniques for Quality & Authenticity*. CRC Press, Boca Raton, 2015, pp. 262-278.

- [28] S. Esslinger, J. Riedl, C. Faul-Hassek, Potential and limitations of non-targeted fingerprinting for authentication of food in official control, *Food Res. Int.* 60 (2014) 189-204.
- [29] J. Riedl, S. Esslinger, C. Faul-Hassek, Review of validation and reporting of non-targeted fingerprinting approaches for food authentication, *Anal. Chim. Acta* 885 (2015) 17-32.
- [30] L.A. Berrueta, R.M. Alonso Salces, K. Héberger, Supervised pattern recognition in food analysis, *J. Chromatogr. A*, 1158 (2007) 196-214.
- [31] P. Oliveri, G. Downey, Multivariate class modeling for the verification of food-authenticity claims, *Trends Anal. Chem.* 35 (2012) 74-86.
- [32] B. Vandeginste, Chemometrics in studies of food origin, in: P. Brereton (Ed.) *New Analytical Approaches for Verifying the Origin of Food*, Woodhead, Cambridge, 2013, pp. 117-145.
- [33] F. Marini (Ed.), *Chemometrics in Food Chemistry*, Elsevier, Amsterdam, 2013.
- [34] E. Borràs, J. Ferré b, R. Boqué, M. Mestres, L. Aceña, O. Busto, Data fusion methodologies for food and beverage authentication and quality assessment - A review, *Anal. Chim. Acta* (in press 2015), doi: 10.1016/j.aca.2015.04.042.
- [35] E. Salvatore, M. Bevilacqua, R. Bro, F. Marini, M. Cocchi, Classification methods of multiway arrays as a basic tool for food PDO authentication, in: M. de la Guardia, A. González (eds), *Food Protected Designation of Origin: Methodologies and Applications*, Elsevier, Amsterdam, 2013, pp. 339-382.
- [36] A. Rinnan, J.M. Amigo, T. Skov, Multiway methods in food science, in: D. Granato, G. Ares (eds), *Mathematical and Statistical Methods in Food Science and Technology*, IFT Press/Wiley Blackwell, Chichester, 2014, pp. 143-174.
- [37] G. Picariello, G. Mamone, F. Addeo, P. Ferranti, Novel mass spectrometry-based applications of the 'omic' sciences in food technology and biotechnology. *Food Technol. Biotechnol.* 50 (2012) 286-305.

- [38] A. Cifuentes, Foodomics: Principles and applications, in: A. Cifuentes (Ed.), *Foodomics: Advanced Mass Spectrometry in Modern Food Science and Nutrition*, A. John Wiley & Sons, Hoboken (NJ), 2013, pp. 1-14.
- [39] A. Cifuentes, Food analysis and foodomics. *J. Chromatogr. A*, 1216 (2009) 7109-7110.
- [40] M. Herrero, C. Simó, V. García-Cañas, E. Ibáñez, A. Cifuentes, Foodomics: MS-based strategies in modern food science and nutrition, *Mass Spectrom. Rev.* 31 (2012) 49-69.
- [41] D.S. Wishart, Metabolomics: applications to food science and nutrition research, *Trends Food Sci. Tech.* 19 (2008) 482-493.
- [42] E. Cubero Leon, R. Peñalver, A. Maquet, Review on metabolomics for food authentication, *Food Res. Int.* 60 (2014) 95-107.
- [43] Document N° SANCO/12571/2013, Guidance document on quality control and validation procedures for pesticide residues analysis in food and feed. European Commission Health & Consumer Protection Directorate-General, 2013, pp. 23-24 (annex A).
- [44] M. Castro Puyana, M. Herrero, Metabolomics approaches based on mass spectrometry for food safety, quality and traceability, *Trends Anal. Chem.* 52 (2013) 74-87.
- [45] G. Oms Oliu, I. Odriozola Serrano, O. Martín Belloso, Metabolomics for assessing safety and quality of plant-derived food, *Food Res. Int.* 54 (2013) 1172-1183.
- [46] Nomenclature for Chromatography, IUPAC Recommendations 1993, *Pure Appl. Chem.* 65 (1993) 819-872.
- [47] K.S. Booksh, B.R. Kowalski, Theory of Analytical Chemistry, *Anal. Chem.* 66 (1994) 782A-791A.

- [48] Concepts and applications of the term "dimensionality" in analytical chemistry, IUPAC Technical Report 1993, *Pure Appl. Chem.* 74 (2002) 1479-1487.
- [49] B. Dejaegher, *Multivariate Data Analysis*, in: R.A. Meyers (Ed.), *Encyclopedia of Analytical Chemistry*, John Wiley & Sons, Online © 2014.
- [50] T.J. Smillie, I.A. Khan, A comprehensive approach to identifying and authenticating botanical products, *Clin. Pharmacol. Ther.* 87 (2010) 175-186.
- [51] W.L. Applequist, J.S. Miller, Selection and authentication of botanical materials for the development of analytical methods, *Anal. Bioanal. Chem.* 405 (2013) 4419-4428.
- [52] Z. Yongyu, S. Shujun, D. Jianye, Quality control method for herbal medicine: chemical fingerprint analysis, in: Y. Shoyama (Ed.), *Quality Control of Herbal Medicines and Related Areas*, InTech, Rijeka (Croatia), 2010, pp. 171-194.
- [53] C. Tistaert, B. Dejaegher, Y. Vander Heyden, Chromatographic separation techniques and data handling methods for herbal fingerprints: A review, *Anal. Chim. Acta* 690 (2011) 148-161.
- [54] D.D. Joshi, *Herbal Drugs and Fingerprints*, Springer India, New Delhi, 2012.
- [55] Y-Z. Liang, P.S. Xie, K. Chan, Quality control of herbal medicines, *J. Chromatogr. B* 812 (2004) 53-70.
- [56] T. Hyötyläinen, M. Oresic, Data handling, in: T. Hyötyläinen and S. Wiedmer (Eds), *Chromatographics Methods in Metabolomics*, Royal Society of Chemistry, Cambridge, 2013, pp. 183-194.
- [57] J.M. Bosque Sendra, L. Cuadros Rodriguez, C. Ruiz Samblás, A.P. de la Mata, Combining chromatography and chemometrics for the characterization and authentication of fats and oils from triacylglycerol compositional data - A review. *Anal. Chim. Acta* 724 (2012) 1-11.

- [58] M. Herrero, E. Ibáñez, A. Cifuentes, J. Bernal, Multidimensional chromatography in food analysis, *J. Chromatogr. A*, 1216 (2009) 7110-7129.
- [59] P.Q. Tranchida, P. Donato, F. Cacciola, M. Beccaria, P. Dugo, L. Mondello, Potential of comprehensive chromatography in food analysis, *Trends Anal. Chem.* 52 (2013) 186-205.
- [60] Z. Da Zeng, H.M. Hugel, P.J. Marriott, Chemometrics in comprehensive multidimensional separations, *Anal. Bioanal. Chem.* 401 (2011) 2373-2386.
- [61] K.M. Pierce, B. Kehimkar, L.C. Marney, J.C. Hoggard, R.E. Synovec, Review of chemometric analysis techniques for comprehensive two dimensional separations data, *J. Chromatogr. A*. 1255 (2012) 3-11.
- [62] L.T. Vaz Freire, M.D.R.G. da Silva, M.C. Freitas, Comprehensive two-dimensional gas chromatography for fingerprint pattern recognition in olive oils produced by two different techniques in Portuguese olive varieties Galega Vulgar, Cobrançosa e Carrasquenha, *Anal. Chim. Acta.* 633 (2009) 263-270.
- [63] G. Purcaroa, C. Cordero, E. Liberto, C. Bicchi, L.S. Conte, Toward a definition of blueprint of virgin olive oil by comprehensive two-dimensional gas chromatography, *J. Chromatogr. A*, (2014) 101-111.
- [64] Z. Zeng, J. Li, H.M. Hugel, G. Xu, P.J. Marriott, Interpretation of comprehensive two-dimensional gas chromatography data using advanced chemometrics, *Trends Anal. Chem.* 53 (2014) 150-166.
- [65] J.A. Murray, Qualitative and quantitative approaches in comprehensive two-dimensional gas chromatography, *J. Chromatogr. A*, 1261 (2012) 58- 68.
- [66] M. Cocchi, C. Durante, G. Foca, D. Manzini, A. Marchetti, A. Ulrici, Application of a wavelet-based algorithm on HS-SPME/GC signals for the classification of balsamic vinegars, *Chemom. Intell. Lab. Syst.* 71 (2004) 129-140.
- [67] C. Durante, M. Cocchi, M. Grandi, A. Marchetti, R. Bro, Application of N-PLS to gas chromatographic and sensory data of traditional balsamic vinegars of Modena, *Chemom. Intell. Lab. Syst.* 83 (2006) 54-65.

- [68] W. Pongsuwan, E. Fukusaki, T. Bamba, T. Yonetani, T. Yamahara, A. Kobayashi, Prediction of Japanese green tea ranking by gas chromatography/mass spectrometry-based hydrophilic metabolite fingerprinting, *J. Agric. Food Chem.* 55 (2007) 231-236.
- [69] M.F. Wang, H.Z. Lian, L. Mao, J.P. Zhou, H.J. Gong, B.Y. Qian, et al., Comparison of various extraction methods for policosanol from rice bran wax and establishment of chromatographic fingerprint of policosanol, *J. Agric. Food Chem.* 55 (2007) 5552-5558.
- [70] L-F. Huang, M-J. Wu, K-J. Zhong, X-J. Sun, Y-Z. Liang, Y-H. Dai, K-L. Huang, F-Q. Guo, Fingerprint developing of coffee flavor by gas chromatography-mass spectrometry and combined chemometrics methods, *Anal. Chim. Acta* 588 (2007) 216-223.
- [71] M. Cocchi, C. Durante, M. Grandi, D. Manzini, A. Marchetti, A. Ulrici, Three-way principal component analysis of the volatile fraction by HS-SPME/GC of aceto balsamico tradizionale of Modena, *Talanta* 74 (2008) 547-554.
- [72] W. Pongsuwan, T. Bamba, T. Yonetani, A. Kobayashi, E. Fukusaki, Quality prediction of Japanese green tea using pyrolyzer coupled GC/MS based metabolic fingerprinting quality prediction of Japanese green tea using pyrolyzer coupled GC/MS based metabolic, *J. Agric. Food Chem.* 56 (2008) 744-750.
- [73] K. Jumtee, T. Bamba, E. Fukusaki, Fast GC-FID based metabolic fingerprinting of Japanese green tea leaf for its quality ranking prediction, *J. Sep. Sci.* 32 (2009) 2296-2304.
- [74] E. Aprea, H. Gika, S. Carlin, G. Theodoridis, U. Vrhovsek, F. Mattivi, Metabolite profiling on apple volatile content based on solid phase microextraction and gas-chromatography time of flight mass spectrometry, *J. Chromatogr. A*, 1218 (2011) 4517-4524.
- [75] C. Pizarro, S. Rodríguez Tecedor, N. Pérez del Notario, J.M. González Sáiz, Recognition of volatile compounds as markers in geographical discrimination of Spanish extra virgin olive oils by chemometric analysis of non-specific chromatography volatile profiles, *J. Chromatogr. A*, 1218 (2011) 518-523.

- [76] J. Jiao, N. Ding, T. Shi, X. Chai, P. Cong, Z. Zhu, Study of chromatographic fingerprint of the flavor in beer by HS-SPME-GC, *Anal. Lett.* 44 (2011) 648-655.
- [77] C. Ruiz Samblás, L. Cuadros Rodríguez, A. González Casado, F.P. Rodríguez García, P. de La Mata Espinosa, J.M. Bosque Sendra, Multivariate analysis of HT/GC-(IT)MS chromatographic profiles of triacylglycerol for classification of olive oil varieties, *Anal. Bioanal. Chem.* 399 (2011) 2093-2103.
- [78] K. Jumtee, H. Komura, T. Bamba, E. Fukusaki, Predication of Japanese green tea (Sen-cha) ranking by volatile profiling using gas chromatography mass spectrometry and multivariate analysis, *J. Biosci. Bioeng.* 112 (2011) 252-255.
- [79] J. Guo, T. Yue, Y. Yuan, Feature selection and recognition from nonspecific volatile profiles for discrimination of apple juices according to variety and geographical origin, *J. Food Sci.* 77 (2012) 1090-1096.
- [80] H. Ochi, T. Bamba, H. Naito, K. Iwatsuki, E. Fukusaki, Metabolic fingerprinting of hard and semi-hard natural cheeses using gas chromatography with flame ionization detector for practical sensory prediction modeling, *J. Biosci. Bioeng.* 114 (2012) 506-511.
- [81] H. Parastar, M. Jalali Heravi, H. Sereshti, A. Mani Varnosfaderani, Chromatographic fingerprint analysis of secondary metabolites in citrus fruits peels using gas chromatography-mass spectrometry combined with advanced chemometric methods, *J. Chromatogr. A*, 1251 (2012) 176-187.
- [82] J.S. Ribeiro, F. Augusto, T.J.G. Salva, M.M.C. Ferreira, Prediction models for Arabica coffee beverage quality based on aroma analyses and chemometrics, *Talanta*. 101 (2012) 253-260.
- [83] C. Ruiz Samblás, F. Marini, L. Cuadros Rodríguez, A. González Casado, Quantification of blending of olive oils and edible vegetable oils by triacylglycerol fingerprint gas chromatography and chemometric tools, *J. Chromatogr. B*, 910 (2012) 71-77.
- [84] G. Fang, J.Y. Goh, M. Tay, H.F. Lau, S.F.Y. Li, Characterization of oils and fats by ¹H NMR and GC/MS fingerprinting: Classification, prediction and detection of adulteration, *Food Chem.* 138 (2013) 1461-1469.

- [85] D. Jacobson, A.R. Monforte, A.C.S. Ferreira, Untangling the chemistry of port wine aging with the use of GC-FID, multivariate statistics, and network reconstruction, *J. Agric. Food Chem.* 61 (2013) 2513-2521.
- [86] R. Malheiro, P. Guedes de Pinho, S. Soares, A. César da Silva Ferreira, P. Baptista, Volatile biomarkers for wild mushrooms species discrimination, *Food Res. Int.* 54 (2013) 186-194.
- [87] C. Ruiz Samblás, C. Arrebola Pascual, A. Tres, S. Van Ruth, L. Cuadros Rodríguez, Authentication of geographical origin of palm oil by chromatographic fingerprinting of triacylglycerols and partial least square-discriminant analysis, *Talanta.* 116 (2013) 788-793.
- [88] S. Song, X. Zhang, K. Hayat, Z. Xiao, Y. Niu, K. Eric, Coordinating fingerprint determination of solid-phase microextraction/gas chromatography-mass spectrometry and chemometric methods for quality control of oxidized tallow, *J. Chromatogr. A*, 1278 (2013) 145-152.
- [89] A.E. Springer, J. Riedl, S. Esslinger, T. Roth, M.A. Glomb, C. Fauhl-Hassek, Validated modeling for German white wine varietal authentication based on HS-SPME-GC-MS fingerprinting, *J. Agric. Food Chem.* 62 (2014) 6844-6851.
- [90] H. Tian, P. Zhan, H. Zhang, Development of a fatty acid fingerprint of white apricot almond oil by gas chromatography and gas chromatography-mass spectrometry, *Eur. J. Lipid Sci. Technol.* 116 (2014) 126-133.
- [91] B.R. Toledo, L.W. Hantao, T.D. Ho, F. Augusto, J.L. Anderson, A chemometric approach toward the detection and quantification of coffee adulteration by solid-phase microextraction using polymeric ionic liquid sorbent coatings, *J. Chromatogr. A*, 1346 (2014) 1-7.
- [92] Z. Xiao, S. Liu, Y. Gu, N. Xu, Y. Shang, J. Zhu, Discrimination of cherry wines based on their sensory properties and aromatic fingerprinting using HS-SPME-GC-MS and multivariate analysis, *J. Food Sci.* 79 (2014) C284-C294.
- [93] C. Ruiz-Samblás, J.M. Cadenas, D.A. Pelta, L. Cuadros Rodriguez, Application of data mining methods for classification and prediction of olive oil blends with other vegetable oils, *Anal. Bioanal. Chem.* 406 (2014) 2591-2601.
- [94] S-D. Lv, Y-S. Wu, Y-Z. Song, J-S. Zhou, M. Liang, C. Wang, L. Liu, Q-X. Meng Multivariate Analysis Based on GC-MS Fingerprint and Volatile Composition for the Quality Evaluation of Pu-Erh Green Tea, *Food Anal. Meth.* (2015) 321-333.

- [95] A. M. van Nederkassel, M. Daszykowski, D.L. Massart, Y. Vander Heyden, Prediction of total green tea antioxidant capacity from chromatograms by multivariate modeling, *J. Chromatogr. A*, 1096 (2005) 177-186.
- [96] N.H. Beltrán, M. a. Duarte-Mermoud, M. a. Bustos, S. a. Salah, E. a. Loyola, a. I. Peña-Neira, et al., Feature extraction and classification of Chilean wines, *J. Food Eng.* 75 (2006) 1-10.
- [97] A.M. van Nederkassel, C.J. Xu, P. Lancelin, M. Sarraf, D. a. MacKenzie, N.J. Walton, et al., Chemometric treatment of vanillin fingerprint chromatograms. Effect of different signal alignments on principal component analysis plots, *J. Chromatogr. A*, 1120 (2006) 291-298.
- [98] M. Daszykowski, Y. Vander Heyden, B. Walczak, Robust partial least squares model for prediction of green tea antioxidant capacity from chromatograms, 1176 (2007) 12-18.
- [99] M. Dumarey, A.M. van Nederkassel, E. Deconinck, Y. Vander Heyden, Exploration of linear multivariate calibration techniques to predict the total antioxidant capacity of green tea from chromatographic fingerprints, *J. Chromatogr. A*, 1192 (2008) 81-88.
- [100] J. Chen, Y-H. Lu, D-Z. Wei, X-L. Zhou, Establishment of a fingerprint of raspberries by LC, *Chromatographia*. 70 (2009) 981-985.
- [101] L. Zheng, D.G. Watson, B.F. Johnston, R.L. Clark, R. Edrada Ebel, W. Elseheri, A chemometric study of chromatograms of tea extracts by correlation optimization warping in conjunction with PCA, support vector machines and random forest data modeling, *Anal. Chim. Acta*. 642 (2009) 257-265.
- [102] M. Dumarey, I. Smets, Y. Vander Heyden, Prediction and interpretation of the antioxidant capacity of green tea from dissimilar chromatographic fingerprints, *J. Chromatogr. B*, 878 (2010) 2733-2740.
- [103] E. Mattarucchi, M. Stocchero, J.M. Moreno-Rojas, G. Giordano, F. Reniero, C. Guillou, Authentication of trappist beers by LC-MS fingerprints and multivariate data analysis, *J. Agric. Food Chem.* 58 (2010) 12089-12095.
- [104] P. de la Mata Espinosa, J.M. Bosque Sendra, R. Bro, L. Cuadros Rodríguez, Olive oil quantification of edible vegetable oil blends using triacylglycerols chromatographic fingerprints and chemometric tools, *Talanta*. 85 (2011) 177-182.

- [105] G. Tomasi, F. Savorani, S.B. Engelsen, Icoshift: An effective tool for the alignment of chromatographic data, *J. Chromatogr. A*, 1218 (2011) 7832-7840.
- [106] Y. Zhao, P. Chen, L. Lin, J.M. Harnly, L. Yu, Z. Li, Tentative identification, quantitation, and principal component analysis of green pu-erh, green, and white teas using UPLC/DAD/MS, *Food Chem.* 126 (2011) 1269-1277.
- [107] C. Sârbu, R.D. Nacu-Briciu, A. Kot-Wasik, S. Gorinstein, A. Wasik, J. Namieśnik, Classification and fingerprinting of kiwi and pomelo fruits by multivariate analysis of chromatographic and spectroscopic data, *Food Chem.* 130 (2012) 994-1002.
- [108] D. Shen, Q. Wu, W.J. Sciarappa, J.E. Simon, Chromatographic fingerprints and quantitative analysis of isoflavones in Tofu-type soybeans, *Food Chem.* 130 (2012) 1003-1009.
- [109] G. Alaerts, J. Van Erps, S. Pieters, M. Dumarey, A. M. van Nederkassel, M. Goodarzi, J. Smeyers-Verbeke, Y. Vander Heyden, Similarity analyses of chromatographic fingerprints as tools for identification and quality control of green tea, *J. Chromatogr. B.* 910 (2012) 61-70.
- [110] D. Serrano Lourido, J. Saurina, S. Hernández Cassou, A. Checa, Classification and characterisation of Spanish red wines according to their appellation of origin based on chromatographic profiles and chemometric data analysis, *Food Chem.* 135 (2012) 1425-1431.
- [111] R. Nescatelli, R.C. Bonanni, R. Bucci, A.L. Magrì, A.D. Magrì, F. Marini, Chemometrics and intelligent laboratory systems geographical traceability of extra virgin olive oils from Sabina PDO by chromatographic fingerprinting of the phenolic fraction coupled to chemometrics, *Chemom. Intell. Lab. Syst.* 139 (2014) 175-180.
- [112] A.E. Schulze, D. De Beer, M. Manley, E. Joubert, Chemometric analysis of chromatographic fingerprints shows potential of *Cyclopia Maculata* (Andrews) kies for production of standardized extracts with high xanthone content, *J. Agric. Food Chem.* 62 (2014) 10542-10551.
- [113] G.M. Leme, I.D. Coutinho, S. Creste, O. Hojo, R.L. Carneiro, V. Bolzani, et al., HPLC-DAD method to metabolic fingerprinting to the phenotyping of sugarcane genotypes, *Anal. Methods.* 6 (2014) 7781-7788.

- [114] J. Zhou, L. Yao, Y. Li, L. Chen, L. Wu, J. Zhao, Floral classification of honey using liquid chromatography-diode array detection-tandem mass spectrometry and chemometric analysis, *Food Chem.* 145 (2014) 941-949.
- [115] M.G. Bagur González, E. Pérez Castaño, M. Sánchez Viñas, D. Gázquez Evangelista, Using the liquid-chromatographic-fingerprint of sterols fraction to discriminate virgin olive from other edible oils, *J. Chromatogr. A*, 1380 (2015) 64-70.
- [116] P.M. Kuś, S. van Ruth, Discrimination of Polish unifloral honeys using overall PTR-MS and HPLC fingerprints combined with chemometrics, *LWT - Food Sci. Technol.* 62 (2015) 69-75.
- [117] J. Li, X. He, M. Li, W. Zhao, L. Liu, X. Kong, Chemical fingerprint and quantitative analysis for quality control of polyphenols extracted from pomegranate peel by HPLC, *Food Chem.* 176 (2015) 7-11.
- [118] X. He, J. Li, W. Zhao, R. Liu, L. Zhang, X. Kong, Chemical fingerprint analysis for quality control and identification of Ziyang green tea by HPLC, *Food Chem.* 171 (2015) 405-411.
- [119] M. Arbulu, M.C. Sampedro, A. Gómez-Caballero, M. A. Goicolea, R.J. Barrio, Untargeted metabolomic analysis using liquid chromatography quadrupole time-of-flight mass spectrometry for non-volatile profiling of wines, *Anal. Chim. Acta.* 858 (2015) 32-41.
- [120] E. Pérez Castaño, C. Ruiz Samblás, S. Medina Rodríguez, V. Quirós Rodríguez A.M. Jiménez Carvelo, L. Valverde Som, A. Gonzalez Casado, L. Cuadros Rodriguez, Comparison of different analytical classification scenarios: application for the geographical origin of edible palm oil by sterolic (NP)HPLC fingerprinting, *Anal. Meth.* 7 (2015) 4192-4201.

2.4. ARTÍCULO 2

Sobre el uso de 'huellas dactilares' cromatográficas en la autenticación de alimentos

(Artículo publicado en Noviembre 2015 en la revista no indexada: Boletín GRASEQA 'Huellas dactilares' mediante el uso de técnicas cromatográficas avanzadas N° 12, del Grupo Regional Andaluz Sociedad Española de Química Analítica)

INVESTIGACIÓN GRASEQA: 'Huellas dactilares'

Sobre el uso de 'huellas dactilares' cromatográficas en la autenticación de alimentos

LUIS CUADROS RODRÍGUEZ[✉], M^a GRACIA BAGUR GONZÁLEZ, ANTONIO GONZÁLEZ CASADO,
CRISTINA RUIZ SAMBLÁS, ESTEFANÍA PÉREZ CASTAÑO, LUCIA VALVERDE SOM, ANA MARÍA JIMÉNEZ
CARVELO

UNIVERSIDAD DE GRANADA. Departamento de Química Analítica, Facultad de Ciencias,
c/ Fuentenueva, s.n., 18071 Granada
E-mail: lcuadros@ugr.es

NOTA DE LOS AUTORES. Este artículo es un extracto parcial de algunos apartados recogidos en un tutorial que ha sido enviado a una revista de química analítica de alto impacto, y que actualmente está en revisión. Aquellos lectores interesados en profundizar en este tema podrán encontrarlo próximamente en las bases de datos introduciendo palabras clave como: "chromatographic fingerprinting" y "food authentication"

1. Autenticación de alimentos

La autenticación (o autenticación) de alimentos implica la comprobación de la veracidad de sus especificaciones declaradas, lo que puede incluir

muchos aspectos como es la determinación analítica de componentes característicos, adulterantes y/o contaminantes, la verificación de los requisitos de calidad diferenciada, o el método de preparación o procesado de dicho alimento [1,2,3]. La autenticidad está ligada a la veracidad y, por tanto, un alimento se considera auténtico (o genuino) cuando no está afectado por ningún tipo de fraude.

1.1. Enfoques y metodologías

La autenticación de alimentos y la detección de fraudes alimentarios son problemas similares que siguen estrategias analíticas paralelas. En un principio, se pueden diferenciar tres enfoques analíticos dependiendo del nivel en la estructura de la materia del que se obtiene la información diana: el nivel 1 obtiene la información a partir de la composición química del alimento (enfoque de componentes químicos, o simplemente enfoque químico); el nivel 2 trabaja a nivel biomolecular, sobre la estructura de proteínas o la secuencia del DNA (enfoque biomolecular); y el nivel 3 está enfocado sobre el estudio de la composición isotópica estable de determinados átomos (enfoque isotópico).

En la Figura 1 se muestra un esquema de los diferentes enfoques.

El enfoque químico (nivel 1) a su vez puede orientarse en tres direcciones que dan lugar a diferentes metodologías:

(a) contenidos de especies químicas características, de naturaleza diversa, y obtenidas mediante técnicas o métodos analíticos iguales o diferentes. Un caso particular de esta metodología está constituida por el

uso de marcadores químicos, definidos como determinados componentes que son específicamente representativos de una característica de calidad directamente relacionada con la autenticidad;

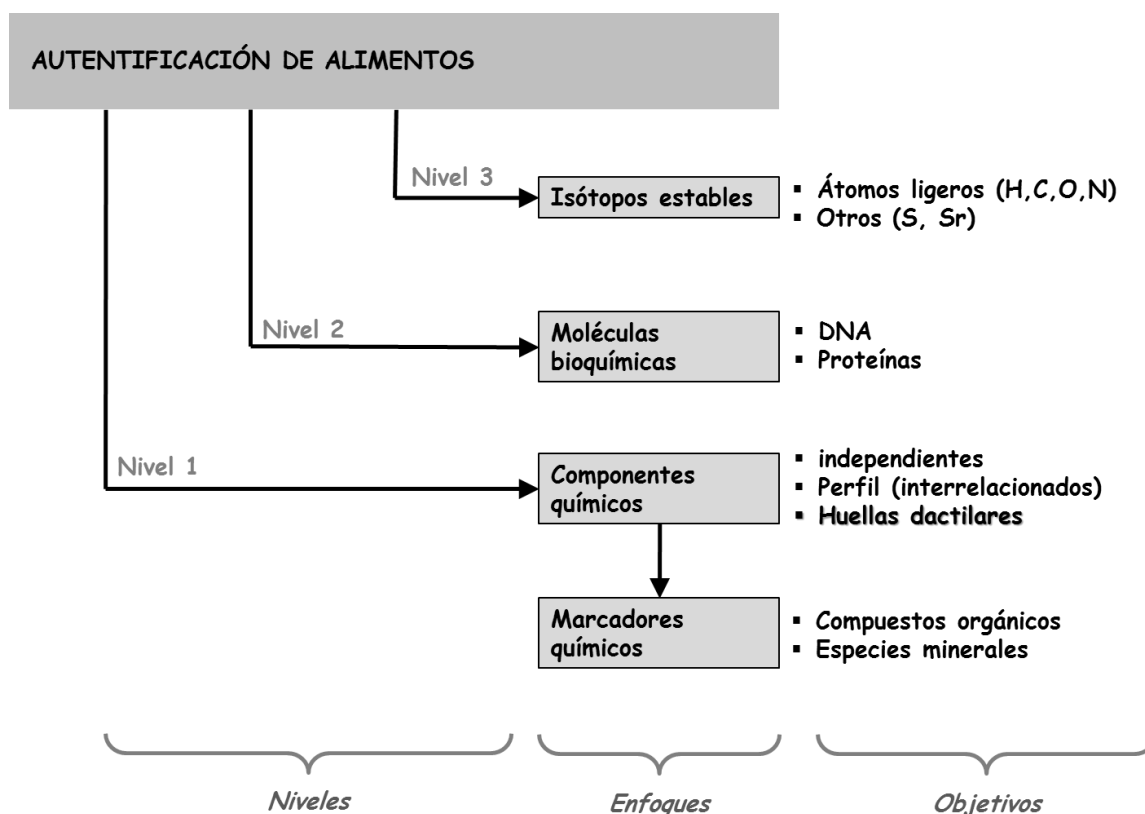


Figura 1. Enfoques analíticos en la autenticación de alimentos según niveles.

(b) perfiles de componentes, definido como un *conjunto de valores que describen los contenidos de una serie de constituyentes relacionados entre sí*, que comparten alguna propiedad física (p.e.: compuestos volátiles, fracción soluble en un determinado disolvente) o química (p.e.: compuestos polares), o incluso que pertenecen a una misma familia química (p.e.: compuestos fenólicos, ácidos grasos, fitosteroles, etc.). Generalmente los perfiles de componentes se obtienen por aplicación de un único método

analítico. Si los contenidos de los constituyentes de interés se expresan como una proporción relativa, entonces se denomina perfil composicional y éste puede estar expresado como contenidos o como relaciones entre intensidades de señal previamente normalizada, p.e.: las áreas o alturas de pico relativas con respecto a la de un patrón interno;

(c) huellas dactilares instrumentales que pueden ser definidas como *todo registro instrumental característico que refleja la composición química compleja de la muestra analizada y que puede ser relacionado con su autenticidad*. Dependiendo del carácter de la información asociada a las huellas instrumentales, esta se puede dividir en dos grupos: (c.1) poco selectivas o inespecíficas: aquellas obtenidas de una señal en las que no se puede diferenciar información individual sobre los diferentes componentes (p.e.: los espectros moleculares); y (c.2) selectivas: las señales están lo suficientemente resueltas como para poder obtener información de componentes individuales o de grupos formados por compuestos de una misma familia. Tradicionalmente, por sus propias características, las señales cromatográficas (cromatograma) se han asociado al grupo de las 'huellas dactilares' selectivas. Pero si se sacrifica la resolución cromatográfica a favor de variables no operacionales de tipo práctico, como tiempo de análisis o coste, es posible obtener señales cromatográficas con un alto grado de inespecificidad.

La aplicación metodológica tradicional de la autenticación de alimentos se ha basado más en demostrar la ausencia de fraude que en verificar la propia autenticidad por lo que, en principio, lo habitual es determinar el valor de uno o más parámetros característicos del fraude y si la respuesta

es positiva, el alimento se declara adulterado, mientras que si la respuesta es negativa el alimento se declara auténtico; ésta constituye la denominada orientación metodológica "negativa" de la autentificación. Las metodologías (a) y (b) están más dirigidas hacia dicha orientación y por ello requieren la aplicación de técnicas analíticas con suficiente sensibilidad y selectividad para detectar, identificar y cuantificar (en su caso) los diferentes compuestos o especies químicas características y, por tanto, para ello es imprescindible disponer de patrones adecuados de cada compuesto. La autentificación se realiza por aplicación de reglas o árboles de decisión previamente desarrollados sobre la base del conocimiento de la composición de los alimentos genuinos.

Sin embargo, recientemente está cobrando cada vez más interés la aplicación de una orientación metodológica "positiva" dirigida directamente a poner de manifiesto la autenticidad del alimento en estudio. La metodología de huellas dactilares instrumentales está enfocada hacia en dirección, ya que considera de forma global la señal suministrada y registrada por un instrumento analítico, p.e.: un espectro o un cromatograma, obtenido directamente sobre el producto de interés o sobre una de sus fracciones. El registro constituye la huella dactilar instrumental del producto, o de la fracción considerada, ya que es característica de él y refleja de forma implícita su composición química.

Las huellas dactilares pueden ser generadas a partir de diferentes técnicas analíticas aunque las más utilizadas han sido las técnicas espectrométricas (ópticas o de masas, en sus diversas modalidades) [4,5], aunque más recientemente se han incorporado técnicas basadas en

sensores, como las denominadas nariz o lengua electrónica [6]. La aplicación de esta metodología no requiere selectividad en el instrumento ya que la caracterización del material se realiza de forma no selectiva, pero sí suficiente sensibilidad [7,8]. Su aplicación práctica requiere disponer de muestras de productos previamente declarados auténticos, en cantidad suficiente para constituir una base de datos representativos de la población de alimentos genuinos.

2. Huellas dactilares cromatográficas: marco metodológico

La versatilidad de los métodos cromatográficos permite al químico analítico interactuar tanto sobre la etapa de separación como sobre la de medida con el objetivo de buscar las condiciones que permitan obtener una señal analítica que recoja el máximo de información útil, y que luego puede ser extraída y filtrada en una etapa posterior de tratamiento de datos.

Como consecuencia, estos métodos son excelentes candidatos potenciales para obtener huellas dactilares características. El marco metodológico para la aplicación de la cromatografía en este contexto ya fue descrito previamente por los autores [9], aunque a continuación se muestra un diagrama actualizado en la Figura 2.

A la vista de la figura puede deducirse que los métodos cromatográficos son adecuados para aplicar cualquiera de las metodologías propias del denominado "enfoque químico" de la autenticación de alimentos descritas en el apartado 1.1.

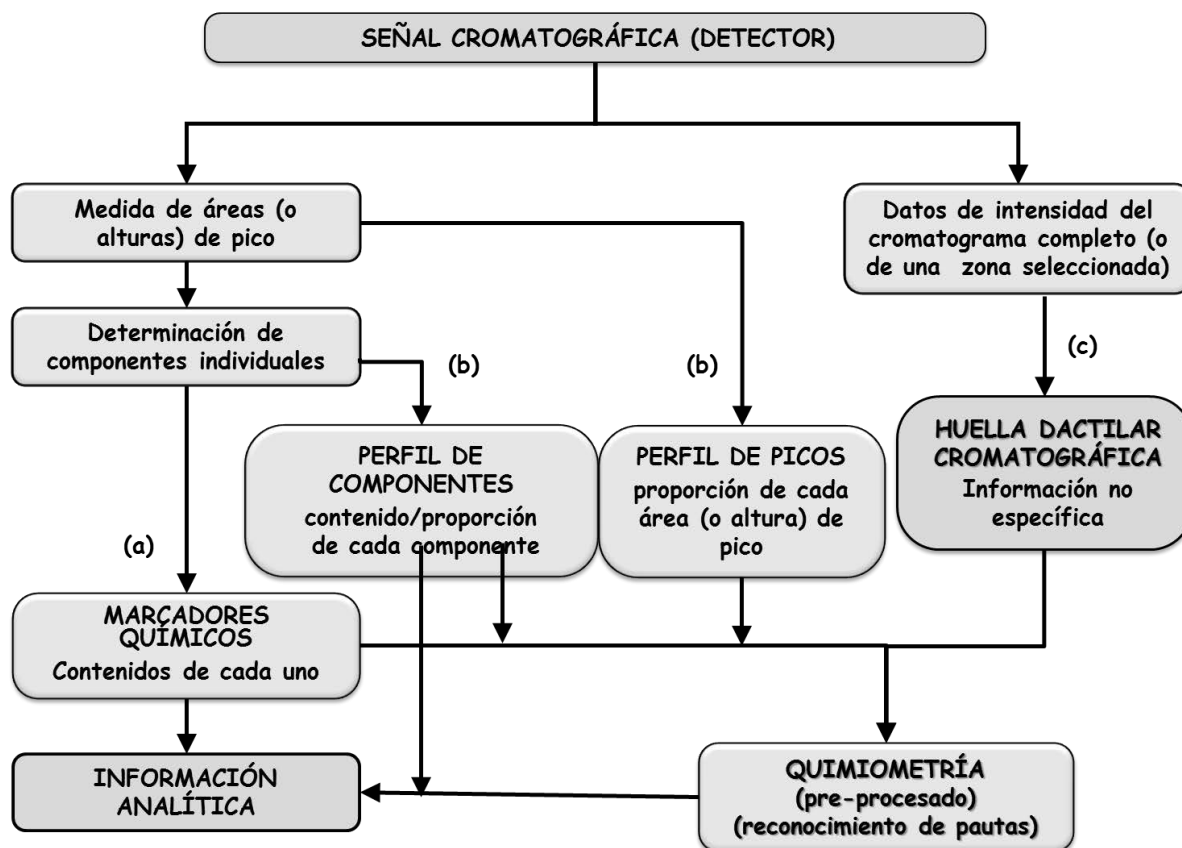


Figura 2. Autentificación de alimentos mediante el empleo de métodos cromatográficos.

No obstante, hay que tener en cuenta que, dependiendo de la metodología a aplicar, el objetivo del método cromatográfico puede ser diferente, lo cual en sí mismo podría constituir un problema si el analista no decide previamente el verdadero objetivo a aplicar. Efectivamente, las metodologías basadas en el uso de marcadores químicos o de perfiles de componentes, que son coherentes con la cromatografía tradicional, requieren que el método cromatográfico se optimice con el objetivo fundamental de encontrar la suficiente separación (resolución de picos) entre los diferentes constituyentes químicos de modo que, idealmente, las

señales sean altamente específicas y describan la composición del material que está siendo analizado. En otras palabras, se busca que cada pico de información sobre un único componente.

Sin embargo, el uso de huellas dactilares cromatográficas implica un cambio metodológico en el paradigma fundamental de la cromatografía, ya que la velocidad de análisis se constituye en el factor prioritario a la hora de realizar la optimización del método en detrimento de la propia separación de los componentes.

Ésta es la única forma que la cromatografía puede competir con los métodos espectrométricos en la aplicación de la metodología de huellas dactilares. En base a nuestra experiencia, los tiempos de análisis cromatográfico no deberían superar los 10 minutos, e incluso, en un marco ideal, debe intentarse que ronden los 5 minutos.

El reto analítico estriba en que, obviamente, el cromatograma obtenido en este corto intervalo de tiempo debe contener suficiente información implícita, aunque no sea evidente, para permitir la caracterización del alimento en estudio. Ello implica, por ejemplo, obtener cromatogramas en forma de "joroba" con una silueta irregular, a modo de "dientes de sierra" que dejen intuir picos cromatográficos solapados y no resueltos.

2.1. Señales cromatográficas y datos analíticos

Antes de continuar, es necesario aclarar algunos términos que luego serán utilizados.

La señal analítica se refiere al registro cuantitativo que un instrumento analítico adquiere mientras realiza la medida continua de una propiedad física o química que, en el caso de los detectores cromatográficos, siempre es un cambio en la composición del eluyente. Cada señal analítica puede ser representada matemáticamente como un tensor de datos de diferente dimensionalidad (escalar, vector, matriz, ..), dependiendo del número de variables que define cada uno de los elementos (números) que constituyen dicho tensor [10,11]. Así por ejemplo, un cromatograma convencional es un registro de dos dimensiones que puede ser representado mediante un vector de datos donde cada elemento representa la intensidad de la señal y es caracterizado por una única coordenada que es el tiempo de retención; por tanto un cromatograma constituye un tensor de orden 1 (un vector). Por el contrario, un espectro-cromatograma UV-Vis o de masas, obtenidos respectivamente por un detector de fila de diodos (DAD) o de espectrometría de masas (MS) acoplados a un cromatógrafo constituye una señal de tres dimensiones que da lugar a un tensor de orden 2 (una matriz) ya que cada elemento representa una intensidad para cada valor de tiempo de retención y longitud de onda de absorción o relación masa/carga. Un escalón más arriba estaría constituido, por ejemplo, por los espectro-cromatogramas de fluorescencia, donde se obtiene señales correspondientes a tensores de orden 3 (un cubo), puesto que cada elemento está definido por el tiempo de retención y las correspondientes longitudes de onda de excitación y emisión.

De forma similar, la cromatografía bidimensional desarrollada en equipos que incorporan dos etapas cromatográficas ortogonales, usualmente

denominada cromatografía 2D exhaustiva ($GC \times GC$, $LC \times LC$, u otras combinaciones) dan lugar a cromatogramas tridimensionales que se corresponden con señales de orden 2. Éstas se suelen representar mediante mapas de contorno o mapas de color, ya que cada valor de intensidad está caracterizado por dos tiempos de retención correspondientes a cada una de las etapas cromatográficas. Obviamente, si la segunda etapa está además acoplada a un espectrómetro de masas o un espectrofotómetro de absorción de fila de diodos, se podrían disponer de señales de orden 3.

A partir de la señal analítica se obtienen datos numéricos, también en forma de tensores, que serán sometidos a procesado (tratamiento de datos) para obtener los correspondientes resultados. Esta conversión de señal en datos puede llevarse a cabo mediante varias estrategias que dan lugar a tensores de diferentes características: (A) un tensor de datos del mismo orden y elementos de la señal original; (B) un tensor del mismo orden pero con menos elementos (reducción de datos); (C) un tensor de datos de orden inferior pero con el mismo número de elementos de la señal original en las dimensiones retenidas; y (D) un tensor de datos de orden inferior al de la señal original y con un número de datos reducido (se corresponde con la yuxtaposición de los resultados de las estrategias B y C).

La reducción de datos a la que se hacer referencia en las estrategias B y D puede realizarse mediante diferentes vías. Algunas de ellas, de carácter genérico, se aplican sobre toda la señal y, entre ellas, la más fácil de aplicar es el muestreo de la señal, por ejemplo, mediante un proceso de

decimación. Sin embargo, hay otras metodologías más específicas: (B1) selección de una zona del cromatograma entre dos valores de tiempo de retención; (B2) recopilación de un único valor representativo para cada pico cromatográfico; y (B3) selección, por aplicación de un algoritmo adecuado, de los elementos del tensor original que contienen la información significativa, por ejemplo, por aplicación de un análisis de componentes principales o de variables latentes. La estrategia A no supone transformación de la señal original mientras que el resto implica que el operador actúa sobre la misma.

Si se considera un caso sencillo dado por una señal de dimensión 2 representada por el cromatograma obtenido mediante un sistema cromatográfico GC/FID, la aplicación de la estrategia A genera un vector de datos (un tensor de orden 1) constituido por tantos elementos como datos haya adquirido el detector, normalmente varios miles. La estrategia B1 deriva de extraer en el anterior vector los elementos que definen uno (o más) de los picos cromatográficos previamente seleccionado. La estrategia B2, posiblemente la más utilizada, ocurre cuando a partir de un cromatograma se obtiene un vector de datos constituido únicamente por las áreas (o alturas) de los picos cromatográficos de interés, o bien por las concentraciones de los diferentes analitos después de una etapa de cuantificación. En cambio, la estrategia B3 da lugar a un vector de pocos elementos que recogen la mayor parte de la información encerrada en el cromatograma original sobre la propiedad en estudio. Y por último, las estrategias C y D conducen, en este ejemplo, al mismo resultado y corresponde con el caso de que solo se considere uno de los picos

cromatográficos, por lo que se obtiene un único valor de área de pico o de concentración que supone un tensor de orden 0 (un escalar). La Figura 3 esquematiza gráficamente las estrategias descritas para el ejemplo considerado.

En el caso de señales tridimensionales, como podría ser un cromatograma obtenido mediante 2D GC (o GC×GC) con un detector convencional, se abren un gran número de nuevas posibilidades a la hora de aplicar las diferentes estrategias y que no serán explicadas aquí.

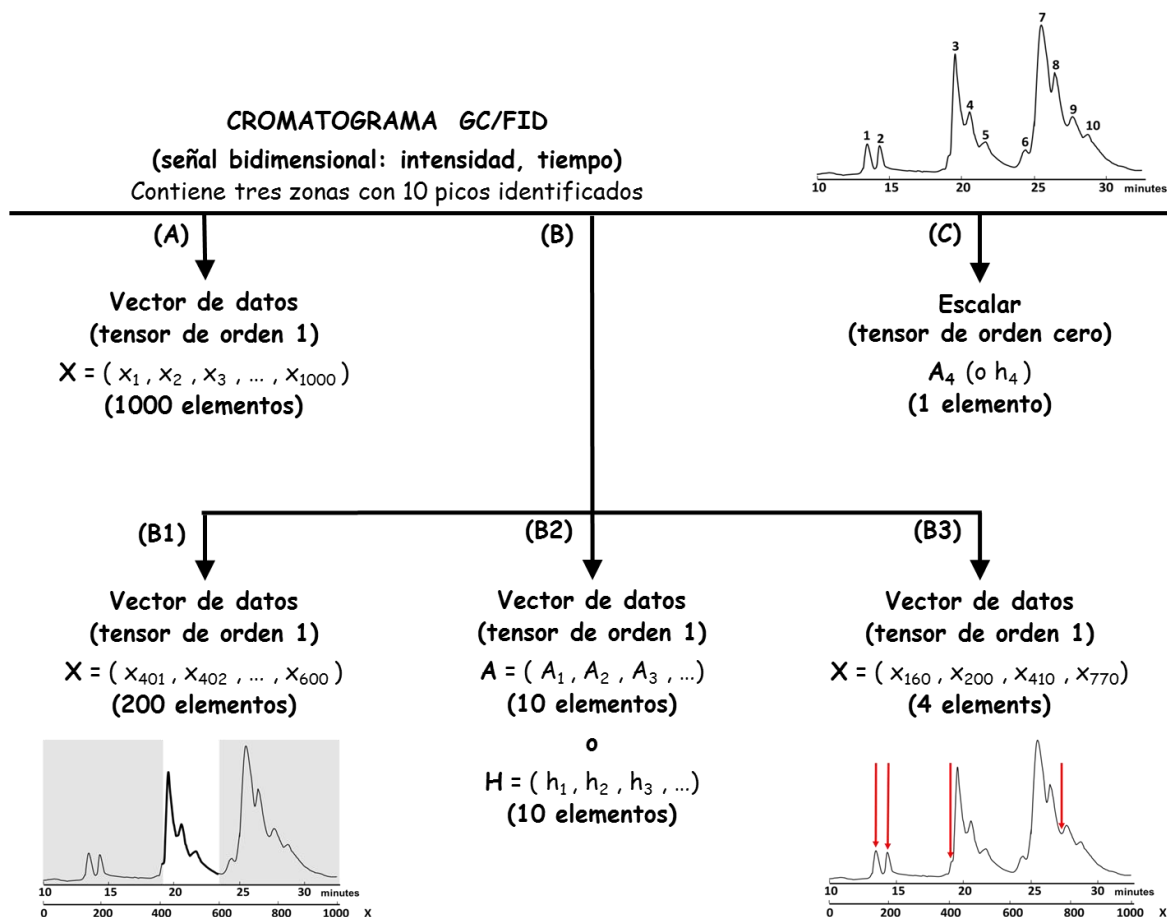


Figura 3. Estrategias para la transformación de señales cromatográficas en datos numéricos.

2.2. Perfiles cromatográficos y huellas dactilares cromatográficas

Es posible encontrar al menos dos significados para el término perfil que pueden ser aplicados en el ámbito de la cromatografía. En el diccionario de la RAE (22^a edición) aparecen los dos siguientes: (P1) *contorno aparente de una figura, representado por líneas que determinan la forma de aquella;* y (P2) *conjunto de rasgos peculiares que caracterizan a alguien o algo.*

El significado P1 está relacionado con la forma de un objeto y mientras que el significado P2 se centra en las características que lo hacen diferente. Por tanto, el término "perfil cromatográfico" puede entenderse como la forma que toma el contorno de un cromatograma (P1) o el conjunto de picos que lo caracterizan (P2). P1 conecta mejor con el concepto de señal mientras que P2 lo hace con los datos. Esta doble forma de entender el término puede generar confusión ya que diversos autores pueden utilizar el término con uno u otro significado y, por tanto, siempre es necesario analizar el contexto en que se está utilizando para encontrar el significado utilizado en cada caso. En este tutorial, siempre se va a utilizar el término con el segundo significado (P2) de modo que perfil cromatográfico se referirá, por ejemplo, al conjunto de picos que caracterizan un cromatograma, y vendrá definido por un conjunto de áreas (o alturas) de pico, o de sus respectivos valores transformados en concentraciones o fracciones másicas. Por lo tanto, el perfil cromatográfico corresponde a un tensor de datos obtenido a partir de la señal cromatográfica (o cromatograma) por aplicación de la estrategia B2 antes definida (véase Figura 3).

El término "huella dactilar" se utiliza en Química Analítica con un significado que es una extensión del uso forense que se le da a las huellas dactilares, o más propiamente, a las impresiones dactilares en la identificación de personas. Por tanto, en su significado original, una impresión dactilar constituye un registro característico de la identidad de una determinada persona pero tiene dos condicionantes: (i) no describe atributos antropomórficos de dicha persona, como la altura y el peso, el color de los ojos o del pelo, etc.; y (ii) por sí sola, no puede ser utilizada para la identificación ya que ésta se basa en una comparación con un registro previo obtenido de la misma persona y que debe estar almacenado en una base de datos.

La huella dactilar cromatográfica se refiere al cromatograma completo obtenido de un determinado sistema material y que es característico de su composición, pero que no muestra información explícita sobre dicha composición u otras propiedades físicas o químicas. Posiblemente el uso más inmediato de dicho cromatograma es el de establecer un rasgo específico y diferenciador que permita caracterizar la identidad de un determinado sistema material, a modo de DNI. Por ejemplo, con este objetivo se aplica actualmente para diferenciar diferentes productos botánicos [12,13].

En algunas ocasiones, por ejemplo en análisis fitoquímico, se utilizan conjuntamente de forma correcta ambos términos, como "perfil de huellas dactilares", para referirse al cromatograma donde el término perfil es utilizado ahora con el significado P1 antes comentado [14]. No obstante, y para evitar confusiones, ambos términos deberían de utilizarse siempre en

relación a datos analíticos derivados de la correspondiente señal cromatográfica (véase Figura 2).

La diferencia conceptual más importante estriba en el tipo de información analítica que conllevan, ya que un perfil implica información específica sobre determinados componentes químicos mientras que una huella dactilar encierra una información inespecífica y no evidente que debe ser extraída mediante la aplicación de métodos y técnicas propias de la quimiometría.

2.3. Métodos y técnicas quimiométricas

La quimiometría es una ciencia aun joven que nace a principios de la década de los años 1970 [15], y que fue definida inicialmente como *la parte de la química que se sirve de las matemáticas, estadística y lógica formal para: diseñar o seleccionar procedimientos experimentales óptimos, proporcionar información química relevante a partir del análisis de señales analíticas y, finalmente adquirir conocimiento de los sistemas químicos* [16].

Más recientemente, en 2009, se establece que la quimiometría es *la ciencia que relaciona las medidas realizadas sobre un sistema químico o proceso con el estado del sistema mediante la aplicación de métodos matemáticos o estadísticos* [17]. Como se puede advertir, esta definición no hace mención al propio proceso de la medida, íntimamente unido al uso que se va a hacer de los resultados de dicha medida [18], por lo que la quimiometría queda como una disciplina complementaria a la metrología y, por extensión, a la cualimetría. De esta forma se puede generar una jerarquía de aplicación

por el cual las tres disciplinas quedan concatenadas según el siguiente esquema:

- Metrología: realización de medidas sobre un sistema químico y obtención de datos.
- Quimiometría: relación entre los datos obtenidos en las medidas y el estado del sistema químico.
- Cualimetría: evaluación y descripción de la calidad de los resultados generados.

Menos compartimentada es la más reciente propuesta de EuChMS-DAC, por la que la quimiometría es *la ciencia de extraer información de sistemas químicos aplicando para ello métodos de análisis y tratamiento de datos* [19].

Este acercamiento quimiométrico al conocimiento químico es consecuencia del hecho constatado de que la mayoría de los sistemas químicos son de naturaleza multivariable (o multivariante) por lo que las metodologías tradicionales de carácter univariable no son siempre capaces de proporcionar soluciones óptimas, como quedó refrendado en el antológico, aunque no por ello demasiado conocido, tutorial de Booksh y Kowalski en el que se establecen las bases de la Teoría de la Química Analítica [9]. La instrumentación analítica actual y el análisis quimiométrico facilitan ver los sistemas químicos desde una perspectiva multivariable y, en este sentido, las herramientas de naturaleza matemática-estadística son utilizadas para establecer relaciones no evidentes entre los datos de las medidas y el estado del sistema en estudio. La quimiometría como ciencia puede ser

resumida en una única palabra: multivariable [20], y son precisamente los desarrollos en quimiometría los que dan sustento a muchos de los avances reales de la química analítica actual.

La quimiometría y la cromatografía constituyen dos potentes herramientas analíticas y posiblemente, junto con la espectrometría de masas, aúnan méritos para ser consideradas pilares básicos responsables de la actual pujanza de la Química Analítica en la mayoría de los campos científicos y técnicos del mundo actual. Sin embargo, ambas disciplinas, aunque se necesitan en la resolución de problemas analíticos complejos y generalmente se complementan [21-24], no se llevan demasiado bien y se miran mutuamente con recelo y desconfianza.

La metodología basada en huellas dactilares, requieren el uso de herramientas quimiométricas [25-28], imprescindibles para encontrar tendencias o información significativa en los datos, y que permiten el desarrollo de modelos de reconocimiento de pautas para exploración, clasificación-discriminación y predicción. El abanico de métodos a aplicar es muy amplio, y todos ellos se pueden encuadrar en dos grandes grupos: métodos matemático-estadísticos: propios de la quimiometría clásica, y métodos de minería de datos resultantes de combinación de métodos de inteligencia computacional, aprendizaje automático ("machine learning"), estadística y sistemas de bases de datos.

Un paso más en el uso de la Quimiometría para mejorar la calidad y fiabilidad de las conclusiones obtenidas a partir de datos experimentales (huellas dactilares) es la aplicación de los métodos de fusión de datos que

implican la generación de modelos a partir de la combinación (fusión) de las señales obtenidas de diferentes técnicas instrumentales instrumentos o métodos analíticos. El principal objetivo de la fusión de datos estriba en aumentar la sinergia entre los diferentes conjuntos de datos a fusionar, y que han sido obtenidos por instrumentos o métodos analíticos suplementarios, de modo que se logre una mayor calidad en los resultados del proceso de autenticación [29].

Y por último, un nivel superior en la etapa de obtención y tratamiento de datos está enfocado al uso de señales con una estructura multidimensional, igual o superior a 3 (señales 3D o mayor) producidos generalmente por espectrometría de fluorescencia molecular o por sistemas constituidos por instrumentos acoplados, por ejemplo cromatografía-espectrometría de masas. El tratamiento de datos requiere el uso de técnicas quimiométricas multivía (o N-vía). Recientemente, las aplicaciones de tales métodos en ciencia de los alimentos han sido revisadas [30,31].

2.4. 'Identificación' de alimentos

La verificación de la autenticidad de un alimento mediante el análisis de las huellas dactilares cromatográficas se basa en la constatación de que la señal instrumental obtenida es coherente, dentro de la variabilidad muestral e instrumental, con una señal de referencia. Por tanto, su aplicación no es inmediata sino que requiere necesariamente la consecución

de una etapa previa que implica el establecimiento de una huella dactilar instrumental característica de la autenticidad.

Para denominar este proceso previo, hemos propuesto el uso del término 'identificación' de alimentos. La necesidad de este neologismo está sustentada por el hecho de que no resulta adecuado utilizar con este fin la palabra 'identificación', que involucra una acción diferente. En efecto, identificación, implica una comparación, de forma similar a la autenticación. El diccionario de la RAE (22^a edición) define identificación como *la acción y efecto de hacer que dos o más cosas, en realidad distintas, aparezcan y se consideren como una misma*. Por tanto, el término 'identificación' de alimentos implica la acción previa de *definir la identidad de un tipo de alimento en base a aquellas características que lo hacen singular o único*. También podría utilizarse con el mismo significado la expresión 'singularización de identidad' de un alimento.

Esto significa que 'identificación' de alimentos y autenticación de alimentos implican dos acciones independientes y consecutivas, de forma que la segunda no es posible sin que la primera haya sido realizada en su totalidad. Además, el nivel de fiabilidad y la confianza sobre las conclusiones de la autenticación de un alimento depende de cómo de bien se haya llevado a cabo su 'identificación', por lo que esta etapa es crucial aunque generalmente es obviada en las publicaciones científicas sobre autenticación de alimentos.

La 'identificación' es aplicable un alimento o un grupo de alimentos concretos, así como a cualquier producto alimenticio básico con los mismos objetivos ya descritos previamente para la autenticación de alimentos.

3. Aplicaciones de las huellas dactilares cromatográficas en 'identificación' y autenticación de alimentos

Cuando se revisa la bibliografía en busca de los artículos publicados que recogen aplicaciones de la metodología de huellas dactilares cromatográficas en la 'identificación' o autenticación de alimentos es habitual encontrar aplicaciones que declaran utilizar huellas dactilares, pero que en realidad están aplicando la metodología de perfiles de componentes.

Todo ello conlleva que las aplicaciones publicadas, que aplican realmente la metodología de huellas dactilares cromatográficas, sean aún escasas, entre 50 y 60 artículos.

Además, entre ellas, es habitual encontrar artículos que aplican la metodología sobre cromatogramas bien resueltos; y entre estos ejemplos figuran los primeros trabajos publicados por algunos de los autores de este tutorial. Sin embargo, la aplicación en estas circunstancias concretas es equivalente a utilizar el perfil cromatográfico. De hecho, la información útil en un vector de datos que recoge un cromatograma bien resuelto, está localizada en aquellos elementos del vector correspondientes a cada uno de los picos cromatográficos, ya que el resto sólo representan línea base. Además, todos los elementos del vector pertenecientes a un mismo pico, si

éste es característico de un único componente, están fuertemente correlacionados y, en la práctica, todos ellos podrían ser sustituidos por un único elemento, como es el correspondiente al elemento de mayor intensidad (altura de pico) o por el área asociada (área de pico). Esta operación de reducción de datos ya ha sido previamente comentada (véase Figura 3).

El abanico de objetivos de los estudios publicados es amplio, fundamentalmente conectados con la autenticación de alimentos aunque también con la 'identificación' de los mismos. En paralelo, dichos estudios emplean herramientas quimiométricas específicas con tres objetivos principales: (i) análisis de la similitud entre señales para identificación de alimentos; (ii) resolución de señales; (iii) clasificación de alimentos en función de diversos criterios (p.e.: especie botánica o animal, origen geográfico); y (iv) predicción de propiedades relacionadas con la calidad alimentaria como, p.e.: la capacidad antioxidante. En muchos casos se aplica también, para diagnóstico previo de los datos, análisis exploratorio y, en algunos pocos, análisis de agrupamientos. Una información detallada sobre cada uno se podrá encontrar en el artículo matriz al que se hace referencia en la nota de los autores, al principio del texto.

Analizando los diferentes artículos, hay varios aspectos que merecen ser destacados:

- Los tiempos de análisis son excesivos para una aplicación efectiva de la metodología de huellas dactilares ya que únicamente en unos pocos casos el tiempo es inferior a 10 minutos, mientras que hay bastantes estudios con

tiempos de análisis superiores a 1 h. Posiblemente este hecho es debido a que el método se ha optimizado buscando la mejor resolución, como se haría en el caso de la metodología de perfiles.

- Salvo en 2 artículos, uno por cada tipo de cromatografía, que aplican técnicas propias de la minería de datos (SVM, RF, M5), en el resto se emplean técnicas convencionales de la quimiometría clásica.
- No hay publicado ningún antecedente del uso de métodos de fusión de datos aplicado a huellas dactilares generadas por diferentes modalidades de cromatografía.
- Tampoco hay ningún ejemplo de uso de huellas dactilares bidimensionales, por ejemplo por aplicación de cromatografía 2D, que requieren aplicar técnicas de 3-vías para el tratamiento de datos.

Estos aspectos indican que la metodología de huellas dactilares cromatográficas aún no está asentada en la comunidad analítica, y los investigadores se resisten aún a desarrollar métodos y aplicaciones *ad-hoc*, lo cual demuestra que aún hay mucho trabajo por hacer en este tema.

Referencias

1. L.M. Reid, C.P. O'Donnell, G. Downey, Recent technological advances for the determination of food authenticity, *Trends Food Sci. Tech.*, 2006, 17, 344-353.
2. D.M.A.M. Luykx, S.M. van Ruth, An overview of analytical methods for determining the geographical origin of food products, *Food Chem.*, 2008, 107, 897-911.

3. E. Capuano, R. Boerrigter-Eenling, G. van der Veer, S.M. van Ruth, Analytical authentication of organic products: an overview of markers, *J. Sci. Food Agric.*, 2013, 93, 12-28.
4. J. Zhang, X. Zhang, L. Dediu, C. Victor, Review of the current application of fingerprinting allowing detection of food adulteration and fraud in China, *Food Control*, 2011, 22 , 1126-1135.
5. D.I. Ellis, V.L. Brewster, W.B. Dunn. J.W. Allwood, A.P. Golovanovc, R. Goodacre, Fingerprinting food: current technologies for the detection of food adulteration and contamination, *Chem. Soc. Rev.*, 2012, 41, 5706-5727.
6. M. Bonnefille, Sensory and Chemical Fingerprinting Aids Quality and Authentication of Ingredients and Raw Materials of Vegetal Origin, in: K.A. Reynertson, K. Mahmood (eds), *Botanicals: Methods and Techniques for Quality & Authenticity*. CRC Press, 2015, cap. 16, pag. 261-278.
7. S. Esslinger, J. Riedl, C. Fahl-Hassek, Potential and limitations of non-targeted fingerprinting for authentication of food in official control, *Food Res. Int.*, 2014, 60, 189-204.
8. J. Riedl, S. Esslinger, C. Fahl-Hassek, Review of validation and reporting of non-targeted fingerprinting approaches for food authentication, *Anal. Chim. Acta*, 2015, 885, 17-32.
9. J.M. Bosque Sendra, L. Cuadros Rodriguez, C. Ruiz Samblas, A.P. de la Mata, Combining chromatography and chemometrics for the characterization and authentication of fats and oils from triacylglycerol compositional data - A review. *Anal. Chim. Acta*, 2012, 724 11-11.
10. K.S. Booksh, B.R. Kowalski, Theory of Analytical Chemistry, *Anal. Chem.*, 1994, 66, 782A-791A.
11. Concepts and applications of the term "dimensionality" in analytical chemistry, IUPAC Technical Report 1993, *Pure Appl. Chem.*, 2002, 74, 1479-1487.

12. T.J. Smillie, I.A. Khan, A comprehensive approach to identifying and authenticating botanical products, *Clin. Pharmacol. Ther.*, 2010, 87, 175-186.
13. W.L. Applequist, J.S. Miller, Selection and authentication of botanical materials for the development of analytical methods, *Anal. Bioanal. Chem.*, 2013, 405, 4419-4428.
14. C. Tistaert, B. Dejaegher, Y. Vander Heyden, Chromatographic separation techniques and data handling methods for herbal fingerprints: A review, *Anal. Chim. Acta*, 2011, 690, 148-161.
15. R.G. Brereton, A short history of chemometrics: a personal view, *J. Chem.*, 2014, 28, 749-760.
16. D.L. Massart, P.K. Hopke, Chemometrics and Distributed Software, *J. Chem. Inf. Comput. Sci.*, 1985, 25, 308-313.
17. D.B. Hibbert, P. Minkkinen, N.M. Faber, B.M. Wise, IUPAC project: A glossary of concepts and terms in chemometrics, *Anal. Chim. Acta*, 2009, 642, 3-5.
18. L. Cuadros Rodríguez, J.M. Bosque Sendra, Mediterranean chemometrics, *Anal. Bioanal. Chem.*, 2011, 399, 1925-1927.
19. R. Tauler, F. Marini, B. Walczak, L. Buydens, R.G. Brereton, European Analytical Column No. 42, *Anal. Bioanal. Chem.*, 2014, 406, 3525-3529.
20. A. Smilde, R. Bro, P. Geladi, Multi-way analysis in chemistry and related fields, John Wiley & Sons, Chichester, 2004.
21. A.C. Duarte, S. Capelo, Application of chemometrics in separation science, *J. Liq. Chrom. Rel. Technol.*, 2006, 29, 1143-1176.
22. M. Daszykowski, B. Walczak, Use and abuse of chemometrics in chromatography, *Trends Anal. Chem.*, 2006, 25, 1081-1096.
23. J.M. Amigo, T. Skov, R. Bro, ChromATHography: solving chromatographic issues with mathematical models and intuitive graphics, *Chem. Rev.* 2010, 110, 4582-4605.

24. J.J. Harynuk, A.P. de la Mata, N.A. Sinkov, Application of chemometrics to the interpretation of analytical separations data. En: K. Varmuza (Ed), *Chemometrics in Practical Applications*, Intech, Rijeka, Croatia, 2012.
25. L.A. Berrueta, R.M. Alonso Salces, K. Héberger, Supervised pattern recognition in food analysis, *J. Chromatogr. A*, 2007, 1158, 196-214.
26. P. Oliveri, G. Downey, Multivariate class modelling for the verification of food-authenticity claims, *Trends Anal. Chem.*, 2012, 35, 74-86.
27. B. Vandeginste, *Chemometrics in studies of food origin*, in: P. Brereton (Ed.) *New Analytical Approaches for Verifying the Origin of Food*, Woodhead, Cambridge, 2013, pag. 117-145.
28. F. Marini (Ed.), *Chemometrics in Food Chemistry*, Elsevier, Amsterdam, 2013.
29. E. Borràs, J. Ferré b, R. Boqué, M. Mestres, L. Aceña, O. Busto, Data fusion methodologies for food and beverage authentication and quality assessment - A review, *Anal. Chim. Acta*, 2015, 891, 1-14.
30. E. Salvatore, M. Bevilacqua, R. Bro, F. Marini, M. Cocchi, Classification methods of multiway arrays as a basic tool for food PDO authentication, in: M. de la Guardia, A. González (eds), *Food Protected Designation of Origin: Methodologies and Applications*, Elsevier, Amsterdam, 2013, pag. 339-382.
31. A. Rinnan, J.M. Amigo, T. Skov, Multiway methods in food science, in: D. Granato, G. Ares (eds), *Mathematical and Statistical Methods in Food Science and Technology*, IFT Press/Wiley Blackwell, Chichester, 2014, pag. 143-174.

2.5. CONCLUSIONES

Teniendo en cuenta la información compilada en este capítulo, se puede afirmar que en los últimos años y cada vez más, se está haciendo uso de las técnicas analíticas con la obtención de "huellas dactilares" instrumentales para la resolución de problemas analíticos. Cada vez son más los analistas que conocen este tipo de señales y que se están familiarizando con el uso y manejo de diferentes herramientas quimiométricas para poder sacarles así el máximo partido.

La principal barrera que nos encontramos en este tema es conseguir que dichas técnicas sean aceptadas por parte de las administraciones u organismos competentes para que este tipo de métodos de agrupación o clasificación de muestras aparezcan como nueva metodología en Reglamentos y normas, mejorando así alguno de los métodos existentes. Por ejemplo, en el caso de su uso para el diagnóstico de enfermedades surgen ciertas reticencias acerca de las implicaciones legales que podrían surgir en el caso en el que los pacientes sean mal diagnosticados. Por otro lado, en el caso de la autenticación de alimentos podría dar soluciones a varios sectores de la industria, pero de nuevo, existen reticencias con la fiabilidad de la metodología y de la posible dificultad que podría entrañar el tratamiento de los datos.

Sin embargo, este tipo de herramientas son muy útiles, como hemos visto en los ejemplos propuestos en los artículos científicos, en el campo del fraude o autenticación de alimentos. Además, cada vez están siendo más usados por la versatilidad de la metodología, sus rápidos tiempos de análisis sin la necesidad de cuantificar ni identificar los compuestos presentes en las señales instrumentales y por los buenos resultados de clasificación de las muestras desconocidas en modelos previamente establecidos, entrenados y validados.

2.6. CONTRIBUCIONES A CONGRESOS

- ❖ L. Cuadros-Rodríguez, A. González-Casado, C. Ruiz-Samblás, E. Pérez-Castaño, L. Valverde-Som, V. Quirós-Rodríguez, F.P. Rodríguez-García. **Aplicación de la metodología de "huellas dactilares" cromatográficas o espectrales (de masas) a la autenticación de aceite de oliva.** XVIII Reunión de la Sociedad Española de Química Analítica (SEQA). **Póster.**

- ❖ S. Medina-Rodríguez, A.M. Jiménez-Carvelo, L. Valverde-Som, E. Pérez-Castaño, C. Ruiz-Samblás, A. González-Casado, L. Cuadros-Rodríguez. **An efficient and user-friendly algorithm for pre-processing of chromatographic data with MATLAB software.** Workshop environmental omics, approaches and technologies for a deeper understanding. **Póster.**

- ❖ L. Valverde-Som, A.M. Jiménez-Carvelo, A. González-Casado, C. Ruiz-Samblás, L. Cuadros-Rodríguez. **Chromatographic fingerprinting for food authentication.** Quality control in the entire food chain. Pre-conference workshop of 9th Euro-Global Summit and Expo on Food and Beverages. **Oral.**

- ❖ L. Valverde-Som, A.M. Jiménez-Carvelo, A. González-Casado, C. Ruiz-Samblás, L. Cuadros-Rodríguez. **Chromatographic fingerprinting for food authentication.** 9th Euro-Global Summit and Expo on Food and Beverages. **Póster.**

CAPÍTULO 3

Monitorización de la estabilidad de patrones de aceite de oliva para análisis sensorial



CAPÍTULO 3

3.1. PRESENTACIÓN

Este capítulo está dedicado a describir diferentes **enfoques para elaborar gráficos de control para monitorizar la estabilidad de patrones de aceite de oliva virgen a partir de datos instrumentales multivariados**.

Las muestras de aceites de oliva empleadas, proceden de un proyecto, titulado: "Certificación y distribución de nueve materiales de referencia certificados (MRC) para análisis organoléptico de aceite de oliva (SensOLEO-MRC 2015)", que ha sido financiado por la Consejería de Agricultura, Pesca y Desarrollo Rural de la Junta de Andalucía y fondos europeos. El proyecto tiene el visto bueno por parte del Consejo Oleícola Internacional (COI) para la certificación de materiales de referencia de aceite de oliva virgen y la armonización de paneles de cata a partir de ellos.

La elaboración de los lotes de los diferentes materiales de referencia (nueve en total), y el envasado de los mismos en muestras unitarias se realizaron en el centro IFAPA (Instituto de Investigación y Formación Agraria y Pesquera) de Venta del Llano en Mengíbar (Jaén). Los materiales de referencia deben ser homogéneos, estables hasta su fecha de caducidad y estar caracterizados para la certificación. Nuestro grupo llevó a cabo los estudios de homogeneidad, estabilidad y caracterización de los materiales SensOLEO-MRC 2015. El estudio de estabilidad forma parte de la investigación que se presenta en este capítulo.

Por tanto, los **objetivos** de la investigación realizada han sido:

- ✓ Optimizar y desarrollar un método analítico para la detección de compuestos orgánicos volátiles en muestras de aceite de oliva.

- ✓ Desarrollar diferentes enfoques que permitan la monitorización de la estabilidad organoléptica de patrones de aceite de oliva mediante gráficos de control multivariados.

3.2. INTRODUCCIÓN

Aceite de oliva virgen

El **aceite de oliva** es un aceite vegetal comestible, que está **compuesto** mayoritariamente (98-99%) **por triglicéridos** (o triacilgliceroles), mientras que el porcentaje restante está formado por una mezcla de **componentes minoritarios** entre los que se pueden destacar algunos **hidrocarburos, ácidos y alcoholes grasos libres, ceras, esteroides, alcoholes triterpénicos, compuestos fenólicos y compuestos volátiles**, entre otros [1]. El Reglamento (CE) nº 1234/2007 [2] por el que se crea una organización común de mercados agrícolas, en su ANEXO XVI establece las *designaciones y definiciones de los aceites de oliva y los aceites de orujo de oliva*, definiendo los **aceites de oliva vírgenes** como:

Aceites obtenidos del fruto del olivo exclusivamente por medios mecánicos u otros procedimientos físicos aplicados en condiciones que excluyan toda alteración del producto, y que no se han sometido a ningún otro tratamiento que no sea su lavado, decantación, centrifugado o filtración, excluidos los aceites obtenidos con el uso de disolventes o de coadyuvantes de acción química o bioquímica, por un procedimiento de reesterificación o como resultado de cualquier mezcla con aceites de otros tipos.

Los aceites de oliva vírgenes solo se clasificarán y designarán de la forma siguiente:

-
- [1] F. Luchetti (2003). "Introducción al estudio del aceite de oliva". En: R. Aparicio, J. Harwood (Eds.), Manual del aceite de olive, AMV Ediciones, pp. 13-31.
- [2] Reglamento (CE) nº 1234/2007 del consejo de 22 de octubre de 2007 por el que se crea una organización común de mercados agrícolas y se establecen disposiciones específicas para determinados productos agrícolas (Reglamento único para las OCM), Diario Oficial de la Unión Europea, L 299, 1-149.

- a) ***Aceite de oliva virgen extra***: *aceite de oliva virgen que tiene una acidez libre máxima, expresada en ácido oleico, de 0.8 g por 100 g y cuyas otras características se ajustan a las establecidas para esta categoría.*
- b) ***Aceite de oliva virgen***: *aceite de oliva virgen que tiene una acidez libre máxima expresada en ácido oleico, de 2 g por 100 g y cuyas otras características se ajustan a las establecidas para esta categoría.*
- c) ***Aceite de oliva lampante***: *aceite de oliva virgen que tiene una acidez libre, expresada en ácido oleico, de más de 2 g por 100 g y/o cuyas otras características se ajustan a las establecidas para esta categoría.*

El resto de características que establecen cada una de las categorías se encuentran recogidas en el ANEXO I acerca de las *características de los aceites de oliva* en el Reglamento (CEE) N° 2568/91, cuya última versión consolidada se ha publicado en diciembre de 2016 [3].

Por tanto, la **clasificación del aceite de oliva en una de sus tres categorías se establece en función de determinados valores, correspondientes a parámetros físico-químicos y atributos sensoriales**, fijados en dicho Reglamento [3].

Análisis sensorial

El **análisis sensorial** es el **examen de propiedades organolépticas** (olor, sabor, textura, etc.) [4]. Al consumir aceite de oliva virgen se perciben una serie de sensaciones debidas a una combinación de **estímulos producidos por los receptores olfativos y gustativos**. Se llama **flavor** a la combinación de las

[3] Reglamento (CEE) N° 2568/91 de la comisión de 11 de julio de 1991 relativo a las características de los aceites de oliva y de los aceites de orujo de oliva y sobre sus métodos de análisis (versión consolidada 01991R2568), Diario Oficial de la Unión Europea (2016), 031.001, 1-128.

[4] J.R. Izquierdo (2008). "Parámetros físico-químicos que definen la calidad del aceite virgen". En: J. Alba, J.R. Izquierdo, F. Gutiérrez, P.Vossen (Eds.), *Aceite de oliva virgen. Análisis sensorial*, segunda edición, Editorial Agrícola Española, pp. 47-77.

sensaciones generadas por el sentido del olfato y del gusto. La descripción del flavor constituye el perfil sensorial de alimentos y bebidas [5].

El aceite de oliva es el único alimento que regula el análisis sensorial para su clasificación en las diferentes categorías y que define la calidad del mismo [6]. El **método sensorial** se denomina "**panel test**" y fue propuesto por el Consejo Oleícola Internacional (COI). El panel test está formado por un grupo de entre 8 y 12 **catadores entrenados**, que funcionan como un análisis instrumental multisensor. Los catadores señalan en la ficha de cata los atributos detectados, cuya intensidad organoléptica aparece representada con una línea horizontal de 10 cm de longitud para cada uno de los atributos positivos y negativos o defectos. Los datos expresados en centímetros se procesan estadísticamente para calcular la mediana de cada atributo dada por el panel [5]. Los atributos positivos se relacionan con el fruto sano, sin enfermedades ni plagas. Las intensidades de los atributos positivos disminuyen y aparecen defectos cuando las etapas del proceso de obtención del aceite de oliva no se realizan correctamente o el fruto del olivo se encuentra en mal estado [7].

El método para la valoración organoléptica de los aceites de oliva vírgenes se recoge y describe en el ANEXO XII de la versión consolidada del Reglamento (CEE) N° 2568/91 [3]. Los atributos que se evalúan en la **ficha de cata** para su clasificación aparecen en el método adoptado por el Consejo Oleícola Internacional (COI/T.20/Doc. No 15/Rev. 8 Noviembre 2015) [8] y corresponden con: (i) **atributos positivos**, como son el frutado, amargo y picante; (ii) **atributos negativos o defectos** que pueden ser atrojado/borras, moho-humedad-tierra, avinado-avinagrado/ácido-agrio, aceituna helada (madera húmeda), rancio, y otros descriptores que aparecen en menor medida y se señalan en el apartado denominado "otros" (por ejemplo, metálico, gusano, esparto, entre otros).

-
- [5] F. Angerosa (2003). "Calidad sensorial de los aceites de olive". En: R. Aparicio, J. Harwood (Eds.), Manual del aceite de olive, AMV Ediciones, pp. 345-380.
- [6] M. Mariotti (2014). "Virgin olive oil: definition and standards". En: C. Peri (Ed.), The extra-virgin olive oil handbook, Wiley, pp. 11-20.
- [7] J.R. Izquierdo (2008). "Las características organolépticas del aceite virgen de oliva". En: J. Alba, J.R. Izquierdo, F. Gutiérrez, P.Vossen (Eds.), Aceite de oliva virgen. Análisis sensorial, segunda edición, Editorial Agrícola Española, pp. 79-111.
- [8] COI/T.20/Doc. No 15/Rev. 8 November 2015, "Sensory analysis of olive oil. Method for the organoleptic assessment of virgin olive oil", pp. 1-20.

Materiales de referencia certificados para aceite de oliva

Los **catadores** que forman el panel han de estar **entrenados**, para que funcionen como un único instrumento y los **paneles de cata, armonizados** para que diferentes paneles den los mismos resultados ante la misma muestra. Así el análisis sensorial no es subjetivo. La única forma de armonizar los paneles de cata sensorial es **mediante el uso de patrones** [9,10], que habitualmente se obtienen de las muestras sobrantes de ensayos interlaboratorios organizadas por el COI u otros organismos que trabajen con dichas muestras, o a partir de materiales de referencia para análisis sensorial.

El uso de materiales de referencia certificados (MRC) de aceite de oliva, permite: (i) obtener material suficiente para que un mismo panel de cata pueda usar estas muestras a lo largo de una campaña; y (ii) armonizar los paneles de diferentes laboratorios. Hoy día no existen en el mercado materiales de referencias certificados de aceite de oliva, y actualmente este hueco es cubierto por el proyecto SensOLEO-MRC.

Un material de referencia es un material suficientemente homogéneo y estable con respecto a propiedades específicas, establecido como apto para su uso previsto en una medida o en un examen de propiedades cualitativas [11].

Un material de referencia certificado es un material de referencia acompañado por la documentación emitida por un organismo autorizado, que proporciona uno o varios valores de propiedades especificadas, con incertidumbres y trazabilidades asociadas, empleando procedimientos válidos [11].

Por tanto, un material de referencia **ha de ser homogéneo, estar caracterizado y ser estable** a lo largo del periodo de uso. Normalmente estos

-
- [9] C. Peri (2014). "Quality excellence in extra virgin olive oils". En: E. Monteleone, S. Langstaff (Eds.), Olive oil sensory science, Wiley, pp. 3-32.
- [10] M. Bertuccioli, E. Monteleone (2014). "The sensory quality of extra-virgin olive oil". En: C. Peri (Ed.), The extra-virgin olive oil handbook, Wiley, pp. 35-58.
- [11] L. Cuadros Rodríguez, L. Gámiz García, A. Carrasco Pancorbo, C. Ruiz Samblás (2013). "Glosario de términos analíticos", primera edición, GRASEQA (Grupo Regional Andaluz de la Sociedad Española de Química Analítica).

tres estudios de homogeneidad, caracterización y estabilidad, se llevan a cabo antes del uso del material. En el certificado del mismo, se indica el tiempo máximo de uso del material, que corresponde con la fecha de caducidad y que garantiza que las propiedades caracterizadas permanecen sin cambios durante ese periodo de tiempo.

En el caso del aceite de oliva la homogeneidad y caracterización se realiza antes del uso del material en los laboratorios, sin embargo, la estabilidad se lleva a cabo a la vez que el material está siendo usado, puesto que es un material vivo que cambia a lo largo de su vida útil. La pérdida de estabilidad podría ser debida al tipo de embotellado usado, así como las condiciones de almacenamiento. Además, podría deberse a un proceso de autooxidación causado por la evolución del aceite de oliva con un contacto prolongado con el aire (atributo rancio).

Los **MRC** utilizados en esta investigación, para el estudio de la estabilidad, se describen en el *apartado 3.3 de muestras*, previo a ello, fueron **caracterizados** y considerados **homogéneos**.

La **caracterización** se realizó en una cata general llevada a cabo por un suprapanel constituido por alrededor de 40 catadores de diferentes paneles de cata de aceite de oliva virgen (paneles oficiales, consejos reguladores, etc.), en septiembre de 2015. Los datos fueron tratados estadísticamente y la valoración de las intensidades organolépticas se llevaron a cabo aplicando el procedimiento establecido en la última versión consolidada, de 04 diciembre 2016, del Reglamento (UE) 2568/91 [3].

El estudio de **homogeneidad** se llevó a cabo en base a los resultados de medidas analíticas instrumentales, usando el mismo método analítico que se usó para el estudio de la estabilidad y que se explica en el *apartado 3.4 de proceso experimental*. Además, en el ANEXO II.1 de este capítulo se presentan los resultados obtenidos de este estudio.

Por último, el estudio de la estabilidad se llevó a cabo durante todo un año, mediante sucesivos análisis químicos instrumentales y a través de gráficos de control multivariables, que se muestran en el ANEXO II.3. Además, al principio y al final de la investigación, se realizó un análisis sensorial por paneles oficiales: Laboratorio Agroalimentario de Granada, Laboratorio Agroalimentario de Córdoba y Laboratorio Arbitral Agroalimentario de Madrid

con el objetivo de corroborar los valores de intensidad de cada uno de los atributos organolépticos certificados.

Compuestos responsables del flavor

Para las medidas instrumentales, en relación a las características sensoriales del aceite de oliva, es importante conocer la relación que existe entre los atributos organolépticos percibidos por las sensaciones generadas en el olfato y el gusto (flavor), y los compuestos responsables de esas sensaciones. El flavor no es sólo resultado de la concentración de un único compuesto, sino que se relaciona con una serie de compuestos que **interaccionan entre sí** [12,13].

Los aldehídos y los alcoholes son los principales compuestos responsables del aroma en el aceite de oliva virgen, así como cetonas, hidrocarburos, ésteres, furanos, ácidos y compuestos aromáticos, que se encuentran en menor concentración [14,15]. El sabor es el resultado de los atributos amargo y picante, que están relacionados con los compuestos responsables de la frescura del aceite de oliva virgen, los compuestos fenólicos. En el aceite de oliva virgen se pueden encontrar: ácidos fenólicos, derivados de la oleuropeína, flavonoides y lignanos [16,17]. Por lo tanto, los **atributos organolépticos** están relacionados con los **compuestos orgánicos volátiles (VOCs)** y los **compuestos fenólicos (PhCs)** los cuales son **responsables del aroma y el sabor**, respectivamente [12,17].

-
- [12] C.M. Kalua, M.S. Allen, Jr.D.R. Bedgood, A.G. Bishop, P.D. Prenzler, K. Robards (2007). "Olive oil volatile compounds, flavor development and quality: a critical review". *Food Chemistry*, 100, 273-286.
 - [13] M.T. Morales, R. Aparicio-Ruiz, R. Aparicio (2013). "Chromatographic methodologies: compounds for olive oil odor issues". En: R. Aparicio, J. Harwood (Eds.), *Handbook of olive oil: analysis and properties*, Springer, pp. 261-309.
 - [14] A. Taticchi, S. Esposto, M. Servili (2014). "The basis of sensory properties of virgin olive oil". En: E. Monteleone, S. Langstaff (Eds.), *Olive oil sensory science*, Wiley, 33-54.
 - [15] G. Luna, M.T. Morales, R. Aparicio (2006). "Characterisation of 39 varietal virgin olive oils by their volatile composition". *Food chemistry*, 98,243-252.
 - [16] N. Kalogeropoulos, A.C. Kaliora (2015). "Effect of fruit maturity on olive oil phenolic composition and antioxidant capacity". En: D. Boskou (Ed.), *Olive and olive oil bioactive constituents*, AOCS, pp. 123-145.
 - [17] M.Z. Tsimidou (2013). "Analytical methodologies: phenolic compounds related to olive oil taste issues". En: R. Aparicio, J. Harwood (Eds.), *Handbook of olive oil: analysis and properties*, Springer, pp. 311-333.

Las **sustancias volátiles** se producen a partir de **ácidos grasos poliinsaturados** a través de **reacciones bioquímicas conocidas como ruta de la lipoxigenasa**. En esta ruta se producen una serie de reacciones enzimáticas durante la molienda de los frutos y el batido de la pasta de aceitunas. A través de estas reacciones se generan una serie de compuestos volátiles durante la extracción del aceite de oliva que son los que le dan el aroma [18].

El **flavor agradable, intrínseco** del aceite de oliva virgen, puede ser alterado por procesos como **lipólisis** y **oxidación**, que dan lugar a notas desagradables de flavor. La **oxidación** se puede dar después de que se haya obtenido el aceite de oliva a causa de un contacto prolongado con el aire o por el almacenamiento prolongado de las aceitunas en montones antes de la extracción del aceite [19].

Además, **el perfil de compuestos orgánicos volátiles también varía en función de la variedad del fruto, su madurez, el sistema de extracción y el origen geográfico**. Existen un gran número de variedades de aceituna, que se caracterizan por su antigüedad, localización y homogeneidad genética. En función del tipo de aceituna escogida se obtiene un aceite con características sensoriales diferentes [20].

Métodos analíticos para el estudio de la estabilidad

Hasta ahora, los métodos analíticos que se usan para el **estudio de la estabilidad** en el aceite de oliva virgen, se basan en parámetros físico-químicos que se relacionan con la oxidación de la muestra.

La **estabilidad oxidativa** de los aceites normalmente se mide con el equipo Rancimat [21,22]. Este equipo se basa en la inducción de la oxidación de las

-
- [18] D.L. García González, M.T. Morales, R. Aparicio (2010). "Olive and olive oil". En: F. Chen, L.M.L. Nollet (Eds.), Handbook of fruit and vegetable flavours, Wiley, pp. 821-847.
- [19] M.T. Morales, R. Prybylski (2003). "Oxidación del aceite de oliva". En: R. Aparicio, J. Harwood (Eds.), Manual del aceite de oliva, AMV Ediciones, pp. 443-473.
- [20] M. Uceda, M.P. Aguilera, A. Jiménez, G. Beltrán (2010). "Variedades de olivo y aceituna. Tipos de aceites". En: A. Fernández Gutiérrez, A. Segura Carretero (Eds.), El aceite de oliva virgen: tesoro de Andalucía, Servicios de Publicaciones de la Fundación Unicaja, pp. 107-139.
- [21] N. Tena, R. Aparicio, D.L. García-González (2017). "Virgin olive oil stability study by mesh cell-FTIR spectroscopy", Talanta, 167, 453-461.

muestras por exposición a elevadas temperaturas y flujo de aire (que conducen al defecto rancio). Permite estimar el tiempo de estabilidad oxidativa, siendo este el momento a partir del cual la muestra ha superado el tiempo en el que permanece estable, por tanto es indicativo de una pérdida de calidad y de vida útil de la muestra. El equipo consta de dos bloques calefactores de aluminio, con calentamiento eléctrico, que puede funcionar a la misma o diferentes temperaturas entre 50 y 200°C.

Además, los métodos de clasificación del aceite de oliva en una de sus tres categorías, que se contemplan en el Reglamento (UE) 2568/91 [3], son métodos de análisis físico-químicos basados en parámetros de oxidación, como son [23,24,25]: (i) contenido de ácidos grasos libres, (ii) índice de peróxido, (iii) absorción ultravioleta de los productos de oxidación primarios y secundarios (K_{232} y K_{270} , respectivamente). Los dos primeros basados en valoraciones volumétricas y la última en la absorción de las longitudes de onda indicadas, que se deben a la presencia de sistemas diénicos y triénicos conjugados que resultan de los procesos de oxidación.

También se ha recurrido al control del contenido en polifenoles y tocoferoles, como indicadores de la exposición de la muestra a la luz [25,26,27], que se puede relacionar con su estabilidad en relación al almacenamiento que se lleve a cabo de la muestra.

-
- [22] S. Gharby, H. Harhar, R. Mamouni, B. Matthäus, E.H. Ait Addi, Z. Charrouf (2016). "Chemical characterization and kinetic parameter determination under Rancimat test conditions of four monovarietal virgin olive oils grown in Morocco", *OCL-Oilseeds and fats, Crops and Lipids*, 23, A401-A408.
- [23] H. Kavuncoglu, T. Dursun Capar, S. Karaman, H. Yalcin (2017). "Oxidative stability of extra virgin olive oil blended with sesame seed oil during storage: an optimization study based on combined design methodology", *Journal of Food Measurement and Characterization*, 11, 173-183.
- [24] I. Javidipour, H. Erinç, A. Bastürk, A. Tekin (2017). "Oxidative changes in hazelnut, olive, soybean, and sunflower oils during microwave heating", *International Journal of Food Properties*, 20, 1582-1592.
- [25] R. Garrido-Delgado, M.M. Dobao-Prieto, L. Arce, J. Aguilar, J.L. Cumplido, M. Valcárcel (2015). "Ion mobility spectrometry versus classical physico-chemical analysis for assessing the shelf life of extra virgin olive oil according to container type and storage conditions", *Journal of Agricultural and Food Chemistry*, 63, 2179-2188.
- [26] S. Esposto, A. Taticchi, S. Urbani, R. Selvaggini, G. Veneziani, I. Di Maio, B. Sordini, M. Servili (2017). "Effect of light exposure on the quality of extra virgin olive oils according to their chemical composition", *Food Chemistry*, 229, 726-733.
- [27] E. Stefanoudaki, M. Williams, J. Harwood (2010). "Changes in virgin olive oil characteristics during different storage conditions", *European Journal of Lipid Science and Technology*, 112, 906-914.

En esta investigación se plantea un método instrumental alternativo para monitorizar la estabilidad de las características organolépticas en muestras de aceite de oliva a partir de la "huella dactilar" de los compuestos orgánicos volátiles.

Métodos de extracción para la determinación de volátiles

Para la determinación de volátiles es necesario extraer, previo a la inyección, los compuestos orgánicos volátiles presentes en la muestra.

A continuación se indican diferentes métodos de extracción de compuestos orgánicos volátiles para cualquier tipo de muestra [12,18]:

- **Inyección directa.** No implica preconcentración de los compuestos orgánicos volátiles. La muestra se coloca en un tubo equipado en el inyector del cromatógrafo de gases, de forma directa, de manera que los volátiles son purgados por un gas portador dentro de la columna. Este método es de baja sensibilidad.
- **Destilación-extracción simultánea.** El procedimiento consta de una destilación de la muestra con éter dietílico, de manera que el vapor de la destilación se condensa en un refrigerante, nitrógeno líquido, en diferentes trampas criogénicas o materiales adsorbentes y después se inyecta en el cromatógrafo de gases. Este método implica una preconcentración de los VOCs.
- **Espacio de cabeza estático (SHS).** No implica preconcentración de los compuestos orgánicos volátiles. La muestra es depositada en un vial sellado, se selecciona una temperatura determinada de manera que la muestra se calienta y tras el equilibrio entre la fase vapor y líquida los VOCs se sitúan en la parte superior del vial en fase vapor. Se toma el gas del espacio de cabeza (que se sitúa sobre el líquido) y se inyecta en el cromatógrafo de gases.
- **Espacio de cabeza dinámico (DHS).** Esta técnica consiste en purgar los volátiles de la muestra sometida a una temperatura determinada, mediante el paso de un gas inerte, helio, a través de una trampa en la que se retienen estos compuestos, de forma que se logra además una preconcentración. Posteriormente, los volátiles se desorben térmicamente de la trampa y se

inyectan en un cromatógrafo de gases. También se denomina "purga y trampa – desorción térmica" (P&T-TD).

- **Extracción con fluidos supercríticos (SFE).** La muestra es colocada en una celda de extracción. Un fluido supercrítico pasa a través de la celda y extrae los compuestos volátiles, preconcentrándolos antes de ser inyectados en el cromatógrafo de gases.
- **Microextracción en fase sólida (SPME).** El dispositivo SPME es una jeringa modificada, compuesto de un soporte de fibra retráctil de 1-2 cm de largo. En el análisis de VOCs la fibra se utiliza para el muestreo de la fase vapor producida mediante el método de espacio de cabeza estático. El recubrimiento de polímero de fibra retráctil, actúa como esponja, concentrando los analitos volátiles por procesos de adsorción. Los compuestos volátiles se desorben en el puerto de inyección del cromatógrafo de gases. SPME integra muestreo, extracción, concentración e introducción de la muestra en el GC en un solo proceso. Al ser un método que se realiza en un solo paso, se evita la pérdida de muestra y la introducción de errores en el análisis.

Instrumentación

Los **equipos analíticos** para la determinación de volátiles se basan en un **cromatógrafo de gases acoplado a un detector**, ya sea de ionización de llama, espectrometría de masas, u olfatometría, entre otros [14,12,28,29].

La **técnica analítica seleccionada** para la determinación de la "huella dactilar" cromatográfica de la fracción de compuestos orgánicos volátiles en muestras de aceite de oliva virgen es la **extracción en fase sólida sobre el espacio de cabeza acoplada un cromatógrafo de gases con un detector de ionización de llama ((HS)SPME-GC-FID)**.

[28] S.M. van Ruth (2011). "Techniques for sampling and identification of volatile compounds contributing to sensory perception". En: L.M.L. Nollet, F. Toldrá (Eds), Sensory analysis of foods of animal origin, CRC, pp. 39-48.

[29] M.D.R. Gomes da Silva, A.M. Costa Freitas, M.J.B. Cabrita, R. Garcia (2011). "Olive oil composition: volatile compounds". En: B. Dimitrios (Ed.), Olive oil – Constituents, quality, health properties and bioconversions, InTech, pp.17-46.

El sistema de extracción seleccionado fue la microextracción en fase sólida por las ventajas que presenta [30]. Además se seleccionó la fibra trifásica: 50/30 μm divinilbenceno/carboxen/polidimetilsiloxano (DVB/CAR/PDMS) StableFlex 24 Ga. Existen un gran número de fibras en el mercado: polidimetilsiloxano, poliacrilato, así como mezclas de ellos junto con el divinilbenceno. En función del tipo de polímero o polímeros seleccionados como fibra de extracción se produce un reparto diferente de los analitos y la muestra y se adsorben en la fibra unos compuestos volátiles u otros [31,32]. Las ventajas de la mezcla trifásica DVB/CAR/PDMS estriban en la detección de un mayor número de compuestos orgánicos volátiles en muestras de aceites de oliva [32].

La **hipótesis** inicial de la que se parte es que la **"huella dactilar" cromatográfica obtenida sobre la fracción de los VOCs puede representar de forma global las características organolépticas de un aceite de oliva (virgen o no) sin que sean necesario resolver, identificar y cuantificar los diferentes VOCs que aparecen en el cromatograma**. Por tanto, no es imprescindible el uso de la espectrometría de masas como sistema de detección, identificación y medida, y se optó por la técnica GC-FID. Además la espectrometría de masas suele presentar problemas de reproducibilidad, la cual es imprescindible en estudios de estabilidad.

Por otro lado, teniendo en cuenta la premisa planteada como hipótesis, no es necesario obtener cromatogramas bien resueltos ni con largos tiempos de análisis. Por lo tanto, se partió del método propuesto en bibliografía por Vichi *et al.* [32] y se optimizó hasta reducir el **tiempo de análisis a 10 minutos**.

Un esquema del **detector FID** se muestra en la figura 3.2. El detector de ionización de llama consiste en un quemador de hidrógeno/oxígeno. Las líneas

-
- [30] J.F. Cavalli, X. Fernandez, L. Lizzani-Cuvelier, A.M. Loiseau (2003). "Comparison of static headspace, headspace solid phase microextraction, headspace sorptive extraction, and direct thermal desorption techniques on chemical composition of French olive oils", *Journal of Agricultural and Food Chemistry*, 51, 7709-7716.
- [31] R. Cela, R.A. Lorenzo, M.C. Casais (2002). "Técnicas de separación en química analítica", Editorial Síntesis, pp.519-555.
- [32] S. Vichi, A.I. Castellote, L. Pizzale, L.S. Conte, S. Buxaderas, E. López-Tamames (2003). "Analysis of virgin olive oil volatile compounds by headspace solid-phase microextraction coupled to gas chromatography with mass spectrometric and flame ionization detector" *Journal of Chromatography A*, 989, 19-33.

de gas que entran en el detector corresponden a hidrógeno y aire, para formar una llama mediante una chispa proporcionada por el mechero. El gas portador introduce los analitos a través de la columna en el sistema de detección. A la salida de la columna, se mezclan con el hidrógeno y llegan a la parte del mechero (figura 3.1) donde se enciende una chispa eléctrica, produciéndose una llama de alta temperatura (el aire se usa para mantener la combustión).

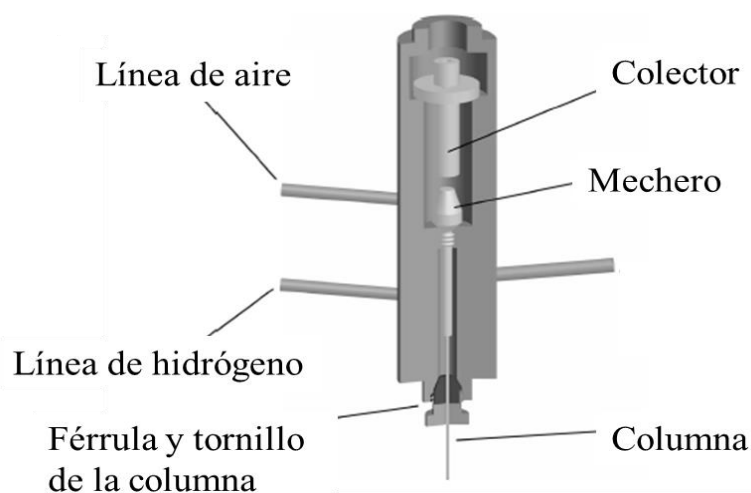


Figura 3.1. Componentes principales de un FID (adaptada [33]).

La mayoría de los compuestos orgánicos se pirolizan en el seno de la llama, y se producen iones y electrones que son conductores eléctricos. Las especies iónicas generadas en la llama pasan al colector donde se produce una pequeña corriente, que es la que se mide como señal. El proceso de ionización que se da en la llama es complejo, pero se puede aproximar el número de iones producidos al número de átomos de carbono transformados en la llama. Esto produce que sea un detector sensible a la masa (al número de átomos de carbono que salen de la columna) más que a la concentración, por lo tanto no le afectan demasiado los cambios en el flujo de salida [33,34].

- [33] D. Rood (2007). "The troubleshooting and maintenance guide for gas chromatographers", Wiley-VCH, Cap. 8, pp. 139-185.
- [34] I.A. Fowlis (2007). "Gas chromatography", segunda edición, ACOI, analytical chemistry by open learning, Cap. 6, pp. 136-157.

3.3. MUESTRAS

Las muestras son MRCs de aceite de oliva procedentes del proyecto SensOLEO-MRC 2015. Se dispuso de un total de nueve materiales de referencias certificados, de diferentes variedades de aceituna, envasados en botellas de PET de calidad alimentaria con tapón de material inerte dotado de precinto (ver Figura 3.2).



Figura 3.2. Materiales de referencia certificados de aceite de oliva.

En la tabla 3.1 se indica la descripción de cada uno de los materiales con su código, la composición en función de la variedad de aceituna (y su proporción en el caso de mezclas), así como una descripción del mismo.

En la tabla 3.2 se muestra el atributo certificado para cada uno de los materiales descritos en la tabla 3.1, así como la intensidad organoléptica y su incertidumbre asociada ($k=2$) en centímetros (unidades de medida de la ficha de cata).

Tabla 3.1. Descripción de los materiales de referencia certificados de aceite de oliva.

Código del material (número de lote)	Variedad (proporción, %)	Descripción
FRU/2/ (010-390)	Hojiblanca (100)	Aceite de oliva obtenido de aceitunas de vuelo maduras, con intensidad media de frutado
FRU/1/ (010-390)	Hojiblanca (100)	Aceite de oliva obtenido de aceitunas de vuelo maduras, con intensidad baja de frutado
BRA/3/ (010-390)	Picual (100)	Aceite de oliva con una intensidad alta del defecto atrojado/borras y media del defecto avinado, que ha sufrido un proceso extensivo de fermentación
BRA/2/ (010-390)	Hojiblanca (88), Picual (12)	Aceite de oliva con una intensidad alta/media del defecto atrojado/borras, obtenido por dilución del material BRA/3/ con el material FRU/1/, en una proporción 12/88
BRA/1/ (010-390)	Hojiblanca (97.5), Picual (2.5)	Aceite de oliva con una intensidad alta/media del defecto atrojado/borras, obtenido por dilución del material BRA/3/ con el material FRU/1/, en una proporción 2.5/97.5
MAH/2/ (010-390)	Picual (100)	Aceite de oliva con defecto aceituna helada obtenido de aceitunas que han sufrido una helada
MAH/1/ (010-390)	Hojiblanca (60), Picual (40)	Aceite de oliva, con una intensidad alta/media del defecto aceituna helada, obtenido por dilución del material MAH/2/ con el material FRU/1/, en una proporción 40/60
COC/2/ (010-390)	Hojiblanca (100)	Aceite de oliva que ha sufrido un excesivo calentamiento durante la obtención del aceite, durante el termo-batido de la pasta, y que incorpora el defecto cocido o quemado. Complementariamente, contiene también el defecto de avinado/avinagrado
HIG/2/ (010-390)	Picual (100)	Aceite de oliva con frutado maduro persistente que se asigna a defecto tierra

Tabla 3.2. Valores de intensidad y atributo certificado para cada uno de los materiales de referencia certificados de aceite de oliva.

Código del material (número de lote)	Atributo (s) certificado (s)	Intensidad (es) organoléptica (s) certificada (s) (cm) ± incertidumbre asociada (cm), septiembre 2015
FRU/1/ (010-390)	Frutado	3.0 ± 0.4
FRU/2/ (010-390)	Frutado	5.7 ± 0.4
BRA/1/ (010-390)	Atrojado/borras	4.3 ± 0.4
	<i>Avinado/avinagrado</i>	<i>(2.0)*</i>
BRA/2/ (010-390)	Atrojado/borras	5.8 ± 0.4
	<i>Avinado/avinagrado</i>	<i>(3.0)*</i>
BRA/3/ (010-390)	Atrojado/borras	8.2 ± 0.4
	<i>Avinado/avinagrado</i>	<i>(4.0)*</i>
MAH/1/ (010-390)	Aceituna helada (madera húmeda)	2.3 ± 0.4
MAH/2/ (010-390)	Aceituna helada (madera húmeda)	3.9 ± 0.4
COC/2/ (010-390)	Avinado/avinagrado	2.8 ± 0.6
	Cocido o quemado	3.3 ± 0.6
HIG/2/ (010-390)	Frutado (higuera sucia)	3.4 ± 0.4
	Moho/humedad/tierra	2.4 ± 0.4

* NOTA: En cursiva se indican los atributos y las intensidades organolépticas de los materiales en los que los catadores que han participado en la certificación no muestran un resultado confiable debido a: (i) una alta dispersión de los datos (coeficiente de variación global superior al 20%); y (ii) un bajo número de datos válidos (inferior al 75% del total). Por este motivo, a estos no se le incluyen la incertidumbre.

3.4. PROCESO EXPERIMENTAL

Tradicionalmente, la forma de llevar a cabo un estudio de estabilidad de las características organolépticas en muestras de aceite de oliva, sería **aplicando**

directamente el método de análisis sensorial "panel test". El problema aparece en que las mismas muestras y atributos certificados que se utilizan para el entrenamiento y la armonización de los paneles de cata han de ser evaluados por los propios catadores de un panel. Esto se puede resolver aplicando un método analítico de determinación de volátiles que permita el control instrumental de los atributos certificados en dichas muestras y su monitorización mediante gráficos de control multivariables.

El estudio de estabilidad abarcó un periodo de 12 meses contados a partir de la fecha de caracterización del material (septiembre 2015). Este estudio se realizó sobre muestras unitarias previamente seleccionadas como representativas de cada lote de material, y una comparación estadística con el cromatograma de referencia obtenido a partir del mismo material recién envasado, metodología que se describe a continuación.

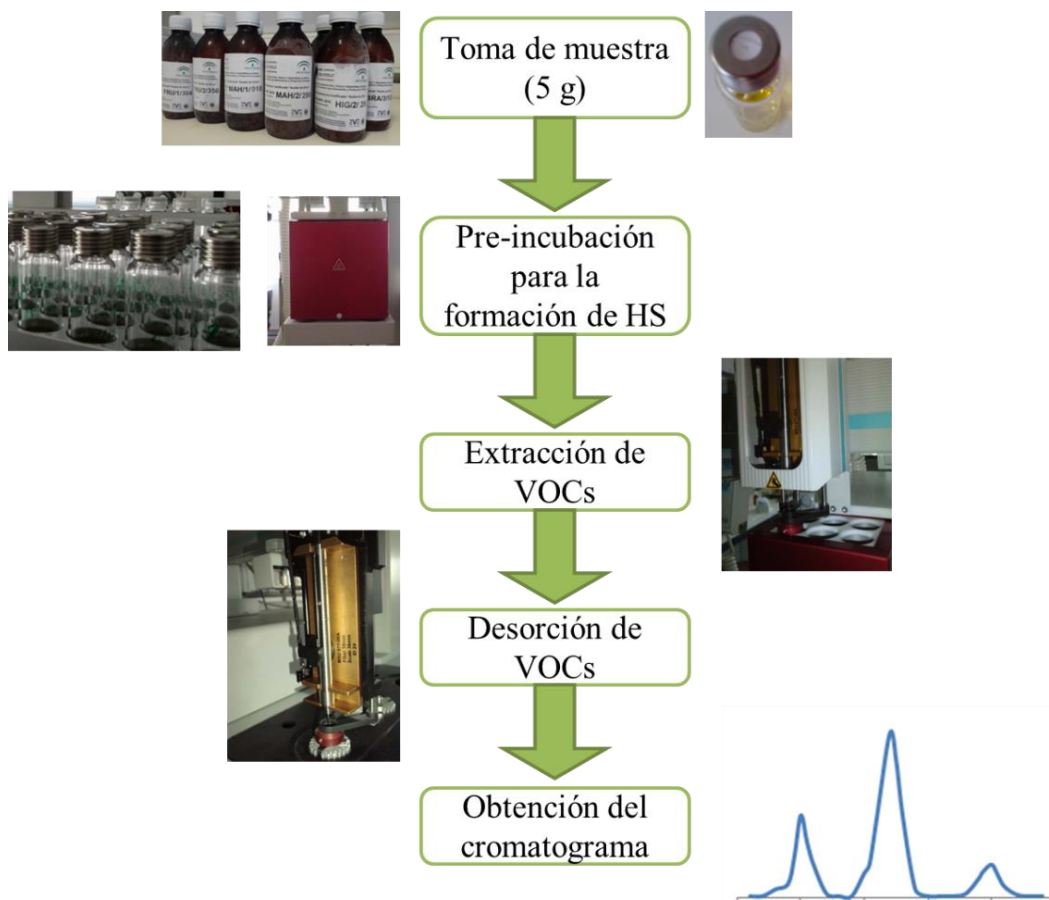


Figura 3.3. Esquema del proceso analítico hasta la etapa de detección de la muestra.

Durante el proceso experimental se utiliza un patrón de control externo, el fluorobenceno, con el objetivo de controlar el equipo cromatográfico. Este patrón es mezclado con aceite de oliva refinado, con una concentración de 100 mg/g y se colocan 100 μ L de dicha disolución en viales de 20 mL con tapón de rosca magnético para su inyección.

Las muestras se preparan pesando 5 gramos en viales de 20 mL con tapón de rosca magnético y se inyectan en el equipo extrayendo la fracción volátil con la fibra SPME, tras la formación del espacio de cabeza por calor. Posteriormente se inyecta en el cromatógrafo de gases y son desorbidos los compuestos en el inyector, separados por el cromatógrafo y detectados en el FID.

Un esquema de este proceso se muestra en la figura 3.3.

En las tablas 3.3 y 3.4 se indican las condiciones de la extracción de los compuestos orgánicos volátiles y las condiciones del cromatógrafo de gases y del detector fijadas en el estudio.

Tabla 3.3. Condiciones del espacio de cabeza estático acoplado a la fibra de microextracción en fase sólida y condiciones del sistema de inyección.

Pre-incubación (horno, formación HS)	Temperatura (°C)	Tiempo (min)	Agitación (rpm)
	40	2	250
Fibra SPME	50/30 μ m DVB/CAR/PDMS, StableFlex 24 Ga (divinilbenceno/carboxen/polidimetilsiloxano)		
Extracción VOCs con SPME	Tiempo (min)		Agitación (rpm)
	30		500
Desorción VOCs en el inyector	Modo	Temperatura (°C)	Tiempo (min)
	Splitless	250	2
Limpieza fibra SPME en el inyector	Split (1:100)	250	6

Tabla 3.4. Condiciones del cromatógrafo de gases y del detector de ionización de llama.

Gas portador	Helio a 1.5 mL/min		
Columna cromatográfica	Restek Rtx-65TG (30 m × 0.25 mm i.d. × 0.10 μm), 65% difenil-35% dimetil polisiloxano		
Rampa de temperatura GC	Temperatura (°C)	Rampa (°C/min)	Tiempo* (min)
	80	0.00	2.50
	240	40.00	3.50
	Tiempo de análisis		10.00
Detector FID	Gas	Flujo (mL/min)	
	Hidrógeno (99.999%)	30	
	Aire líquido (99.999%)	300	

* Tiempo que se mantiene la temperatura final a la que llega el gradiente.

Pre-tratamiento de los datos

El objetivo de la investigación se basa en comprobar que los atributos organolépticos certificados se mantienen estables en el tiempo. Para ello, se comparan dos cromatogramas entre sí: la señal obtenida al tiempo de referencia o tiempo cero (cuando se recibieron las muestras), y la señal obtenida al tiempo de control (un tiempo determinado "x"); en este caso, el estudio se hace cada tres meses. Si ambos cromatogramas se mantienen iguales, o con pequeñas variaciones en el tiempo que pudieran ser achacables a efectos de tipo aleatorio durante la medida, implica que la composición de compuestos volátiles no se ha modificado. Como consecuencia, este resultado se puede extrapolar a que sus características sensoriales también han permanecido inalterables, y se podrá concluir que el material es estable.

Una vez obtenidos los cromatogramas o vectores de datos, se guardan en formato CSV y son pre-procesados con una función programada para el entorno MATLAB® (versión 8.2), llamada "Medina" (versión 14) [35].

Para poder comparar ambas señales cromatográficas es necesario hacer un pre-tratamiento de los datos. Los pasos que se llevan a cabo son los siguientes:

1) Agrupamiento de las réplicas de un mismo material.

Los cromatogramas obtenidos de un mismo material se agrupan, de manera que se genera una matriz de datos con todas esas réplicas. En el caso del cromatograma a tiempo cero se ponen un total de 10 réplicas, con el objetivo de obtener un cromatograma representativo. Cuando se hace el control a un tiempo determinado, se realizan 2 réplicas de cada material.

2) Eliminación de la parte final del cromatograma, que no tiene información cromatográfica sobre compuestos orgánicos volátiles.

El análisis cromatográfico dura 10 minutos, pero a partir del minuto 6, aproximadamente, no aparecen más picos responsables del perfil de los compuestos orgánicos volátiles. Así, se reduce la matriz de datos a un total de 2800 variables.

3) Eliminación de ruido.

La eliminación de ruido, se conoce como suavizado de la señal y se realiza de forma automática mediante uno de los comandos implementados en la función "Medina".

4) Corrección de la línea base.

Permite corregir la posible deriva en la señal cromatográfica, para los posibles desplazamientos o desviaciones que pudieran darse en la línea base entre las señales de las distintas muestras o entre réplicas de una misma muestra.

[35] E. Pérez-Castaño, C. Ruiz-Samblás, S. Medina-Rodríguez, V. Quirós-Rodríguez, A.M. Jiménez-Carvelo, L. Valverde-Som, A. González-Casado, L. Cuadros-Rodríguez (2015). "Comparison of different analytical classification scenarios: application for the geographical origin of edible palm oil by sterolic (NP)HPLC fingerprinting" *Analytical Methods*, 7, 4192-4201.

5) Alineación de los picos cromatográficos.

El alineamiento de los picos cromatográficos se hace respecto al tiempo de retención de los mismos, por intervalos de tiempo, para corregir el pequeño desplazamiento que se producen entre inyecciones diferentes.

Después de este pre-tratamiento de los datos, se genera una **señal promedio** de las réplicas obtenidas al tiempo de referencia y al tiempo de control. Estas señales promedio son las que se comparan mediante los correspondientes índices de similitud, que se representan en un gráfico de control para la monitorización de la estabilidad.

Gráficos de control

Los **gráficos de control** permiten **analizar el comportamiento de diferentes procesos y detectar tendencias en los datos** para prever posibles fallos mediante métodos estadísticos. Existen **dos tipos de gráficos** en función del número de variables con el que se trabaje. Así, los **gráficos de control univariable** se usan habitualmente para controlar una única característica de calidad en función del tiempo; y los **gráficos de control multivariable** permiten controlar simultáneamente más de una característica o variables, que se condensan en un único parámetro. Esta última herramienta se usa principalmente para monitorizar y como diagnóstico de control de procesos industriales [36].

Los gráficos de control incluyen un **límite de aceptación** a partir del cual es posible decidir si un proceso está o no fuera de control. Existen dos formas de establecer este límite: (i) mediante estadística a partir de datos repetidos en los datos obtenidos para la muestra de referencia; o (ii) por una decisión determinista, que se basa en la experiencia del profesional.

Así, los gráficos de control pueden aplicarse para estudiar la estabilidad de las características organolépticas de aceite de oliva virgen.

[36] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. de Jong (1998). "Handbook of chemometrics and qualimetrics: Part A. Data handling in science and technology", vol. 20, parte A, Elsevier, cap.20, pp. 587-641.

Los **gráficos de control univariantes** se pueden aplicar a las **muestras analizadas mediante el panel de cata**. La intensidad organoléptica del atributo certificado es la característica a controlar. Existen dos formas de representar este tipo de gráficos: (i) o bien directamente el valor de la intensidad de cada atributo positivo o defecto; (ii) o el resultado de restar el valor de la intensidad organoléptica en el tiempo de control de la intensidad de referencia, a tiempo cero. En ambos casos, los valores se representan frente al tiempo durante el que se controla la estabilidad del material.

Los límites de aceptación estadísticos se pueden calcular como tres veces la desviación estándar del método de medición [37]. Y también, para este caso particular, se puede considerar un valor de 0.7 cm, que es representativo de la reproducibilidad del método "panel test" como límite de control determinista. Ese valor ha sido fijado tras analizar un gran número de resultados de diferentes pruebas de análisis sensorial en aceite de oliva virgen, mediante un gran número de paneles por análisis interpanel [38].

En esta investigación sólo se disponía de los datos de análisis sensorial al principio y al final del estudio, por lo que no se pudo llevar a cabo el control univariante de los atributos organolépticos certificados en los MRCs. La tabla 3.2 muestra los valores de intensidad organoléptica certificada en septiembre de 2015 y en la tabla 3.5 se indican los valores de intensidad organoléptica obtenidos al final del estudio (noviembre 2016).

Tomando el valor de 0.7 cm como límite de aceptación, todos los materiales de referencia certificados se mantienen estables.

Los **gráficos de control multivariantes** se pueden utilizar para controlar la estabilidad de las **muestras analizadas mediante el método instrumental**. Así, el perfil de compuestos orgánicos volátiles se adquiere para cada material como un vector de datos, donde por medio del tratamiento de la señal pueden compararse a través de diferentes índices de similitud, la "huella dactilar"

[37] ISO 7870-2:2013, Control chart. Part 2: Shewhart control charts, International Organization for Standardization (ISO). Geneva.

[38] T.28/Doc. No 1/Rev. 1 (2017), Guidelines for the accomplishment of the requirements of the norm ISO 17025 by the laboratories of sensory analysis of virgin olive oil, International Olive Council. Madrid.

cromatográfica a tiempo cero y un tiempo "x". En este caso, se representa el índice de similitud frente al tiempo. IUPAC define índice de similitud como *una cantidad que describe la equivalencia de dos objetos caracterizados por datos multivariantes y que puede estar comprendido en el intervalo [0,1], donde el valor 0 representa una diferencia completa y el valor 1 una completa equivalencia [39]*.

Tabla 3.5. Valores de intensidad organoléptica de cada uno de los atributos certificados de los materiales de referencia aceite de oliva pasado un año desde su elaboración.

Código del material (número de lote)	Atributo (s) certificado (s)	Intensidad organoléptica (cm), noviembre 2016
FRU/2/ (010-390)	Frutado	5.2
FRU/1/ (010-390)	Frutado	3.7
BRA/3/ (010-390)	Atrojado/borras	7.8
	<i>Avinado/avinagrado</i>	<i>0.0*</i>
BRA/2/ (010-390)	Atrojado/borras	5.4
	<i>Avinado/avinagrado</i>	<i>0.0*</i>
BRA/1/ (010-390)	Atrojado/borras	4.4
	<i>Avinado/avinagrado</i>	<i>0.0*</i>
MAH/2/ (010-390)	Aceituna helada (madera húmeda)	3.6
MAH/1/ (010-390)	Aceituna helada (madera húmeda)	2.6
COC/2/ (010-390)	Avinado/avinagrado	2.2
	Cocido o quemado	2.6
HIG/2/ (010-390)	Frutado (higuera sucia)	3.2
	Moho/humedad/tierra	1.7

*NOTA: En cursiva se indican los atributos y las intensidades organolépticas de los materiales en los que los catadores que han participado en la certificación no muestran un resultado confiable. En esta segunda cata, ese atributo no fue detectado por los catadores.

[39] D.B. Hibbert (2016). "Vocabulary of concepts and terms in chemometrics (IUPAC Recommendations 2016)", Pure and Applied Chemistry, 88, 407-443.

En esta investigación se recopilaron diferentes índices de similitud que se dividen en **dos tipos**: (1) **análisis de la proximidad**, basado en las distancias entre los vectores; y (2) **otros análisis de similitud**, que corresponden a estrategias alternativas basadas en la correlación entre los elementos de los vectores o en la orientación espacial de los mismos.

A continuación se explican cada uno de los índices de similitud con un ejemplo sencillo de vectores (ver tabla 3.6), para su cálculo no es necesario tener en cuenta el vector de intensidades. Los vectores de datos del ejemplo, se denominan: A_0 y A_1 , y podrían corresponder con los cromatogramas obtenidos a un tiempo cero y a un tiempo de control "1", respectivamente en el estudio de estabilidad. Cada uno de ellos aparece replicado, $A_{0,1}$, $A_{0,2}$ y $A_{1,1}$, $A_{1,1}$, donde el segundo subíndice es indicativo de la réplica considerada. Además, para cada pareja se obtiene su vector medio: \bar{A}_0 y \bar{A}_1 ; así como el valor medio (\bar{a}) de cada vector de datos $A_{0,1}$, $A_{0,2}$ y $A_{1,1}$, $A_{1,1}$.

Tabla 3.6. Valores numéricos de las réplicas de los vectores de datos: A_0 y A_1 , junto con sus correspondientes valores medios.

Tiempo (min)	$A_{0,1}$, intensidad (mV)	$A_{0,2}$, intensidad (mV)	\bar{A}_0 , intensidad (mV)	$A_{1,1}$, intensidad (mV)	$A_{1,2}$, intensidad (mV)	\bar{A}_1 , intensidad (mV)
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	2	2	2	3	3	3
4	5	5	5	9	6	7.5
4.5	10	10	10	15	15	15
5	19	21	20	30	30	30
5.5	14	14	14	20	20	20
6	8	8	8	12	12	12
7	1	1	1	2	1	1.5
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	5	7	6	9	9	9

Tabla 3.6. (continuación).

Tiempo (min)	$A_{0,1}$, intensidad (mV)	$A_{0,2}$, intensidad (mV)	\bar{A}_0 , intensidad (mV)	$A_{1,1}$, intensidad (mV)	$A_{1,2}$, intensidad (mV)	\bar{A}_1 , intensidad (mV)
10.4	14	14	14	12	12	12
11	25	25	25	38	37	37.5
12	40	40	40	60	60	60
12.5	30	30	30	50	50	50
13	20	20	20	30	30	30
14	2	2	2	3	3	3
15	0	1	0,5	0	0	0
16	0	0	0	0	0	0
17	0	0	0	0	0	0
18	1	1	1	2	1	1.5
19	5	5	5	8	7	7.5
19.3	7	7	7	9,6	9	9.3
20	7	9	8	13	11	12
20.5	6	6	6	8	8	8
21	4	4	4	7	5	6
22	2	2	2	3	3	3
23	0	0	0	0	0	0
24	0	0	0	0	0	0
\bar{a}	7,57	7,80		11,45	11,07	

La figura 3.4 muestra la superposición de los valores medios de las réplicas de ambos cromatogramas.

- Análisis de proximidad (distancias)

Como medida de la similitud entre los cromatogramas que se desean comparar, se calcula la distancia entre ambos vectores de datos, que permite determinar la cercanía que existe entre ellos.

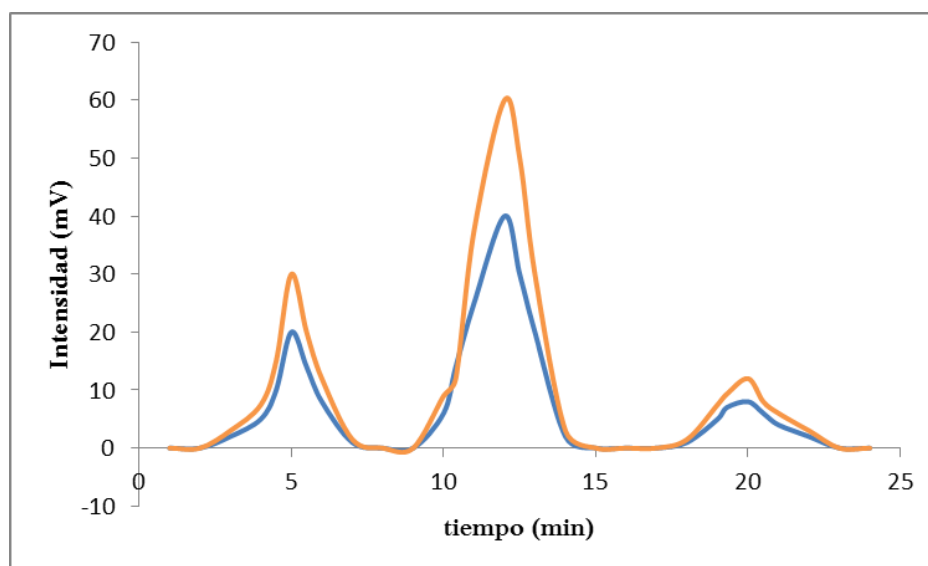


Figura 3.4. Representación gráfica de los vectores de datos \bar{A}_0 (azul) y \bar{A}_1 (naranja).

La **distancia Euclídea (d(eucl))** es la distancia entre los dos vectores A_0 y A_1 [39,40], y se calcula como la raíz cuadrada de la suma de las diferencias cuadrada entre los elementos correspondientes de cada vector. En este caso, se calcula el valor medio \bar{A}_0 y \bar{A}_1 de las réplicas de cada vector y se hace el cálculo de la distancia Euclídea. La notación matricial de este índice es:

$$d(\text{eucl})(A_1, A_0) = \sqrt{(\bar{A}_1 - \bar{A}_0) \times (\bar{A}_1 - \bar{A}_0)^T} \quad (1)$$

Cuando se trabaja con vectores de datos en quimiometría, normalmente cada objeto o muestra se representa por filas, como se ha comentado anteriormente. Por tanto, en este ejemplo, los vectores de datos son de dimensiones (1x30).

Una $d(\text{eucl})$ igual a cero, indica la máxima similitud entre los vectores de datos comparados.

Teniendo en cuenta la definición de índice de similitud de la IUPAC [39], la equivalencia de dos vectores se encuentre comprendida en valores entre 0 y 1, indicando el valor uno la similitud completa. Por lo que es necesario convertir la $d(\text{eucl})$ en un nuevo parámetro normalizado que cumpla con las

[40] R.G. Brereton (2009). "Chemometrics for pattern recognition", Wiley, capítulo 3, pp. 47-106.

recomendaciones de la IUPAC. Este parámetro se denomina **índice de proximidad, NEAR** [41].

$$\text{NEAR}(A_1, A_0) = 1 - \left[\sqrt{\frac{((\bar{A}_1 - \bar{A}_0)) \times (\bar{A}_1 - \bar{A}_0)^T}{(\bar{A}_1 + \bar{A}_0) \times (\bar{A}_1 + \bar{A}_0)^T}} \right] \quad (2)$$

La distancia Euclídea no tiene en cuenta la posible correlación entre las variables, por esto, se recomienda el uso de la **distancia de Mahalanobis (d(mahal))**, en el caso en el que la correlación sea significativa. La distancia de Mahalanobis incluye la matriz varianza-covarianza (Cov), de dimensiones (n×n), que estima la covarianza entre los vectores, que requiere réplicas del vector considerado [36].

$$\text{Cov}(A) = \frac{1}{n-1} \times A_c^T \times A_c \quad (3)$$

Donde A_c es el vector de datos centrado con respecto al vector media ($A - \bar{a}$).

La matriz de varianza-covarianza $\text{Cov}(A_1, A_0)$ [42], se calcula como sigue:

$$\text{Cov}(A_1, A_0) = \frac{(n_1-1) \times \text{Cov}(A_1) + (n_0-1) \times \text{Cov}(A_0)}{(n_1 + n_0 - 2)} \quad (4)$$

En el caso en el que el número de variables en ambos vectores sean iguales, la ecuación pasa a ser:

$$\text{Cov}(A_1, A_0) = \frac{\text{Cov}(A_1) + \text{Cov}(A_0)}{2} \quad (5)$$

Así, la distancia de Mahalanobis se calcula como:

$$d(\text{mahal})(A_1, A_0) = \sqrt{(\bar{A}_1 - \bar{A}_0) \times \text{Cov}(A_1, A_0)^{-1} \times (\bar{A}_1 - \bar{A}_0)^T} \quad (6)$$

Pero, solo se puede calcular la $d(\text{mahal})$ si el número de variables es menor al número de muestras, ya que un número mayor de variables conduce a una Cov cuya matriz inversa da como resultado valores de infinito.

Sin embargo, se puede llevar a cabo un cálculo aproximado de $d(\text{mahal})$, en el caso de tener dos réplicas de cada muestra. En el ejemplo propuesto hay una

[41] R. Pérez-Robles, N. Navas, S. Medina-Rodríguez, L. Cuadros-Rodríguez (2017). "Method for the comparison of complex matrix assisted laser desorption ionization-time of flight mass spectra. Stability of therapeutical monoclonal antibodies", *Chemometrics and Intelligent Laboratory Systems*, 170, 58-67.

[42] R.D. Maesschalck, D. Jouan-Rimbaud, D.L. Massart (2000). "Tutorial the Mahalanobis distance", *Chemometrics and Intelligent Laboratory Systems*, 50, 1-18.

réplica de cada muestra A_0 y A_1 . Para hacer el cálculo de la matriz de varianza-covarianza se calculan los vectores de datos $A_{0,c}$ y $A_{1,c}$ de cada uno de las parejas de vectores: $A_{0,1}$, $A_{0,2}$, y $A_{1,1}$, $A_{1,2}$. A partir de ahí, se obtiene la Cov para A_0 y A_1 y la Cov (A_0, A_1) promedio. Así, se obtiene una matriz de varianza-covarianza de (2x2).

Para poder aplicar la ecuación 6 es necesario calcular previamente el valor medio de cada uno de los vectores (\bar{a}) y generar un nuevo vector de datos (1x2). Por ejemplo, el vector de datos A_0M es un vector de datos (1x2) formado por los valores $\bar{a}_{0,1}$ y $\bar{a}_{0,2}$, para el caso de los datos pertenecientes a cada réplica del vector de datos A_0 . Lo mismo se hace para obtener el vector de datos A_1M . Por lo tanto, la distancia de Mahalanobis pasa a ser:

$$d(\text{mahal})(A_1, A_0) = \sqrt{(A_1M - A_0M) \times \text{Cov}(A_1, A_0)^{-1} \times (A_1M - A_0M)^T} \quad (7)$$

Una $d(\text{mahal})$ igual a cero, indica la máxima similitud entre los vectores de datos comparados. Por tanto, y al igual que para la $d(\text{eucl})$, se puede obtener el valor normalizado. Este parámetro ha sido desarrollado en esta investigación y no se encuentra en bibliografía, lo hemos llamado **índice de manejabilidad, HAND**.

$$\text{HAND}(A_1, A_0) = 1 - \left[\sqrt{\frac{(A_1M - A_0M) \times \text{Cov}(A_1, A_0)^{-1} \times (A_1M - A_0M)^T}{(A_1M + A_0M) \times \text{Cov}(A_1, A_0)^{-1} \times (A_1M + A_0M)^T}} \right] \quad (8)$$

Cuando no es posible calcular la distancia de Mahalanobis porque el número de variables es mayor al número de muestras, se puede recurrir a una reducción de variables previa.

Hotelling propuso una ecuación basada en la $d(\text{mahal})$ bajo el nombre **T² de Hotelling** [43,44], donde se hace una reducción de variables mediante PCA.

$$T^2\text{Hotelling}(A_1, A_0) = n (A_1 - A_0) \times \text{Cov}(A_1, A_0)^{-1} \times (A_1 - A_0)^T \quad (9)$$

El valor de T^2 de Hotelling se obtiene directamente del PCA en el software correspondiente. Por lo general, se normaliza respecto al valor crítico tabulado

[43] H. Hotelling (1947). "Multivariate quality control". En: C. Eisenhart, M.W. Hastay, W.A. Wills (Eds.), Techniques of statistical analysis, Oxford Science publications, pp. 111-184.

[44] A.J. Ferrer-Riquelme (2009). "Statistical control of measures and process. En: S.D. Brown, R. Tauler, B. Walczak (Eds.), Comprehensive chemometrics: chemical and biomedical data analysis, Elsevier, pp. 97-126.

en 0.95 (95% de confianza) para que varíe entre 0 y 1 y se conoce como T^2 de Hotelling reducido.

Para ampliar la información provista por T^2 de Hotelling, se calcula el **Q²-residual**, que se basa en los residuos (r_i) de cada componente principal para el modelo PCs [45]. Al igual que el T^2 de Hotelling, su valor Q²-residual reducido, se obtiene del PCA, en este caso la reducción se hace respecto al valor χ^2 tabulado en 0.95, para cada vector de datos.

$$Q^2 = \sum r_i^2 \quad (10)$$

En ambos casos, el valor cero corresponde con la similitud completa. Para hacer el PCA se genera una matriz de datos que contenga los vectores $A_{0,1}$, $A_{0,2}$, y $A_{1,1}$, $A_{1,2}$, obteniendo así valores de T^2 de Hotelling y Q²-residual reducido, para cada uno de los vectores de datos. Por lo que es necesario hacer el valor medio de para cada uno de esos índices de similitud.

Los valores T^2 y Q² reducidos podrían dibujarse individualmente en gráficos de control. Sin embargo, la desventaja de este tipo de gráficos es que el valor de T^2 de Hotelling no aporta información de los residuales, dando lugar a valores de similitud que no corresponden con la realidad. Por lo que se propone el cálculo de la **media cuadrática (QM)** de los valores reducidos de ambos [46] y su representación en el gráfico de control.

$$QM = \sqrt{\frac{T^2_{red} + Q^2_{red}}{2}} \quad (11)$$

Por cada muestra se obtiene un valor individual de estos índices de similitud. El valor final, representativo del material, es el correspondiente a la media aritmética de todos ellos.

Como alternativa, se pueden calcular los parámetros T^2 de Hotelling y Q²-residual reducidos con otras herramientas quimiométricas como son SIMCA y PLS (descritas en el capítulo 2).

[45] P. Phaladiganon, S.B. Kim, V.C.P. Chen, W. Jiang (2013). "Principal component analysis-based control charts for multivariate nonnormal distributions", *Expert Systems with Applications*, 40, 3044-3054.

[46] M. Bevilacqua, R. Bucci, A.D. Magri, A.L. Magri, R. Nescatelli, F. Marini (2013). "Classification and class-modelling". En: F.Marini (Ed.), *Chemometrics in food chemistry*, Elsevier, pp. 171-233.

- Otros análisis de similitud

Existen otros índices de similitud que también pueden aplicarse a señales procedentes de "huellas dactilares" cromatográficas o instrumentales. Estas son: (i) el **coeficiente de determinación, R^2** , que indica la proporción de la variabilidad total que es común entre los elementos de dos vectores, lo que implica que la diferencia hasta 1 es una medida de la diferencia real entre ambos; está relacionado con el grado de correlación entre ellos [40]; y (ii) el **coseno del ángulo determinado por dos vectores, COS**, que indica la orientación espacial de dichos vectores [40].

Las ecuaciones de cada uno de ellos en forma matricial y para el ejemplo sencillo propuesto, serían:

$$R^2(A_1, A_0) = \frac{(A_{1,c} \times A_{0,c}^T)}{(A_{1,c} \times A_{1,c}^T) \times (A_{0,c} \times A_{0,c}^T)} \quad (12)$$

$$\text{COS}(A_1, A_0) = \frac{\overline{A_1} \times \overline{A_0}^T}{\sqrt{(\overline{A_1} \times \overline{A_1}^T) \times (\overline{A_0} \times \overline{A_0}^T)}} \quad (13)$$

En ambos casos, los valores próximos a la unidad indican la similitud entre los vectores de datos que se están comparando.

Seguidamente, y continuando con el ejemplo antes planteado (tabla 3.6), se calculan cada uno de los índices de similitud (ver tabla 3.7).

Normalmente, los vectores de datos cromatográficos se disponen de la forma $(1 \times n)$, donde el número 1 corresponde con la fila del vector, y representa el vector de datos del cromatograma de la muestra seleccionada, y "n" es el número de columnas que corresponde con las variables de intensidad adquiridas en el cromatógrafo. En el ejemplo, tenemos vectores de datos de dimensiones (1×30) .

A la vista de los resultados obtenidos en la tabla 3.7 concluir que los índices R^2 y COS únicamente controlan la correlación y orientación de los vectores de datos que se están comparando; y el resto de índices propuestos relacionan los

vectores de datos en relación a la intensidad o altura, en el ejemplo con cromatogramas.

Tabla 3.7. Valores de cada uno de los índices de similitud calculados para comparar los vectores de datos A_0 y A_1 .

Índice de similitud	Valor obtenido
Distancia Euclídea, $d(\text{eucl})$	35.9415
Índice de proximidad, NEAR	0.2084
Distancia de Mahalanobis, $d(\text{mahal})^*$	0.9476
Índice de manejabilidad, HAND	0.3595
T^2 Hotelling reducida	0.02632
Q^2 -residual reducida	0.3237
Media cuadrática, QM	0.4154
Coefficiente de determinación, R^2	0.9847
Coseno del ángulo, COS	0.9949

*Cálculo aproximado siguiendo el procedimiento antes descrito

En los apartados 3.5 y 3.6 aparecen los artículos científicos resultado de esta investigación, bajo el título: *Multivariate approaches for chemical chromatographic control of the stability of the organoleptic quality of olive oil reference materials for sensory analysis* y que ha sido dividido en dos partes. En la primera (*framework and fundamentals*) se exponen y se explican las ecuaciones generales para aplicar todos estos índices de similitud; en la segunda (*applications*) se presentan los ejemplos de algunos de los índices de similitud seleccionados para la representación del gráfico de control multivariable de las muestras FRU/1 y BRA/1. Todos los gráficos de control, de todas las muestras y para todos los índices de similitud descritos, se muestran en el ANEXO II.3. En todos estos casos se aplica un límite de aceptación determinista.

Además, en el ANEXO II.2 se explica la forma en la que se seleccionaron diferentes picos cromatográficos característicos de cada MRC con el uso de la herramienta quimiométrica PCA. La selección de dichos picos permite calcular

el índice de proximidad a partir del valor del área de cada pico y obtener así su gráfico correspondiente.

3.5. ARTICULO 1

Multivariate approaches for chemical chromatographic control of the stability of the organoleptic quality of olive oil reference materials for sensory analysis – Part I: Framework and fundamentals

(Artículo enviado a revisión en Noviembre 2017 a la revista: Journal of the Science of Food and Agriculture)



Multivariate approaches for chemical chromatographic control of the stability of the organoleptic quality of olive oil reference materials for sensory analysis – Part I: Framework and fundamentals

Journal:	<i>Journal of the Science of Food and Agriculture</i>
Manuscript ID	JSFA-17-3324
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	13-Nov-2017
Complete List of Authors:	Valverde-Som, Lucía; Universidad de Granada Facultad de Ciencias, Analytical Chemistry Ruiz-Samblás, Cristina; Universidad de Granada Facultad de Ciencias, Analytical Chemistry Rodríguez-García, Francisco P.; Agricultural and Fishery Research Institute, Consejería de Agricultura, Pesca y Desarrollo Rural Cuadros-Rodríguez, Luis ; University of Granada, Department of Analytical Chemistry
Key Words:	Olive oil stability monitoring, organoleptic quality, multivariate control chart, similarity analysis

Abstract

Virgin olive oil is the only food product whose sensory analysis is regulated in order to classify it in three quality categories: extra virgin, virgin and lampante olive oil. Though, to harmonize the results of the sensorial method, is crucial the use of standards or reference materials. The stability of sensory reference materials is required to carry out a suitable control of them in order to confirm that their specific target values are kept on an ongoing basis. Currently, such stability is monitored by mean of sensory analysis and the sensory panels are in the paradoxical situation of controlling the standards which are devoted to control the panels.

In this manuscript, several approaches based on similarity analysis are exploited and different similarity indexes could be applied. For each approach, the specific methodology to build a multivariate control chart to monitor the stability of the sensory properties is explained and discussed.

Keywords

Olive oil stability monitoring, organoleptic quality, multivariate control chart, similarity analysis.

1. Introduction

Virgin olive oil (VOO) is the only food product whose sensory analysis is regulated [1]. The Commission Regulation (EEC) No 2568/91 *on the characteristics of olive oil and olive-residue oil and on the relevant*

methods of analysis includes a section for sensory analysis which describes the organoleptic assessment of virgin olive oil which proposed a sensorial method to classify the produced olive oil in three quality categories: extra virgin olive oil, virgin olive oil, and lampante olive oil, in conjunction with physical and chemical characteristics. This Commission Regulation has been modified over the years (last consolidated version on 04 December 2016) [2]. The 'panel test' is a sensory method proposed by International Olive Council (IOC) to classify the olive oil in function of its quality category, and should be composed between 8 up to 12 tasters. This panel is equivalent to an analytical instrument including a multisensor array, for this, it must be harmonized [3,4].

The only way to harmonize the olive oil sensory panels is by using recognised reference standards [4,5], which could be obtained from the remained samples of the proficiency interlaboratory testing organised by IOC or other provider bodies, or from reference materials (RMs) for sensory analysis, although the use of these last ones are not very extended since they are hard to obtain. Reference material (RM) is *a material, sufficiently homogeneous and stable with respect to one or more specified properties, which has been established to be fit for its intended use in a measurement process* [6].

The stability is *a characteristic of a reference material, when stored under specified conditions, to maintain a specified property value within specified limits for a specified period of time* [6]. The stability is established to prove that the reference values remain unchanged from the production until ready to use. Usually the stability study must be

carried out before to the use of the material. However, as VOO is a 'live' material which changes along its useful life [7], the study of stability should also be carried out at the same time, while the samples are in use. The loss of stability could be produced by bottling and storage conditions. Also, it could be due to a process of auto-oxidation caused by evolution of olive oil or by prolonged contact with air (this is a rancid attribute) [2].

So far the stability study of a VOO was focused on the oxidative stability which is measured by estimating the oxidation induction time in a Rancimat equipment [8,9,10]. Other studies are based on physicochemical analysis through chemical composition. Free fatty acid content, peroxide value and refractive index are the main oxidation parameters. Free fatty acids are responsible for the development of rancidity during storage and its content is assayed by non-aqueous acid-base titration; in the same way, peroxide value is determined by redox titration; and refractive index is measured directly by a refractometer [11,12,13]. Also, the evolution of different indirect parameters as: UV absorption characteristic of primary and secondary oxidation products (K_{232} and K_{270} , respectively); content of polyphenols and tocopherol, as indicator of the light exposure [13,14,15,16]. Other indirect way to analyse the oxidation of VOO is monitoring the volatile compounds which are responsible for rancid aroma [13,14,15]. All these studies are based in oxidation parameters and storage conditions.

The organoleptic attributes are related with volatile organic compounds (VOCs) and phenolic compounds (PhCs) which are responsible of the aroma and taste of the VOO [17,18,19,20]. VOC and PhC profiles are usually

characterised by gas or liquid chromatography, respectively [21,22,23]. The conventional methodology involves the identification and quantification of all these compounds or, alternatively, some of them which are previously selected (markers).

Chromatographic fingerprinting is an innovative approach which enables to monitor the changes in the VOC or PhC profiles without identifying or quantifying the compounds. The fundamentals and main applications of chromatographic fingerprinting for both food 'identification' and food authentication have been recently reviewed [24]. This methodology is being increasingly used because the data information could be present in a part of the chromatogram without being necessarily a chromatographic peak.

In this paper, the first one of a serial of two, the framework and fundamentals for carrying out the monitoring of the stability of the olive oil organoleptic quality, by using chromatographic measures, are explained and discussed. In addition, chemometrics tools are used to establish the multivariate control charts. This is based on the assumption that the aroma of VOO is related to both VOC and PhC profiles, so that if the intensity of the organoleptic attributes remains stable, these profiles will also remain unchanged. Thus, the analytical signal (or chromatographic fingerprint) will be monitored by a specific multivariate control chart. On the contrary, if the intensity of one or more of the organoleptic attributes of the material change, the chromatographic fingerprinting will show a decreasing or increasing systematic trend.

The chromatographic control of the stability of VOO for sensory analysis has not been described in bibliography yet. In addition, this strategy could mean an important innovation for the process control of the olive oil industry.

2. Fundamentals

The stability of the organoleptic attributes in VOO could be controlled by sensory analysis or instrumental chemical analysis. The study of stability by sensory analysis involves the organoleptic assessment of each material by the panel test method. Each attribute positive or negative and their intensities are ranked by the tasters. However, instrumental chemical analysis is based on observing the changes over time of an instrumental signal related to the composition of specific chemical compounds. In both cases, the stability of these materials is stated by the use of appropriate control chart showing the evolution of a quantity related to the intensities of the attributes which are being monitored.

The VOCs, and additionally the PhCs for pungent or bitter attributes, could be selected as targeted compound profiles to acquire and monitor the correspondent chromatographic signals. The proposed methodology assumes that the targeted compounds profiles will not change if the organoleptic attributes of a VOO remain stable. On the same way, any modification on such profiles should indicate a change in the organoleptic characteristics of the material. Consequently, chromatographic methods

showing VOC and, if the case, PhC profiles could be used to detect the lack of stability.

Figure I.1 shows a flow chart describing a general proposal which is suggested for monitoring the organoleptic quality through chromatographic chemical analysis.

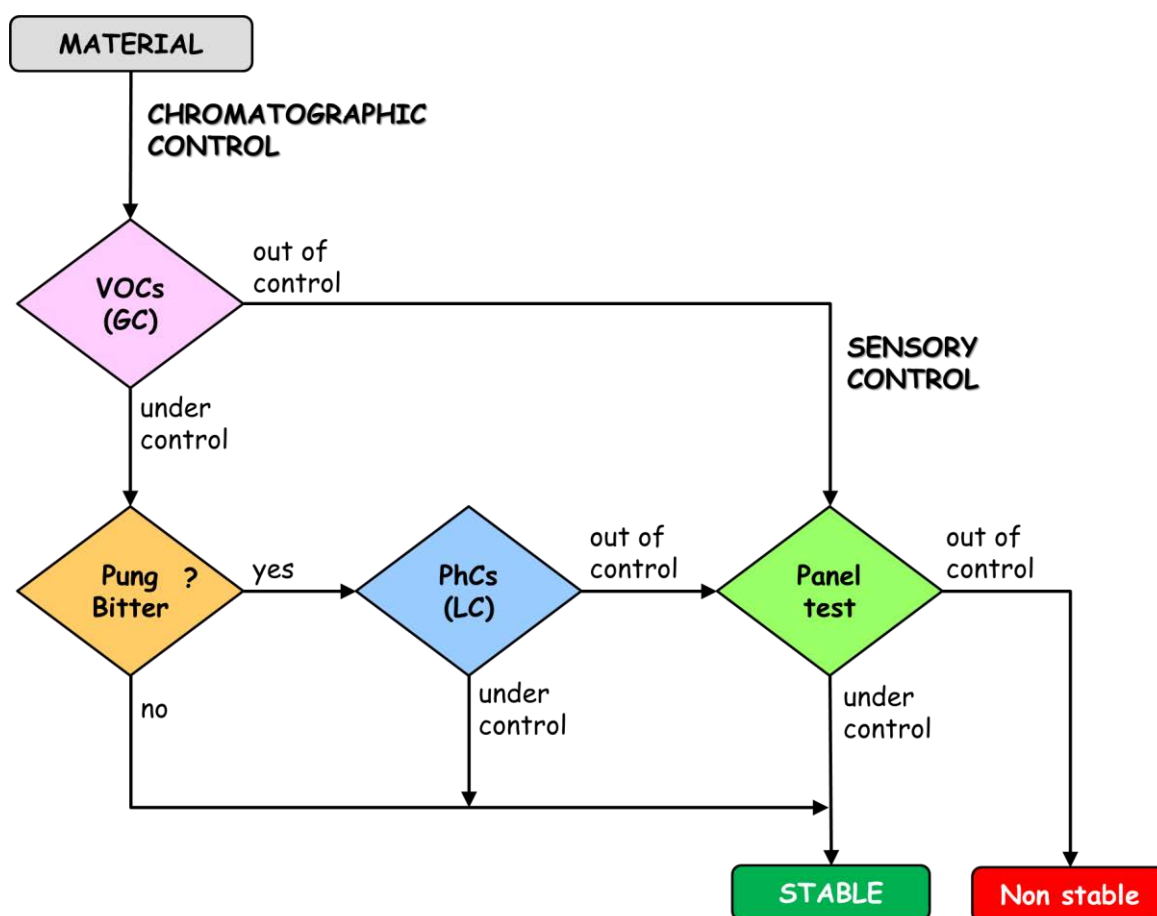


Figure I.1. Flow chart showing the whole methodology to monitor the organoleptic quality through chromatographic chemical analysis. Notice that sensory analysis is used only for confirmatory purpose.

Firstly, the VOC profile from each material is characterised by gas chromatography. This profile is related to all odour/aroma attributes,

positive and negative ones, and provides a comprehensive approach of the organoleptic characteristics. Additionally, and only if the pungent or bitter attributes have to be controlled, the PhC profile are determined in a second step by liquid chromatography. Mostly, it is not necessary to carry out this second analysis because these taste attributes are not very stable and usually they are not used as control characteristic though, in order to ensure that all the information on the whole organoleptic quality is available, the PhC profile could be characterised too.

If the VOC profile and, when appropriated PhC profile, are out of control, the materials should be studied by sensory analysis to confirm the deviation of the control conditions. Finally, the material will be declared 'non stable' only if the outcome from sensory analysis confirms that the intensity of the controlled attribute or attributes have changed.

The chromatographic signals (or chromatograms) can be staked into data vectors which can be handled and compared by means of a suitable mathematical treatment [24]. A comprehensive similarity metrics gathering the result of the comparison would be used as quantity of control to build a multivariate control chart. Univariate control charts are habitually used to monitor a single quality characteristic versus time. Nevertheless, multivariate control charts are required to monitor simultaneously more than one quality characteristic or variable. This graphical tool is largely used to monitor and diagnose industrial process [25]. In addition, these control charts could be also applied to study the organoleptic stability of olive oil using the information collected from successively acquired chromatographic fingerprints. There are diverse

ways of data treatment to create a proper multivariate control chart depending on the control metrics to be graphed.

Firstly, a sufficiently informative instrumental signal should be acquired. Secondly, the signals should be pre-processed in order to remove the unhelpful information and minimize the differences between signals due to non-significant random effects. This previous data treatment involves five steps: (1) grouping data vectors (replicates) from the same material to build a single matrix collecting all chromatographic signals; (2) removing the end of chromatogram which has not information; (3) filtering and removing noise; (4) drift correction and zeroing of the baseline; (5) peaks alignment regarding the retention times. Pre-processing was carried out using MATLAB® version 8.2 (Mathworks Inc., Natick, MA, USA), and a home-programmed MATLAB® function named 'Medina' (version 14) [26]. This function implements several algorithms from MATLAB Bioinformatics Toolbox™, and 'icoshift' (interval correlation optimized shifting) algorithm to align the peaks of the chromatograms [27]. Thirdly, the similarity of chromatographic signals is stated to monitor the stability of targeted profiles, with 'n' variables: (i) at time zero, which is established when samples are received and is symbolised by M_0 (zero or reference); and (ii) at time 'x', which is established at every control time on study and is symbolised by M_x ; an averaged signal is generated for each one of them. As a result a similarity index is calculated for each comparison. In all cases, a similarity index is an output which was defined by IUPAC as *a quantity that describes the equivalence of two objects characterised by multivariate data and may be in the interval [0,1], where 0 is complete*

dissimilarity and 1 is complete equivalence [28]. Fourthly, the multivariate control chart is drawn and the stability is monitored.

Two basic approaches could be applied in order to establish the multivariate control chart: (1) closeness analysis, *i.e.* similarity analysis based on distances; and (2) other similarity analysis, *i.e.* alternative strategies based on correlation between the elements of the signal vectors, or on spatial orientation of such vectors. The flowchart shown in Figure I.2 describes several strategies under those two-approaches of similarity analysis.

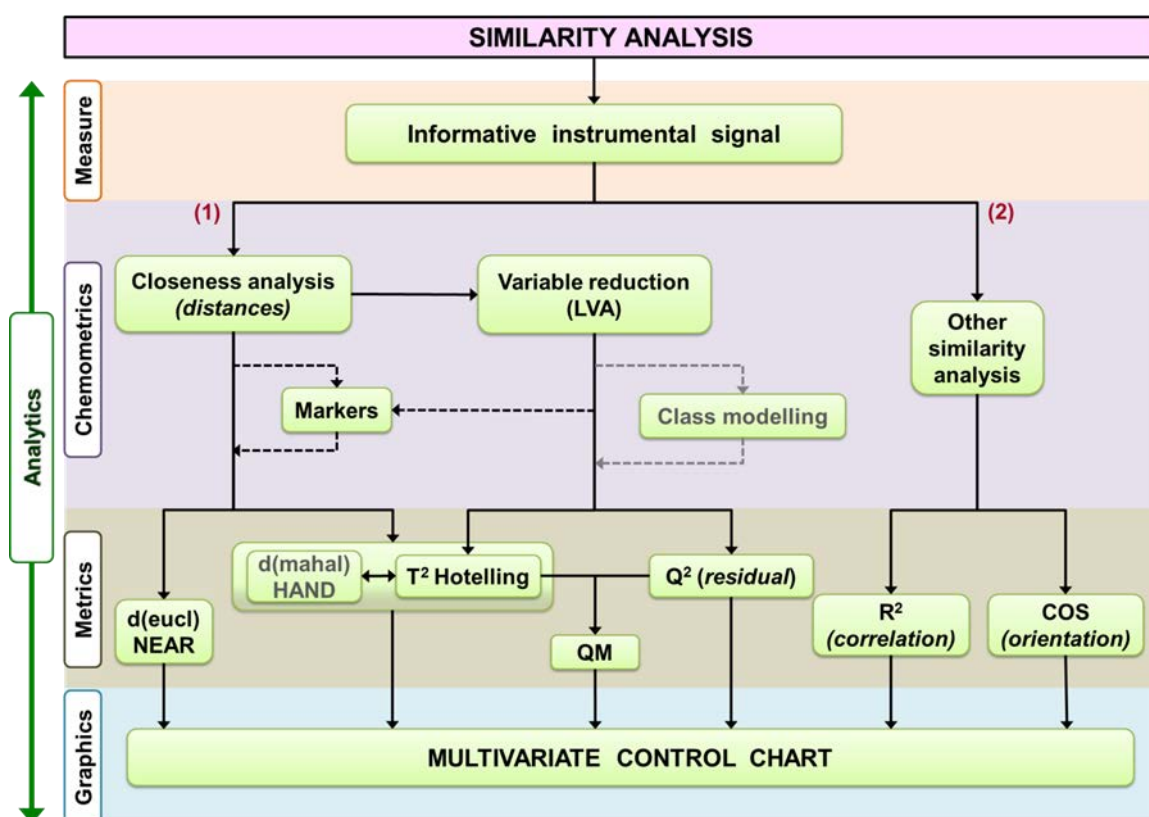


Figure I.2. Flow chart of the analytical framework collection of the multivariate approaches to chromatographic control of the stability of reference materials for organoleptic attributes (LVA: latent variables analysis; QM: quadratic mean).

The following is a detailed explanation of each approach including the basis for calculating the quantity of control to be charted.

Approach 1. Closeness analysis (distances)

The closeness between the chromatograms, expressed as a mathematical distance, was routinely used as a measure of the similarity.

Euclidean distance ($d(\text{eucl})$) is the ordinary distance between two vectors [28,29]. If $M_0 = (M_{01}, M_{02}, \dots, M_{0n})$ and $M_x = (M_{x1}, M_{x2}, \dots, M_{xn})$, where 'n' is the number of elements of the vector, are the two vectors in the Euclidean space, the $d(\text{eucl})$ is calculated as the square root of the sum of the squared differences between the corresponding elements of each vector:

$$\begin{aligned} d(\text{eucl})(M_0, M_x) &= d(\text{eucl})(M_x, M_0) = \\ &= \sqrt{(M_{x,1} - M_{0,1})^2 + (M_{x,2} - M_{0,2})^2 + \dots + (M_{x,n} - M_{0,n})^2} = \sqrt{\sum_{i=1}^n (M_{x,i} - M_{0,i})^2} \quad (1) \end{aligned}$$

Equation 1 could be also re-formulated in matrix way as:

$$d(\text{eucl})(M_x, M_0) = \sqrt{(M_x - M_0) \times (M_x - M_0)^T} \quad (2)$$

The closer the $d(\text{eucl})$ is to zero, the more similar are both vectors. Euclidean distance could be normalized to describe the equivalence of two vectors in the interval [0,1], as it is recommended by IUPAC, thus 1 indicates the complete similarity. This new similarity metrics was named 'nearness index' (NEAR) [30].

$$\text{NEAR}(M_x, M_0) = 1 - \left[\sqrt{\frac{(M_x - M_0) \times (M_x - M_0)^T}{(M_x + M_0) \times (M_x + M_0)^T}} \right] \quad (3)$$

The right application of this matrix equation requires that each one of the data vectors, containing a chromatogram, is written in a single row, *i.e.*, a (1×n) matrix (rows×columns), then the result is a scalar; otherwise the order of the terms of the vector product must be exchanged. The same shall apply to all matrix equations of the rest of the paper.

Euclidean distance ignores the possible correlation between the variables. When this correlation is significant, the use of the Mahalanobis distance (d(mahal)) is recommended.

Mahalanobis distance includes the variance-covariance matrix (Cov) that estimates the covariance between the elements of the vectors [25]. Obviously, the establishment of the variance-covariance matrix, as any measurement based on data variability, requires setting replicates of the vector in consideration.

$$\text{Cov}(M) = \frac{1}{n-1} \times M_c^T \times M_c \quad (4)$$

where M_c is the centred data vector ($M - \bar{M}$). Notice that variance-covariance matrix is always a square (n×n) matrix.

Mahalanobis distance can be of different kinds [31]: (i) distance between a random variable and the centroid (arithmetic mean); (ii) distance between two random variables; and (iii) distance between two vectors; this last one has been considered in this paper.

Thus, in order to calculate the $d(\text{mahal})$ between M_x and M_0 vectors, a pooled variance-covariance matrix denoted as $\text{Cov}(M_x, M_0)$ should be considered [32].

$$\text{Cov}(M_x, M_0) = \frac{(n_x - 1) \times \text{Cov}(M_x) + (n_0 - 1) \times \text{Cov}(M_0)}{(n_x + n_0 - 2)} \quad (5)$$

If the number of variables in both vectors are the same ($n_x = n_0$), Equation 5 could be simplified as:

$$\text{Cov}(M_x, M_0) = \frac{\text{Cov}(M_x) + \text{Cov}(M_0)}{2} \quad (6)$$

Thus, Mahalanobis distance is calculated as it is showed in Equation 7.

$$d(\text{mahal})(M_x, M_0) = \sqrt{(M_x - M_0) \times \text{Cov}(M_x, M_0)^{-1} \times (M_x - M_0)^T} \quad (7)$$

The Mahalanobis distance could be normalized as the Euclidean distance. This new similarity metrics, expressed in Equation 8, could be called 'handiness index' (HAND).

$$\text{HAND}(M_x, M_0) = 1 - \left[\sqrt{\frac{(M_x - M_0) \times \text{Cov}(M_x, M_0)^{-1} \times (M_x - M_0)^T}{(M_x + M_0) \times \text{Cov}(M_x, M_0)^{-1} \times (M_x + M_0)^T}} \right] \quad (8)$$

The $d(\text{mahal})$ can be only calculated if the number of variables is less than the number of objects or samples; if this constraint is not achieved it is necessary to reduce the dimensionality of the vector. A larger number of variables in the data, also called multicollinearity, lead to a Cov which inverse matrix could not be calculated. When the data vectors are chromatographic fingerprints, such as those that are being taken into consideration, they are composed by a large number of variables (several thousand) and it is difficult to process them mathematically. For this reason, a previous variable reduction step should be applied. Thus,

Hotelling proposed a control chart based on $d(\text{mahal})$ with a previous variable reduction [33].

Latent variables analysis (LVA), are used to reduce the dimensionality of the vector, or remove some of the original variables [34]. One example of this is principal component analysis (PCA). PCA is chosen as method to reduce the original data matrix M and the new variables are determined by the number of principal components (PCs) which are linear combinations of the original variables of each object which are broken down into loadings and scores [35]. The characteristic equation of PCA is:

$$M = T \times P^T \quad (9)$$

where T is the score matrix and P is the loading matrix which is transposed in the equation. By definition, PCs are uncorrelated variables; therefore it is no necessary to correct the covariance matrix between the variables [32].

In this LVA point, and according to the flow chart of figure II.2, there are two possibilities: (i) to select markers; or (ii) to reduce the number of variables in the chromatographic fingerprint to calculate other quantities of control.

Chemical markers were defined as particular components which are representative of a quality feature [24]. Markers could be searched on the M_0 vector applying any chemometric method of selection of significant variables. For this purpose, the PC-loadings are particularly an useful tool. PCA is generated from samples which have one or several common attributes. Also specific attributes which difference each sample, are

necessary to generate the PCA. Thus, the PC-loading plot should show the more significant peaks responsible of each specific attribute. The areas of such peaks (one or more) are chosen as characteristic variables.

In this case, the multivariate control charts would be built from the Euclidean distance (or from the nearness index) between the corresponding areas versus time (at each control time).

The second alternative allows to compute with few variables the Mahalanobis distance by means of the Hotelling's T^2 , for this reason this last statistic is often called D^2 -statistic [31]. Hotelling's T^2 measures the variation in each sample within the PCA model [36]. The expression to be applied is [37]:

$$T^2_{\text{Hotelling}}(M_x, M_0) = n (M_x - M_0) \times \text{Cov}(M_x, M_0)^{-1} \times (M_x - M_0)^T \quad (10)$$

Hotelling's T^2 value is directly obtained from PCA. It is usually normalized respect to the critical (tabulated) value of Hotelling's T^2 at 0.95 to range from 0-1; then it is known as reduced Hotelling's T^2 .

In addition, Q^2 -residual is used to extend the information provided by Hotelling's T^2 , since it is based on the residuals of PCs (r_i) to the model.

$$Q^2 = \sum r_i^2 \quad (11)$$

Q^2 -residual defines the overall squared orthogonal distance between the objects or samples and the model, in the PC space [36]. In the same way that Hotelling's T^2 , residual Q^2 value is obtained from PCA. A reduced Q^2 value could be also achieved respect to the critical value of the χ^2 statistic at 0.95 of confidence.

Since the reduced values of both T^2 and Q^2 are squared distances, the corresponding 0 values indicate the complete similarity. Reduced T^2 and Q^2 would be used as control quantities to build multivariate control T^2 -charts and Q^2 -charts, respectively.

Quadratic mean (QM) from the reduced values of both statistics, which is calculated through Equation 12, appears as a new joint control metric [38].

$$QM = \sqrt{\frac{T^2_{red} + Q^2_{red}}{2}} \quad (12)$$

An alternative way to attain both parameters, is applying a one-class classification method or class-modelling strategy, for example, 'soft independent modelling of class analogies' (SIMCA), [39] or 'partial least-square regression' (PLS) [40], which are based on building a PCA model or a PLS model, respectively, using the initial data set from the samples at time zero. After building this model, the samples of each control time are inserted into the model in order to have the respective T^2 - and Q^2 -statistic values.

Approach 2. Other similarity analysis

There are other similarity indexes which could be also applied on the whole instrumental signals (fingerprints). Two data vectors could be too compared evaluating: (i) the coefficient of determination, denoted by R^2 , which indicates the degree of correlation between the elements of two vectors; and (ii) the cosine of the angle determined by two vectors, denoted by COS , which indicates the spatial orientation of such vectors.

The coefficient of determination represents the square of the Pearson correlation coefficient. The main purpose of R^2 is to predict future results or test a hypothesis and therefore obtain the correlation between data.

The matrix equation of R^2 between M_x and M_0 is defined as [29]:

$$R^2(M_x, M_0) = \frac{(M_x - \overline{M_x}) \times (M_0 - \overline{M_0})^T}{((M_x - \overline{M_x}) \times (M_x - \overline{M_x})^T) \times ((M_0 - \overline{M_0}) \times (M_0 - \overline{M_0})^T)} \quad (13)$$

The coefficient of determination ranges from 0 to 1 where the value of 1 indicates the complete similarity.

The cosine of the angle between two vectors is 1 when the angle is 0° . It is a judgment of the orientation: two vectors with the same orientation have a cosine similarity of 1, two vectors at 90° have a similarity index of 0. Cosine similarity is used in positive space. In the case of two chromatograms equals, its corresponding data vectors will be the same and will have the same orientation. COS ranges from 0 to 1 following the directive from IUPAC. The matrix equation is defined as [29,41]:

$$\text{COS}(M_x, M_0) = \frac{M_x \times M_0^T}{\sqrt{(M_x \times M_x^T) \times (M_0 \times M_0^T)}} \quad (14)$$

Both indexes work better if previously each vector is recalculated as cumulative instrumental signal (CIS). CIS is defined as the cumulative sum (cusum) of each data of the vector. Each new element of the vector is added to the previous ones. For instance, the last element in $\text{cusum}(M_0)$ is the sum of all vector of M_0 . Thus, all points on the baseline will disappear. Figure I.3 shows graphically the difference between both vectors (M and $\text{cusum}(M)$).

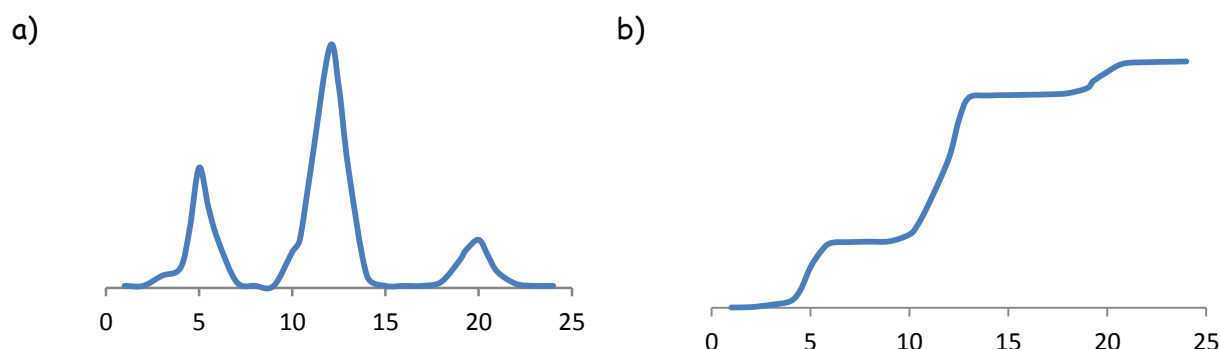


Figure I.3. An example of: (a) an instrumental signal, vector M ; and (b) the cumulative instrumental signal (CIS), vector $\text{cumsum}(M)$.

Both similarity indexes have been widely used to differentiate or authenticate specimens of medicinal plants of traditional Chinese medicine from the respective liquid chromatographic fingerprints. In addition, in this context, another similarity index named 'similarity ratio' was used [42] which is related to the changing patterns of characteristic variables in fingerprints. The 'similarity ratio' (SR) is processed by:

$$\text{SR}(M_x, M_0) = \frac{M_x \times M_0^T}{(M_x \times M_x^T) + (M_0 \times M_0^T) - (M_x \times M_0^T)} \quad (15)$$

Acceptance limits in multivariate control charts

The acceptance limit determines the threshold from which it is possible to decide if a process is or not out of control. There are two ways to establish this limit: (i) by statistics from repeated data vectors of the reference sample; or (ii) by a deterministic decision which is based on the experience of the practitioner.

The statistical threshold describes the control limit (CL) as three times sigma (target value $\pm 3\sigma$), where sigma is the true standard deviation of the considered control metrics. In addition, It is possible to draw a more restrictive 2-sigma limit as 'warning limits' (target value $\pm 2\sigma$) [43]. For instance, a criterion of 6-10 measurements of the reference sample in a very short period of time (a few days) could be established to set down a sufficiently representative estimation of the standard deviation.

These control limits could be only applied when d(eucl), d(mahal), R^2 and COS are used as quantities of control. However, T^2 -statistic, Q^2 -statistic and QM have a specific equation to calculate the CL properly [37,44]. T^2 - and Q^2 -limits depend on the critical values of F and χ^2 , respectively, to calculate each limit according to:

- CL for T^2 -chart:

$$CL(T^2) = \frac{k(m+1)(m-1)}{m^2-mk} F_{(\alpha,k,m-k)} \quad (16)$$

where k is the number of PC, m is the number of samples and F is the one-sided value at 1-a confidence level, and k and m-k degrees of freedom.

- CL for Q^2 -chart:

$$CL(Q^2) = \frac{v}{2b} \chi^2_{(\alpha,2b^2/v)} \quad (17)$$

where b and v are the sample mean and variance of the Q^2 -statistics and χ^2 is the critical value of the chi-square distribution at 1-a confidence level, and $2b^2/v$ degrees of freedom.

- CL for QM-chart:

It is the quadratic mean of both limits.

$$CL(QM) = \sqrt{\frac{CL(T^2) + CL(Q^2)}{2}} \quad (18)$$

Furthermore, the deterministic threshold is based in the data and the experience of the practitioner which knows the characteristic of the samples and the process. It is usually defined as a percentage of the target value, for instance, a 5-10%.

Once the limits are established, it is necessary to define the acceptability criteria to state when the samples are under control. The main criterion is based on the lack of a significant drift. As a general rule, a particular material or product will be stated out of control when at least two correlative points go beyond the stabilised threshold. If this is the case, the stability is doubted and a representative sample should be submitted to sensory analysis in order to confirm the reasons for such a deviation (see Figure I.1).

Abbreviations used

CIS: Cumulative instrumental signal

CL: Control limit

COS: Cosine similarity index

Cov: Variance-covariance matrix

Cusum: Cumulative sum

d(eucl): Euclidean distance

d(mahal): Mahalanobis distance

HAND: Handiness similarity index

IOC: International Olive Council

LVA: Latent variables analysis

NEAR: Nearness similarity index

PCA: Principal component analysis

PC: Principal component

PhC: Phenolic compound

PLS: Partial least-square regression

RM: Reference material

QM: Quadratic mean

SIMCA: Soft independent modelling of class analogies

SR: Similarity ratio

VOC: Volatile organic compound

VOO: Virgin olive oil

Acknowledgments

The authors would like to thank the Andalusia Regional Government (Consejería de Agricultura, Pesca y Desarrollo Rural, Spain) and European funds (ERDF, European Regional Development Fund in Andalusia) for financial assistance with the project SensOLEO-2015. The authors are very grateful to the olive oil samples producer Centro Venta del Llano, Mengíbar, Spain.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- [1] Mariotti, M. (2014). Virgin olive oil: definition and standards. In C. Peri (Ed.), *The extra-virgin olive oil handbook* (pp. 11-20). Chichester: Wiley.
- [2] Commission Regulation (EEC) No 2568/91 on the characteristics of olive oil and olive-residue oil and on the relevant methods of analysis (consolidated version 01991R2568), Official Journal of the European Union (2016), 031.001, 1-128.
- [3] Lerma Garcia, M.J. (2012). *Characterization and authentication of olive and other vegetable oils*. Berlin: Springer, (Chapter 1).
- [4] Peri, C. (2014). Quality excellence in extra virgin olive oils. In E. Monteleone, S. Langstaff (Eds.), *Olive oil sensory science* (pp. 3-32). Chichester: Wiley.
- [5] Bertuccioli, M., Monteleone, E. (2014). The sensory quality of extra-virgin olive oil. In C. Peri (Ed.), *The extra-virgin olive oil handbook* (pp. 35-58). Chichester: Wiley.
- [6] ISO Guide 31:2015 Reference materials - Contents of certificates, labels and accompanying documentation, International Organization for Standardization (ISO). Geneva.
- [7] Kalua, C.M., Allen, M.S., Bedgood Jr, D.R., Bishop, A.G., Prenzler, P.D., Robards, K. (2007). Olive oil volatile compounds, flavour development and quality: A critical review. *Food Chemistry*, 100, 273-286.
- [8] Tena, N., Aparicio, R., García-González, D.L. (2017). Virgin olive oil stability study by mesh cell-FTIR spectroscopy. *Talanta*, 167, 453-461.
- [9] Gharby, S., Harhar, H., Mamouni, R., Matthäus, B., Ait Addi, E.H., Charrouf, Z. (2016). Chemical characterization and kinetic parameter determination under

- Rancimat test conditions of four monovarietal virgin olive oils grown in Morocco. *OCL-Oilseeds and fats, Crops and Lipids*, 23, A401-A408.
- [10] Cuadros-Rodríguez, L., Bosque-Sendra, J.M., Mata-Espinosa, A.P., González-Casado, A., Rodríguez-García, F.P. (2008). Elaboration of four olive oil certified reference materials. InterOleo-CRM 2006 certification study. *Food Analytical Methods*, 1, 259-269.
- [11] Kavuncoglu, H., Dursum Capar, T., Karaman, S., Yalcin, H. (2017). Oxidative stability of extra virgin olive oil blended with sesame seed oil during storage: an optimization study based on combined design methodology. *Journal of Food Measurement and Characterization*, 11, 173-183.
- [12] Javidipour, I., Erinc, H., Baştürk, A., Tekin, A. (2017). Oxidative changes in hazelnut, olive, soybean, and sunflower oils during microwave heating. *International Journal of Food Properties*, 20, 1582-1592.
- [13] Garrido-Delgado, R., Dobao-Prieto, M.M., Arce, L., Aguilar, J., Cumplido, J.L., Valcárcel, M. (2015). Ion mobility spectrometry versus classical physic-chemical analysis for assessing the shelf life of extra virgin olive oil according to container type and storage conditions. *Journal of Agricultural and Food Chemistry*, 63, 2179-2188.
- [14] Esposito, S., Taticchi, A., Urbani, S., Selvaggini, R., Veneziani, G., Di Maio, I., Sordini, B., Servili, M. (2017). Effect of light exposure on the quality of extra virgin olive oils according to their chemical composition. *Food Chemistry*, 229, 726-733.
- [15] Stefanoudaki, E., Williams, M., Harwood, J. (2010) Changes in virgin olive oil characteristics during different storage conditions. *European Journal of Lipid Science and Technology*, 112, 906-914.

- [16] Serrano, L., Cruz, A., Sousa, S., Morais, Z. (2016). Alterations in monovarietal, blended and aromatized Portuguese virgin olive oils under four storage conditions for 12 months. *European Food Research and Technology*, 242, 1041-1055.
- [17] Morales, M.T., Aparicio-Ruiz, R., Aparicio, R. (2013). Chromatographic methodologies: compounds for olive oil odor issues. In R. Aparicio, J. Harwood (Eds.), *Handbook of olive oil: analysis and properties* (pp. 261-309). 2nd ed. New York: Springer.
- [18] Tsimidou, M.Z. (2013). Analytical methodologies: phenolic compounds related to olive oil taste issues. In R. Aparicio, J. Harwood (Eds.), *Handbook of olive oil: analysis and properties* (pp. 311-333). 2nd ed. New York: Springer.
- [19] García-González, D.L., Morales, M.T., Aparicio, R. (2010). Olive and olive oil. In Y.H. Hui (Ed.), *Handbook of fruit and vegetable flavours* (pp. 821-847). Chichester: Wiley.
- [20] Taticchi, A., Esposto, S., Servili, M. (2014). The basis of sensory properties of virgin olive oil. In E. Monteleone, S. Langstaff (Eds.), *Olive oil sensory science* (pp. 33-54). Chichester: Wiley.
- [21] Sales, C., Cervera, M.I., Gil, R., Portolés, T., Pitarch, E., Beltran, J. (2017). Quality classification of Spanish olive oils by untargeted gas chromatography coupled to hybrid quadrupole-time of flight mass spectrometry with atmospheric pressure chemical ionization and metabolomics-based statistical approach. *Food Chemistry*, 216, 365-373.
- [22] Reboredo-Rodríguez, P., González-Barreiro, C., Cancho-Grande, B., Valli, E., Bendini, A., Gallina Toschi, T., Simal-Gandara, J. (2016). Characterization of virgin olive oils produced with autochthonous Galician varieties. *Food Chemistry*, 212, 162-171.
- [23] Zhu, H., Wang, S.C., Shoemaker, C.F. (2016). Volatile constituents in sensory defective virgin olive oils. *Flavour and Fragrance Journal*, 31, 22-30.

- [24] Cuadros-Rodríguez, L., Ruiz-Samblás, C., Valverde-Som, L., Pérez-Castaño, E., González, A. (2016). Chromatographic fingerprinting: An innovative approach for food "identification" and food authentication - A tutorial. *Analytica Chimica Acta*, 909, 9-23.
- [25] Massart, D.L., Vandeginste B.G.M., Buydens L.M.C, de Jong S. (1998). *Handbook of chemometrics and qualimetrics: Part A. Data handling in science and technology volume 20, part A*. Oxford: Elsevier, (Chapter 20).
- [26] Pérez-Castaño, E., Ruiz-Samblás, C., Medina-Rodríguez, S., Quirós-Rodríguez, V., Jiménez-Carvelo, A.M., Valverde-Som, L., González-Casado, A., Cuadros-Rodríguez, L. (2015). Comparison of different analytical classification scenarios: application for the geographical origin of edible palm oil by sterolic (NP) HPLC fingerprinting. *Analytical Methods*, 7, 4192-4201.
- [27] Tomasi, G., Savorani, F., Engelsen, S.B. (2011). IcoSHIFT: An effective tool for the alignment of chromatographic data. *Journal of Chromatography A*, 1218, 7832-7840.
- [28] Hibbert, D.B. (2016). Vocabulary of concepts and terms in chemometrics (IUPAC Recommendations 2016). *Pure and Applied Chemistry*, 88, 407-443.
- [29] Brereton, R.G. (2009). *Chemometrics for pattern recognition*. Chichester: Wiley, (Chapter 3).
- [30] Pérez-Robles, R., Navas, N., Medina-Rodríguez, S., Cuadros-Rodríguez, L. (2017). Method for the comparison of complex matrix assisted laser desorption ionization-time of flight mass spectra. Stability of therapeutical monoclonal antibodies. *Chemometrics and Intelligent Laboratory Systems*, 170, 58-67.
- [31] Mukhopadhyay, A.R. (2008). Multivariate attribute control chart using Mahalanobis D2 statistic. *Journal of Applied Statistics*, 35, 421-429.
- [32] Maesschalck, R.D., Jouan-Rimbaud, D., Massart, D.L. (2000). Tutorial The Mahalanobis distance. *Chemometrics and Intelligent Laboratory Systems*, 50, 1-18.

- [33] Hotelling, H. (1947). Multivariate quality control. In: C. Eisenhart, M.W. Hastay, W.A. Wills (Eds), *Techniques of statistical analysis* (pp. 111-184). New York: Oxford Science Publications.
- [34] Brereton, R.G. (2009). *Chemometrics for pattern recognition*. Chichester: Wiley, (Chapter 6).
- [35] Bro, R., Smilde, A. (2014). Principal component analysis. *Analytical Methods*, 6, 2812-2831.
- [36] Phaladiganon, P., Kim, S.B., Chen, V.C.P., Jiang, W. (2013). Principal component analysis-based control charts for multivariate nonnormal distributions. *Expert Systems with Applications*, 40, 3044-3054.
- [37] Ferrer-Riquelme, A.J. (2009). Statistical control of measures and process. In: S.D. Brown, R. Tauler, B. Walczak (Eds.), *Comprehensive chemometrics: chemical and biochemical data analysis* (pp. 97-126). Oxford: Elsevier.
- [38] Bevilacqua, M., Bucci, R., Magrì, A.D., Magrì, A.L., Nescatelli, R., Marini, F. (2013). Classification and class-modelling. In: F. Marini (Ed.), *Chemometrics in food chemistry* (pp. 171-233). Amsterdam: Elsevier.
- [39] Brereton, R.G. (2011). One-class classifiers. *Journal of Chemometrics*, 25, 225-246.
- [40] Xu, L., Cai, C-B., Deng, D-H. (2011). Multivariate quality control solved by one-class partial least squares regression: identification of adulterated peanut oils by mid-infrared spectroscopy. *Journal of Chemometrics*, 25, 568-574.
- [41] Andrić, F., Héberger, K. (2017) How to compare separation selectivity of high-performance liquid chromatographic columns properly? *Journal of Chromatography A*, 1488, 45-56.
- [42] Cai, B., Ong, S.P., Liu, X. (2012). *High performance liquid chromatography fingerprinting technology of the commonly-used traditional Chinese medicine herbs*, Singapore: World Scientific, (Chapter 1).

- [43] ISO 7870-2:2013, Control chart. Part 2: Shewhart control charts, International Organization for Standardization (ISO). Geneva.
- [44] Nijhuis, A., de Jong, S., Vandeginste, B.G.M. (1997). Multivariate statistical process control in chromatography. *Chemometrics and Intelligent Laboratory Systems*, 38, 51-62.

3.6. ARTICULO 2

Multivariate approaches for chemical chromatographic control of the stability of the organoleptic quality of olive oil reference materials for sensory analysis – Part II: Applications

(Artículo enviado a revisión en Noviembre 2017 a la revista: Journal of the Science of Food and Agriculture)



Multivariate approaches for chemical chromatographic control of the stability of the organoleptic quality of olive oil reference materials for sensory analysis – Part II: Applications

Journal:	<i>Journal of the Science of Food and Agriculture</i>
Manuscript ID	JSFA-17-3325
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	13-Nov-2017
Complete List of Authors:	Valverde-Som, Lucía; Universidad de Granada Facultad de Ciencias, Analytical Chemistry Ruiz-Samblás, Cristina; Universidad de Granada Facultad de Ciencias, Analytical Chemistry Rodríguez-García, Francisco P.; Agricultural and Fishery Research Institute, Consejería de Agricultura, Pesca y Desarrollo Rural Cuadros-Rodríguez, Luis ; University of Granada, Department of Analytical Chemistry
Key Words:	Virgin olive oil, stability, monitoring multivariate control chart, PCA, similarity analysis, distances

Abstract

The organoleptic quality of virgin olive oil depends on positive and negative sensory attributes. These attributes are related to volatile organic compounds and phenolic compounds which represent the aroma and taste (flavour) of the virgin olive oil. The flavour is the characteristic which may be measured by a taster panel. However, as any analytical measuring device, the tasters, individually, and the panel, as a whole, should be harmonized and validated and proper olive oil standards are needed.

In this manuscript, the multivariate approaches which have been discussed in Part I, are put into practice and the rules to build a multivariate control chart from chromatographic volatile fingerprinting and chemometrics. Fingerprinting techniques provide analytical information without identify and quantify the analytes. This methodology is used in order to monitor the stability of sensory reference materials.

Keywords

Virgin olive oil, stability, monitoring multivariate control chart, PCA, similarity analysis, distances.

1. Introduction

The organoleptic quality of virgin olive oil (VOO) depends on both positive and negative sensory attributes that are characteristic of each VOO. Positive attributes are fruity, bitter and pungent; negative attributes or defects can be fusty/muddy sediment, musty-humid-earthy, winey-vinegary-acid-sour, rancid, frostbitten olives (wet wood) and other ones which appears in minor extent (i.e., burnt, greasy, metallic, grubby, among others) [1]. The enzymatic oxidation generated by the lipoxygenase pathway is the main producer for pleasant aromas (positive aroma perception). Whereas the over-ripening of the fruit, chemical oxidation and exogenous enzymes, are primarily responsible for unpleasant aromas (sensory defects) [2].

The organoleptic attributes are related with volatile organic compounds (VOCs) and phenolic compounds (PhCs) which are responsible for the aroma and taste of the virgin olive oil [2,3,4,5]. Aroma and taste are perceived as a single sensation and it is a complex sensory impression of many individual compounds [2]. Thus, the relationship between VOO chemical composition and sensory attributes is a complex process and it remains an unfinished matter of flavouring studies. Aldehydes and alcohols are the main compounds in aroma of VOO. In addition, other ones as ketones, hydrocarbons, esters, furans, acids and aromatics compounds were found in smaller quantities [3,6,7,8].

The taste is mainly the result of bitterness and pungency attributes which are linked to the presence of PhCs. These compounds are related

to the freshness of virgin olive oil, since they are responsible for the fruity attribute in VOO. Phenolic acids, complex oleuropein derivatives (catecholic groups in a molecule), flavonoids, lignans, and hydroxyl-isocromans are the five classes of PhCs which could be found in VOO [5,9]. There are also some volatile phenolic compounds which are associated with fusty and musty attributes [5,10].

These compounds are analysed by chromatographic methods [11,12,13]. But the analytical procedures to characterize the VOC and PhC profiles include a previous separation step. For the case of VOC there are different extraction techniques such as: static headspace, dynamic headspace (purge and trap, and thermic desorption); headspace solid-phase microextraction; supercritical fluid technologies; high capacity headspace sorptive extraction [2]; and other ones as stir bar sorptive extraction; or liquid phase microextraction [14]. In order to find the relationships between VOCs and organoleptic attributes, many authors suggest the use of gas chromatography-olfactometry [3,15,16,17]. Next, mass spectrometry (MS) is used to select and identify the compounds and to describe their corresponding sensory perception [18,19].

Furthermore, PhCs need different procedures of extraction and clean-up to isolate these compounds before injection. The principally methods for extraction is based on liquid-liquid extraction; and solid-phase extraction. After that, liquid chromatographic is universally used to determine the PhCs profile. The most common detectors are

UV/Vis/diode array detection and mass spectrometry which includes mainly electrospray ionization [20,21,22].

Over the years different VOCs and PhCs were considered as odorant characteristics of some positive and negative attributes. For example, there are several VOC such as butanal, pentanal, pentanol, E-3-hexen-1-ol, butyl acetate, hexyl acetate, 1-penten-3-one, 6-methyl-5-hepten-2-one, among others compounds that are responsible for green flavour. As well as compounds such as heptanal, octanal, 2,4-heptadienal, 2-methyl butan-1-ol, heptan-2-ol, butanoic acid, octanoic acid, among other that are responsible for negative attributes as fatty, woody, rancid, winey or mushroom [2,23,24]. In the case of PhCs, derivate of oleuropein aglycon, isomer of ligstroside aglycon, hydrophobic polyphenols, among other compounds are responsible for bitter and astringent attributes [5,25,26]. All these compounds could be selected as potential chemical markers of such attributes.

There is no direct relationship between a particular chemical marker and a single sensory attribute, so any control strategy based on quantifying VOCs is bound to be unsuccessful. For this, the use of fingerprinting chromatographic signal is proposed in order to avoid loss of information. Two additional advantages could be highlighted: (i) minor run time to acquire a fingerprinting signal, and (ii) no need of identifying and quantifying each VOC individually. However, chromatographic fingerprinting presents some disadvantages, the main one is the data treatment which requires specialized software and to know how the

relevant but non-evident information could be extracted from the data [27].

In recent years, researchers are looking for different analytical methods to classify fresh olive oil in the three recognized quality categories (extra virgin olive oil, virgin olive oil and lampante olive oil) by analysing the VOC profiles to supplement the sensory analysis with instrumental methods. Some examples of these efforts are: (i) the use of atmospheric pressure chemical ionization source in combination with gas chromatography (GC) coupled to hybrid quadrupole time-of-flight mass spectrometry for determining VOCs in olive oil in order to classify extra virgin, virgin and lampante categories [11]; (ii) Fourier-transform infrared coupled to chemometrics was suggested to classify and predict the quality of commercial virgin olive oil [28]; (iii) GC coupled to ion mobility spectrometry was applied for the characterisation of different olive oil categories [29]. In all these cases the data treatment is the most critical steps in order to obtain the full relevant information from the instrumental data.

In this Part II, the practical methodology for monitoring the stability of two real certified reference materials of olive oil using chromatographic and chemometric methods are explained. The fundamentals were discussed in detail in the Part I of this serial, and now they are applied on real samples. Additionally, the proposed chromatographic method presents a novelty in the analysis of VOCs as well as the use of chromatographic fingerprinting to evaluate the similarity between samples over time. The application is focused on

VOCs because the taste attributes are not very stable and usually they are not used as control feature. Therefore the monitoring of the PhCs profile was not carried out.

The proposed methodology could be extensible to other kinds of foodstuffs with organoleptic attributes.

2. Material and methods

2.1. Olive oil reference materials

The olive oil standards used as working samples in this study were provided from a project titled SensOLEO-CRM 2015, implemented for the Andalusia Regional Government (in the south of Spain). The olive oil batches were obtained from milling fresh olive fruits and deliberately prepared with different attributes at different intensity levels.

Olive oil reference materials were previously characterized by official recognized sensory testing laboratories and supra-panels testing (panel arranged by a larger number of tasters from different panels) composed of around forty tasters. They make a determination of the organoleptic attributes assigning different levels of intensity. The certificated value of each material was issued in September 2015. Three experimental studies (homogeneity, characterization and stability) are involved to certify the materials. The homogeneity and characterization studies were carried out before to be used; and then the stability study was carried out during the use of these materials and it is presented here.

Table II.1 shows the description of FRU and BRA olive oil materials which have been used as representative examples of the nine CRMs from SensOLEO-CRM 2015.

Table II.1. Description of FRU/1 and BRA/1 /2 and /3 certified reference materials of olive oil used.

Material code	Varietal composition	Description
FRU/1/ (010-390)	100% 'Hojiblanca'	Olive oil was obtained from ripe olives of flight
BRA/3/ (010-390)	100% 'Picual'	Olive was subjected to a fermentation process to generate the 'fusty/muddy sediment' defect in the olive oil
BRA/2/ (010-390)	88% 'Hojiblanca', 12% 'Picual'	Mixture 88% FRU/1 and 12% BRA/3
BRA/1/ (010-390)	97.5% 'Hojiblanca', 2.5% 'Picual'	Mixture 97.5% FRU/1 and 2.5% BRA/3

2.2. Sensory analysis

Sensory analysis was carried out by an officially recognized laboratory of sensory analysis (Agrofood Laboratory of Granada-Atarfe, Spain) applying the EU 'panel test' method, at the beginning and at the end of the project (September 2015 and November 2016) to confirm if the certified attributes were kept stable over time. They are showed in table II.2. For that, the intensity of the defects and the fruitiness were determined on a 10 cm scale according to the Commission Regulation (EEC) No 2568/91 [1].

Table II.2. Attributes and organoleptic intensities of FRU/1 and BRA/1, /2 and /3 olive oil certified reference materials, measured in September 2015 and November 2016.

Material code	Attribute	Certified organoleptic intensity \pm (k=2) uncertainty (cm) September 2015	Organoleptic intensity (cm) November 2016
FRU/1/ (010-390)	Fruity	3.0 \pm 0.4	3.7
BRA/3/ (010-390)	Fusty/muddy sediment	8.2 \pm 0.4	7.8
BRA/2/ (010-390)	Fusty/muddy sediment	5.8 \pm 0.4	5.4
BRA/1/ (010-390)	Fusty/muddy sediment	4.3 \pm 0.4	4.4

The table shows that there are no large differences between the certified values at the beginning of the project and at the end.

2.3. Analysis of VOCs in the samples

Olive oil reference materials (5 g approx.) were introduced in a 20 mL glass vial with magnetic screw-top. Firstly, ten portions from ten independent samples from each material, evenly distributed throughout the whole batch, were analysed. These experimental data set constitute the reference chromatograms which are representative of the month

zero, M_0 (September 2015). After that, every three months (December 2015, March 2016; June 2016 and September 2016), two independent aliquots of the same sample from each material were analysed. These chromatograms are representative of the month 'x', M_x , where 'x' is an ordinal number which symbolised to the moth in study (3, 6, 9 and 12).

The samples were stored in the dark, frozen at -20°C , until ready to analysis.

2.4. Quality control solution

A 0.100 mg/g quality control solution of fluorobenzene (99%, Sigma Aldrich, United States) was prepared in refined olive oil (odourless and tasteless). 100 μL of this solution were introduced in a 20 mL vial with magnetic screw-top and analysed in the same conditions that the olive oil samples.

The quality control solution was cool stored in the dark, at 4°C , until ready to analysis.

2.5. GC-FID instrument and method conditions

All chromatographic separations were performed with a VARIAN GC 3800 (PA, USA) equipped with a split/splitless injector coupled to a FID (PA, USA). The GC equipment uses a robotized system autosampler module (CombiPal, CTC ANALYTICS, Switzerland).

The analytical column was a capillary column coated with 65% diphenyl-35% dimethylpolysiloxane stationary phase (Restek Rtx-65TG; 30 m × 0.25 mm i.d. × 0.10 μm film thickness, maximum temperature 370°C; Restek Corp., Bellefonte, PA, USA).

The column oven temperature of 80°C was held for 2.5 min, then programmed at 40°C/min to 240°C and keeping at this temperature for 3.5 min. Hydrogen (99.999%, Air Liquid, Spain) and air (99.999%, Air Liquid, Spain) were used as the flame, flow rate 30 mL/min and 300 mL/min, respectively. Helium (99.999%, Air Liquid, Spain) was used as the carrier gas, flow rate 1.5 ml/min.

Starting with the experimental conditions previously reported by Cavalli *et col.* [30], the method was re-optimised in order to minimize the run time until 10 min. In this case, the chromatogram has not to be fully resolved because a low resolution is enough for applying fingerprinting methodology properly.

2.6. HS-SPME and injection conditions

Previous to the chromatographic analysis, the VOCs were extracted from the headspace using a fibre 50/30 μm divinylbenzene/carboxen/polydimethylsiloxane (DVB/CAR/PDMS) StableFlex 24 Ga (Supelco Analytical, Bellefonte, PA, USA).

The samples of olive oil and the quality control solution were incubated in a shaker oven at 40 °C during 2 min (250 rpm). The extraction was carried out during 30 min (500 rpm). Fibre thermal desorption into GC

was at 250°C during 2 min in splitless injection mode, and then the fibre was cleaned into GC with 1:100 split ratio during 6 min. Thus, the fibre remains into the injector 8 min.

2.7. Chemometrics - Data treatment

Scan control and data acquisition were performed by version 6.9 MS Workstation (VARIAN) software data system in CSV format. Data treatment was performed using MATLAB® version 8.2 (Mathworks Inc., Natick, MA, USA), and a MATLAB® function programmed named 'Medina' (version 14) [31]. This function implements several algorithms from MATLAB Bioinformatics Toolbox™, and the 'icoshift' (interval correlation optimized shifting) which is an algorithm to align the chromatogram peaks [32]. Previous to the chemometric data analysis, a pre-processing step of the data is required. This consists of (1) grouping data vectors (replicates) from the same material in order to build a single matrix that collects all chromatographic signals; (2) removing the end of the chromatograms which has not information; (3) filtering and removing noise; (4) drift correction and zeroing of the baseline; (5) peaks alignment regarding to the retention times. After these steps, an averaged signal is generated to compare the similarity of targeted profiles.

Then, for each comparison, a similarity index is calculated which is graphed in a multivariate control chart so, the stability is monitored.

3. Results and discussion

The results are going to be showed in two parts which explain the way to carry out the stability monitoring from: (i) univariate control chart of the sensory results of panel test; and (ii) multivariate control chart built from the experimental chromatographic data.

3.1. Sensory analysis

Typically, univariate control charts are used to monitor a single quality characteristic versus time. In this case, the organoleptic intensity of the certified attribute is enough to prove the stability of samples. There are two ways of plotting the univariate control chart, with: (i) either the value of the organoleptic intensity directly; or (ii) the result of subtracting from the organoleptic intensity at the control time, the reference organoleptic intensity. In this last case the value of the reference organoleptic intensity is set to zero. Also, it is possible to monitor in a single chart several materials with different intensities. In both cases, the values are plotted *vs* time (month).

The acceptance limits would be calculated as three times the standard deviation of the considered control measuring method [33]. In this case, a value of 0.7 cm is considered as representative of the reproducibility of the panel test method and it could be selected as control limits. This value was found by analysing the results of a large number of interpanel proficiency tests [34].

Table II.2 shows the organoleptic intensities of the olive oil samples which were chosen as example. The organoleptic intensities remain the same, taking into account the data uncertainties, during their time of use.

3.2. Chromatographic analysis

The VOC profile was selected as target to acquire the chromatographic signals and monitor the stability. Such signals are staked in data vectors which can be handled and compared by means of a suitable mathematical treatment.

Figure II.1 shows the chromatogram averaging from the ten analyses carried out in month zero of FRU/1 and BRA/1 materials. The chromatogram has 2800 variables approximately, which are the point of the chromatogram at each acquisition time. This variable set is a data vector which contains the reference signal information.

To the eye of the observer, both chromatograms look the same. This is due to the sample of BRA/1 contains 97.5% of FRU/1 (see table II.1). However, when data treatment is applied, some differences appear in the variables 100, 200, 750, and 1800. These variables are linked to chromatographic retention times of approximately 1.4, 1.6, 2.5, and 4.3 minutes, respectively.

The representative chromatogram signals from the reference and from the month in study are compared by pair and, in each case, a similarity index is calculated. Figure I.2, which was showed in the Part I of this

serial, described the strategies and approaches to be applied. Next, detailed examples of each approach are described.

Approach 1. Distance-based closeness analysis

The distance between the chromatograms could be calculated from: (i) all the data of the whole chromatographic signal; (ii) a few variables previously selected by applying a proper variable reduction method.

As representative similarity indexes, nearness index (NEAR), which could be calculated from the whole signals or applying the control markers strategy on the selected variables. Also, the quadratic mean (QM) which includes both Hotelling's T^2 and Q^2 -residual, are shown in these examples. Then, multivariate control charts are built versus control time (each month). In all cases the acceptance limits are settled with a deterministic threshold.

Figure II.2 shows the multivariate control charts from FRU/1 and BRA/1 materials which were built with the whole chromatographic data vector.

As it may be seen, in the first two data, the value of similarity is almost the same (value of 0.90), and although in month 9 the values descend to 0.85 for FRU/1 and 0.83 for BRA/1, the stability remains under control.

Next, principal components analysis (PCA) was applied to reduce the dimensionality of the data vectors and to select control markers. The identification of these control markers is a specific strategy which is carried out on the reference chromatograms from month zero.

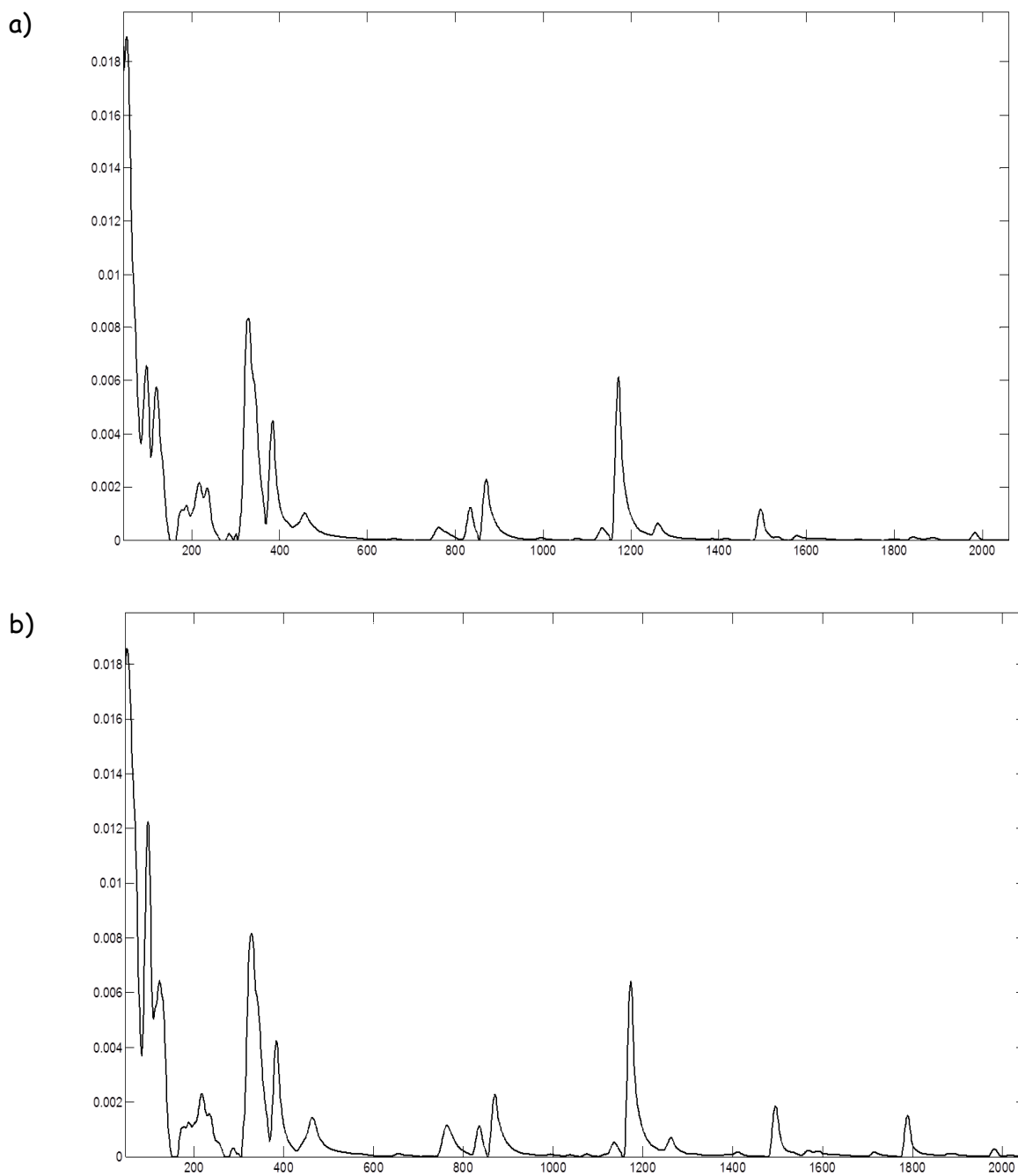


Figure II.1. Initial chromatogram averaging, at time zero, from the samples (certified reference materials): a) FRU/1, b) BRA/1.

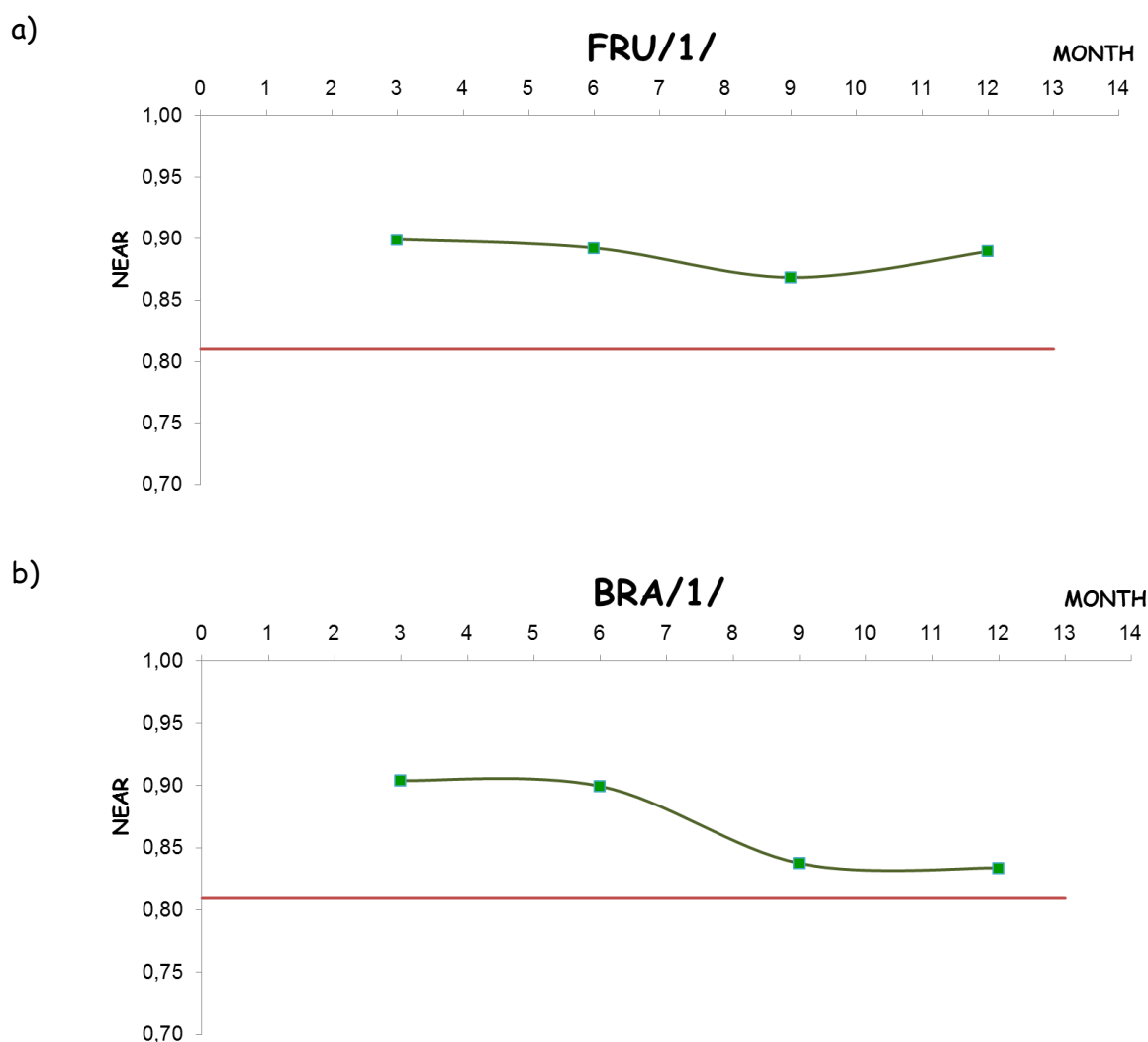


Figure II.2. Nearness index (NEAR) multivariate control chart from the whole chromatographic data vector: a) FRU/1, b) BRA/1.

Once the control markers are selected, the areas of the peaks gathering such control markers are chosen as characteristic variable during the entire study and NEAR index is then calculated to build the multivariate control chart.

In this example, FRU/1 and BRA/3 materials have a specific attribute, fruity and 'fusty/muddy sediment', respectively. In addition, BRA/1 and BRA/2 materials are blends of FRU/1 and BRA/3 in varying proportions

(see table II.1), therefore both materials share common attributes. Figure II.3 shows the two first principal components scores plot from the whole chromatographic data vector of this set of olive oil reference materials.

Four groupings, which correspond to the four reference materials, could be easily distinguished. Moreover, PC1 explained 94.86% of the variance and it differentiates between fruity attribute (left or negative region) and 'fusty/muddy sediment' defect (right or positive region).

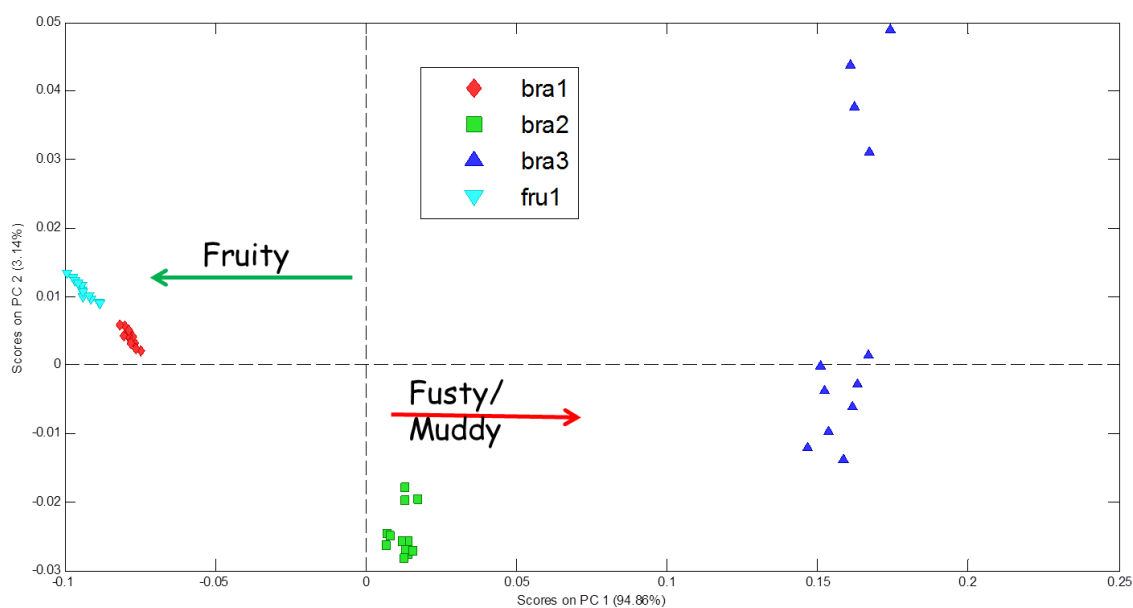


Figure II.3. PC-scores plot from the VOC chromatographic fingerprint data of the initial analysis, at time zero, of FRU/1, BRA/1, BRA/2 and BRA/3 reference materials.

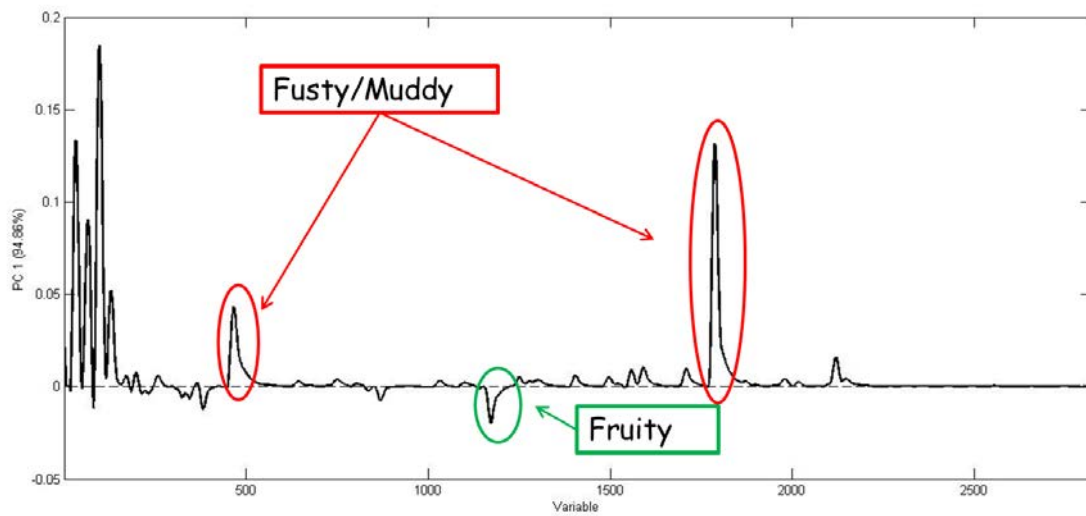


Figure II.4. Plot of PC1-loading versus data vector variable number of FRU/1, BRA/1, BRA/2 and BRA/3 materials.

Furthermore, from the PC-loadings plot, the peaks responsible for each specific attribute can be identified. The plot of the PC1-values versus the variable number (Figure II.4) shows two positive peaks which could be considered as specific of the 'fusty/muddy sediment' defect and one negative peak as specific of the fruity attribute.

The 'fusty/muddy sediment' peaks match to the variables 465 and 1785 from the data vector (approximately, 2.0 and 4.2 minutes in the chromatogram, respectively), whereas the fruity peak matches to the variable 1170 (approximately, a retention time of 3.2 minutes). Then, the areas of these peaks are selected as control markers to calculate the NEAR index and to build the multivariate control chart, which can be seen in figure II.5. Notice that these control markers are just specific for these samples and it is not appropriate to extend this selection to others olive oil samples.

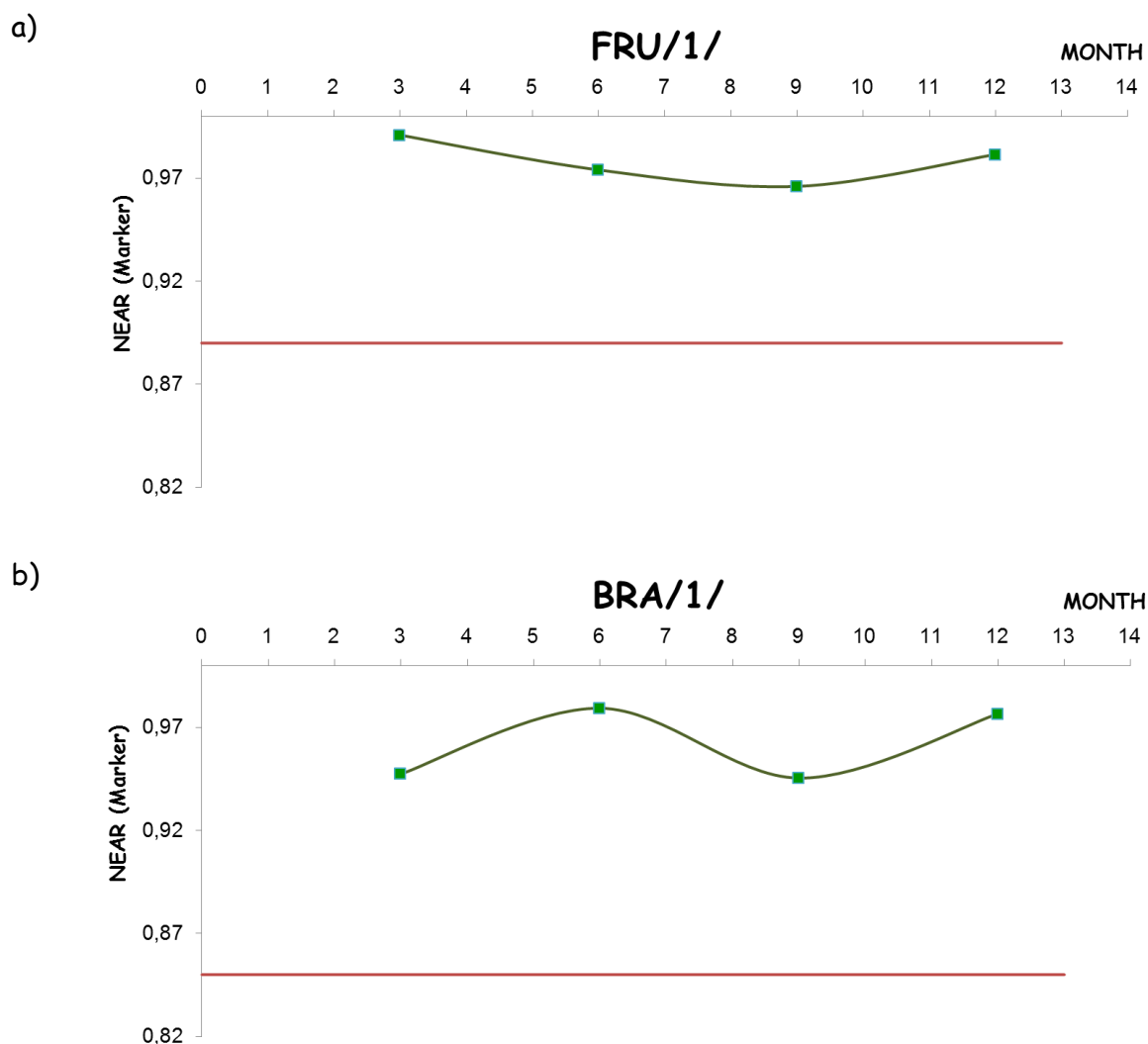


Figure II.5. Nearness index (NEAR) multivariate control chart from the selected markers: a) FRU/1, b) BRA/1 materials.

As conclusion, with the multivariate control charts, it can be said that both FRU/1 and BRA/1 remain stable because the NEAR index values are within the limits.

Finally, PCA is applied again to reduce the dimensionality of the vector. Figure II.6 shows the PC-scores plot of FRU/1 and BRA/1 materials, collecting all the acquired chromatograms for each month.

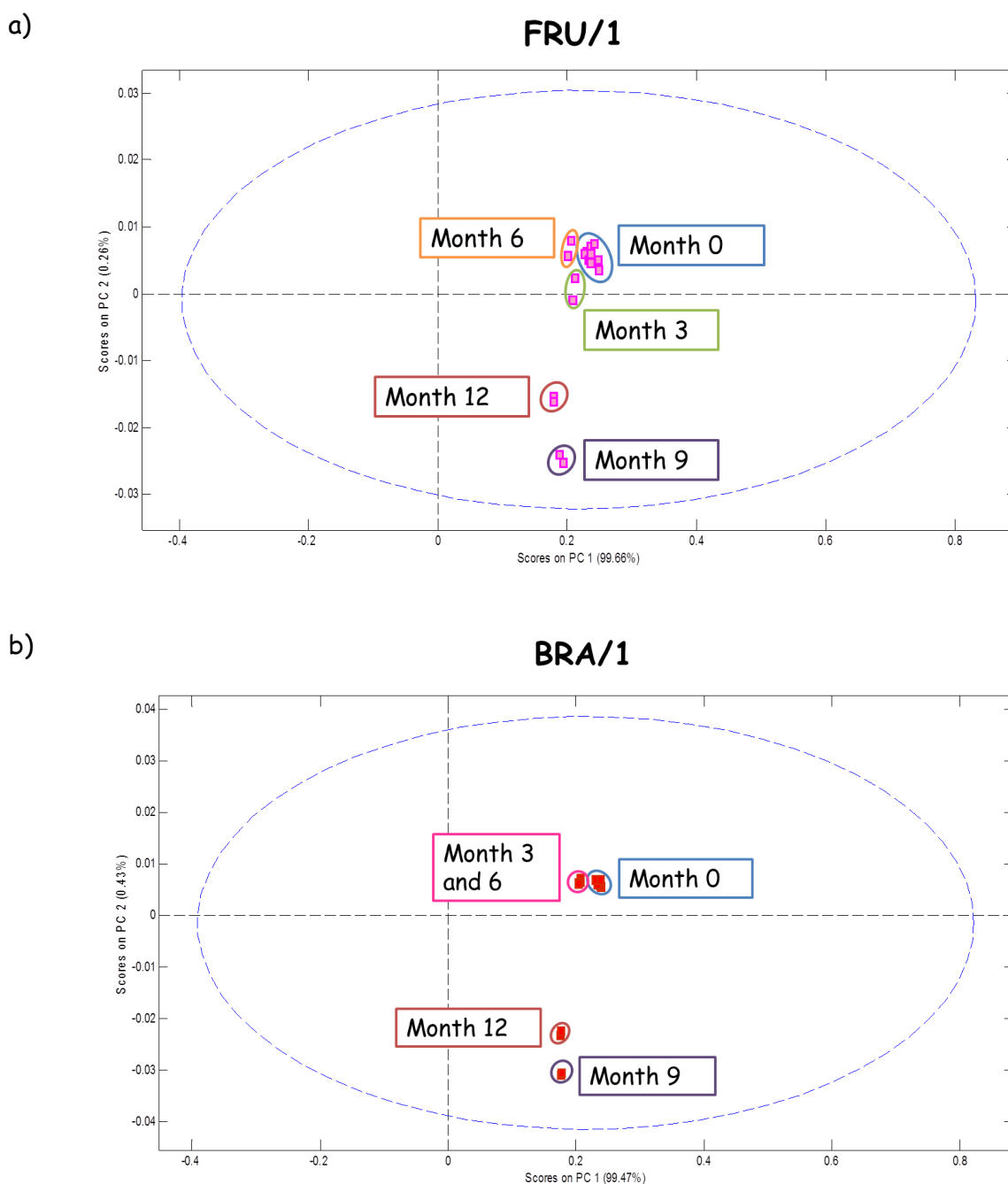


Figure II.6. PC2-PC1 plane obtained from the fingerprint data of VOC extraction. PC-scores plot from the VOC fingerprint data of each control month: a)FRU/1, b) BRA/1 materials.

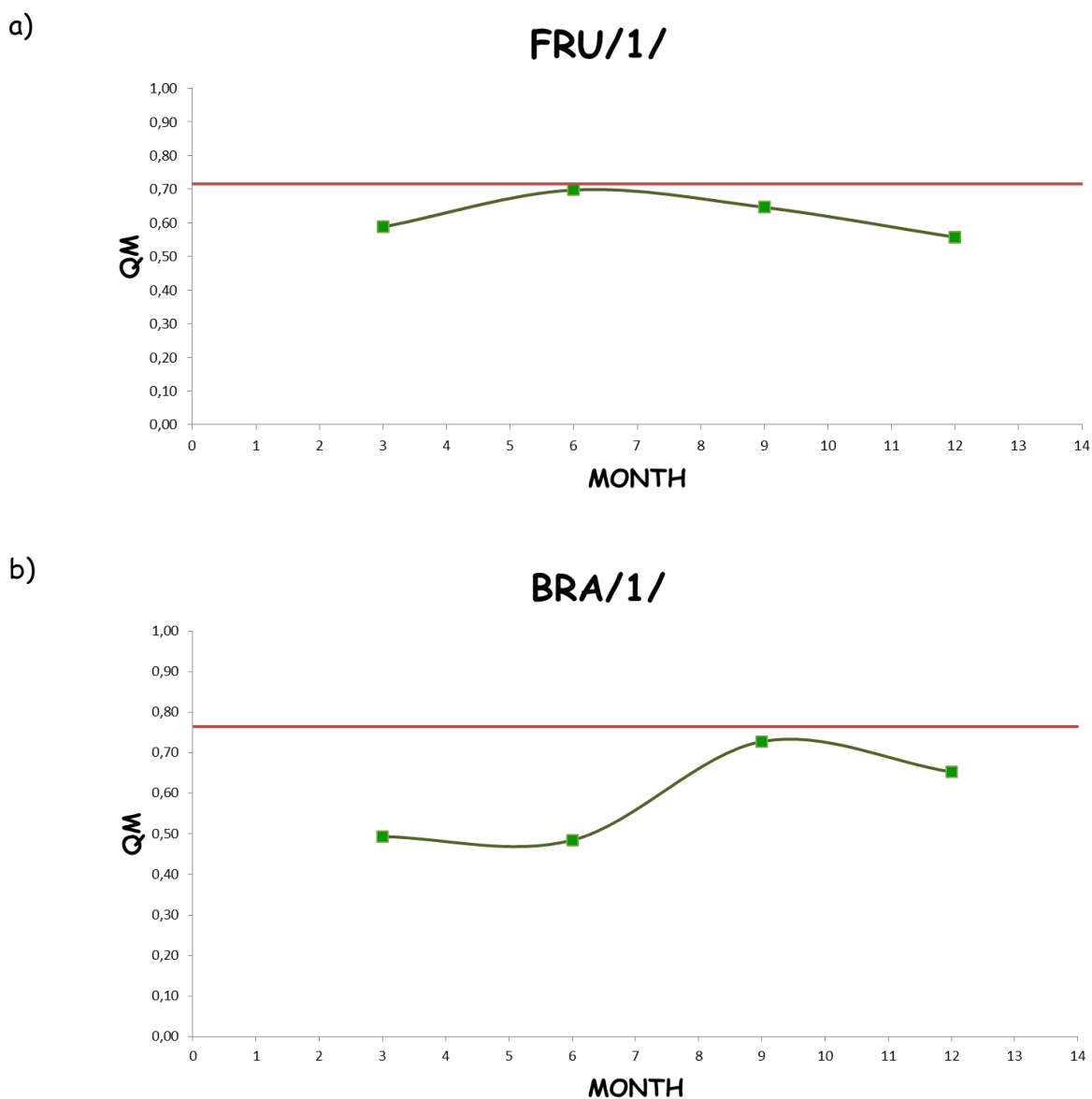


Figure II.7. T^2 - Q^2 quadratic mean (QM) multivariate control chart from the whole chromatographic data vector: a) FRU/1, b) BRA/1.

The reduced Hotelling's T^2 and Q^2 -residual values, normalized respect to the critical value of them at 0.95 confidence level (tabulated), are directly obtained from PCA. These normalized values range from 0-1. Figure II.7 shows the T^2 - Q^2 quadratic mean chart. Once again, both

FRU/1 and BRA/1 materials should be considered as stable, since all data appear under the limit.

Approach 2. Other similarity analysis

For this approach, the entire data vector from each cumulative chromatographic signal (CIS) is considered.

Figures II.8 and II.9 show the multivariate control charts when the determination coefficient, R^2 , and the cosine similarity, COS , are respectively used as control metrics. In all cases, the values are always close to one indicating a high similarity.

To sum up, a valuable and rigorous methodology for monitoring the organoleptic stability of olive oils, by means of the VOC chromatographic profile has been proposed. The main merit of the proposed study is the integration of multivariate control charts to monitor the stability of reference materials for sensory analysis of olive oil with chromatographic analysis which has not been described in bibliography yet. Thus, this methodology could be applied in parallel with panel test method but having the advantage that sensory analysis would not be necessary each control time, only when the data are out of control. However, sensory analysis could be carried out at the beginning and at the end of the study to check the organoleptic intensities of the control attributes. For this reason, the combination of chromatographic analysis and multivariate control charts provides a supplementing methodology to panel test method for monitoring the stability of the

organoleptic quality and it would be an important advance for the quality control of olive oil products.

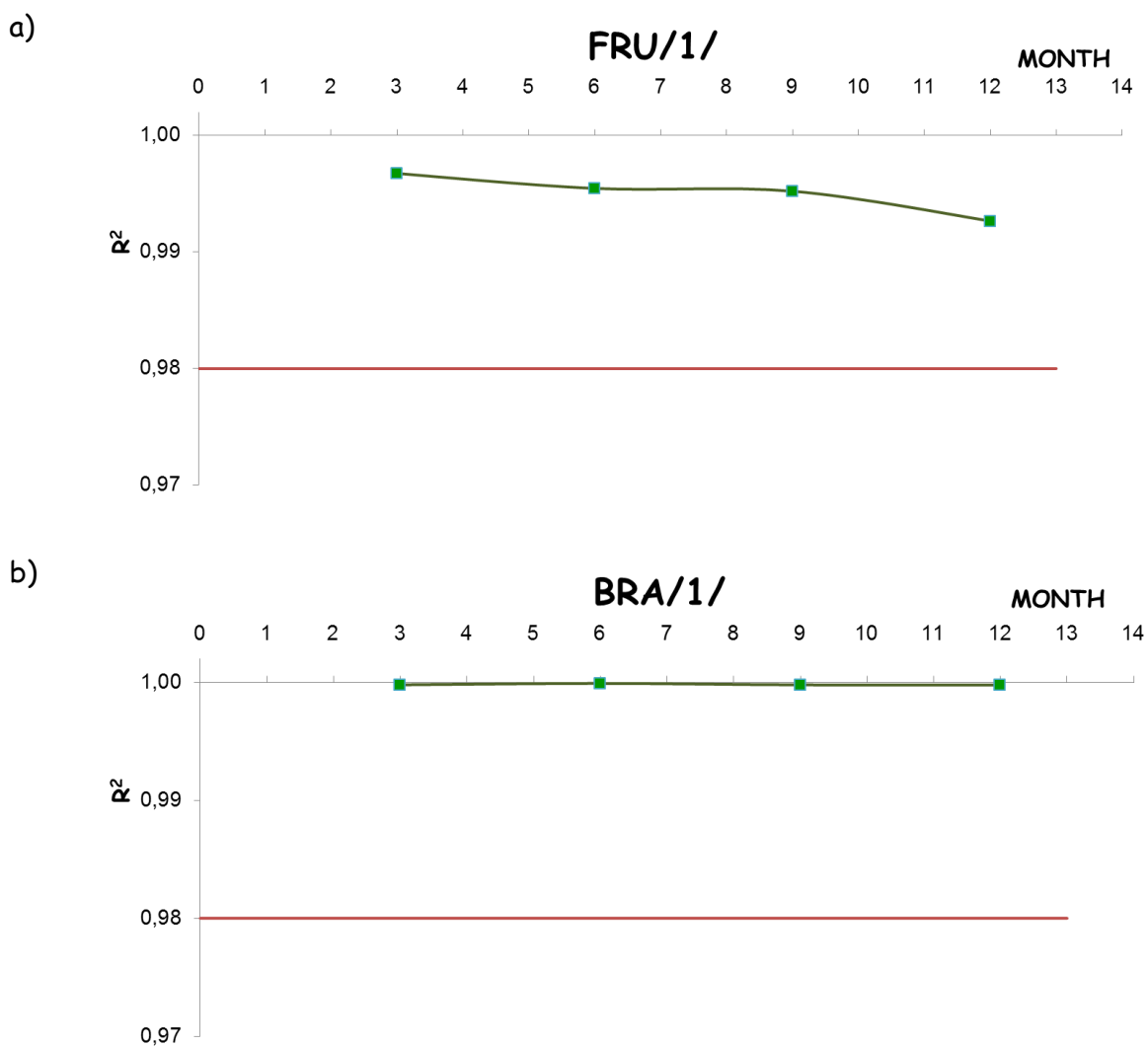


Figure II.8. Determination coefficient (R^2) multivariate control chart from the whole chromatographic data vector: a) FRU/1, b) BRA/1.

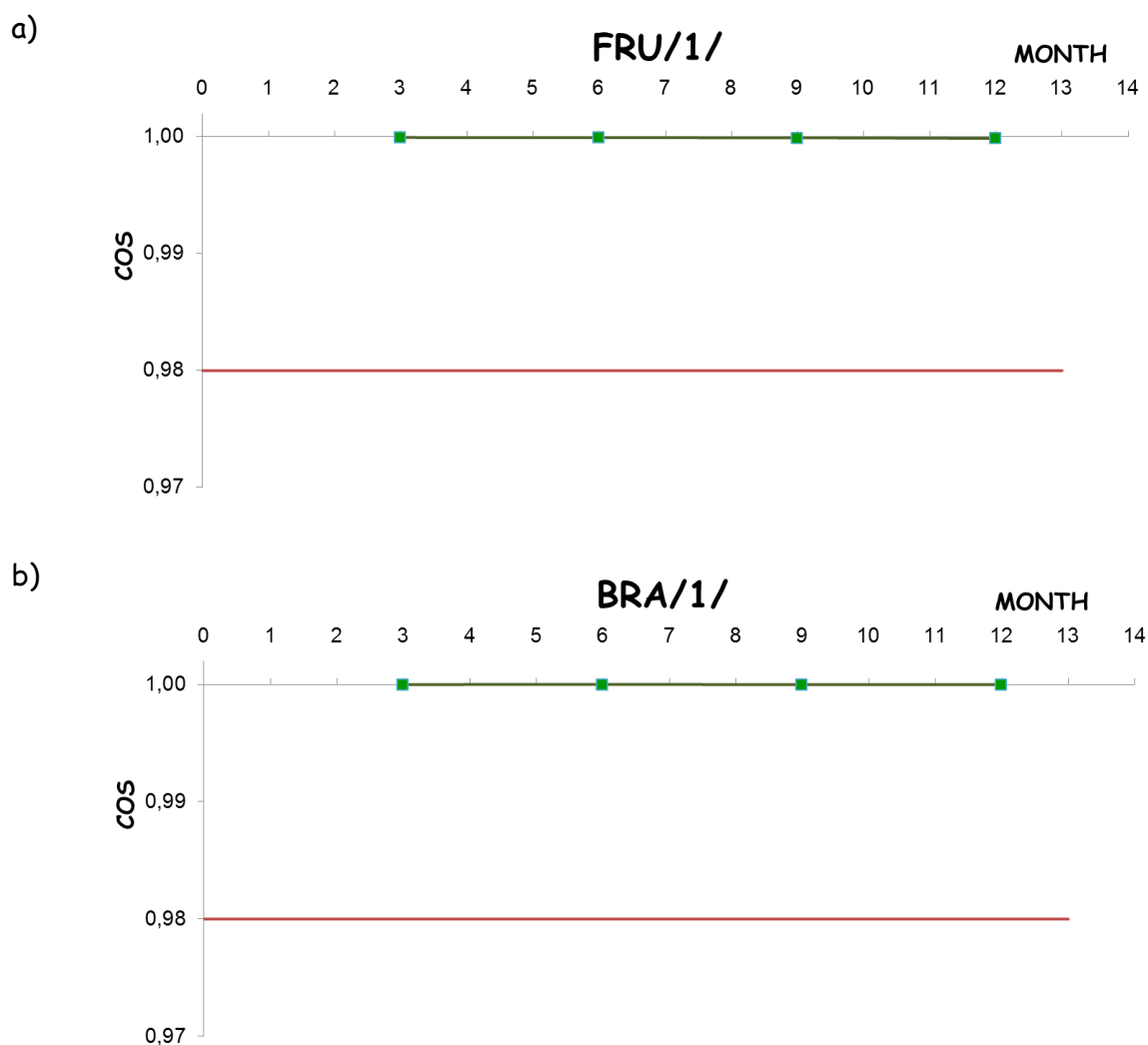


Figure II.9. Cosine similarity (COS) multivariate control chart from the whole cumulative chromatographic data vector: a) FRU/1, b) BRA/1.

Abbreviations used

CIS: Cumulative instrumental signal

COS: Cosine similarity index

CRMs: Certified reference materials

d(eucl): Euclidean distance

d(mahal): Mahalanobis distance

FID: Flame ion detector

GC: Gas chromatography

HAND: Handiness similarity index

HS-SPME: headspace solid-phase microextraction

MS: Mass spectrometry

NEAR: Nearness similarity index

PCA: Principal component analysis

PC: Principal component

PhC: Phenolic compound

QM: Quadratic mean

VOC: Volatile organic compound

VOO: Virgin olive oil

Acknowledgments

The authors would like to thank the Andalusia Regional Government (Consejería de Agricultura, Pesca y Desarrollo Rural, Spain) and European funds (ERDF, European Regional Development Fund in Andalusia) for financial assistance with the project SensOLEO-2015. The authors are very grateful to the olive oil samples producer Centro Venta del Llano, Mengíbar, Spain.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- [1] Commission Regulation (EEC) No 2568/91 on characteristics of olive oil and olive-residue oil and on the relevant methods of analysis (consolidated version 01991R2568), Official Journal of the European Union (2016), 031.001, 1-128.
- [2] Morales, M.T., Aparicio-Ruiz, R., Aparicio, R. (2013). Chromatographic methodologies: compounds for olive oil odor issues. In R. Aparicio, J. Harwood (Eds.), *Handbook of olive oil: analysis and properties* (pp. 261-309). 2nd ed. New York: Springer.
- [3] Taticchi, A., Esposto, S., Servili, M. (2014). The basis of sensory properties of virgin olive oil. In E. Monteleone, S. Langstaff (Eds.), *Olive oil sensory science* (pp. 33-54). Chichester: Wiley.
- [4] García-González, D.L., Morales, M.T., Aparicio, R. (2010). Olive and olive oil. In Y.H. Hui (Ed.), *Handbook of fruit and vegetable flavours* (pp. 821-847). Chichester: Wiley.
- [5] Tsimidou, M.Z. (2013). Analytical methodologies: phenolic compounds related to olive oil taste issues. In R. Aparicio, J. Harwood (Eds.), *Handbook of olive oil: analysis and properties* (pp. 311-333). 2nd ed. New York: Springer.
- [6] Luna, G., Morales, M.T., Aparicio, R. (2006). Characterisation of 39 varietal virgin olive oils by their volatile composition. *Food Chemistry*, 98, 243-252.
- [7] Romero, I., García-González, D.L., Aparicio-Ruiz, R., Morales, M.T. (2015). Validation of SPME-GCMS method for the analysis of virgin olive oils volatiles responsible for sensory defects. *Talanta*, 134, 394-401.
- [8] Kotsiou, K., Tasioula-Margari, M. (2015). Changes occurring in the volatile composition of Greek virgin olive oils during storage: oil variety influences stability. *European Journal of Lipid Science and Technology*, 117, 514-522.

- [9] Kalogeropoulos, N., Kaliora, A.C. (2015). Effect of fruit maturity on olive oil phenolic composition and antioxidant capacity. In D. Boskou (Ed.), *Olive and olive oil bioactive constituents* (pp. 123-145). Urbana: AOCS.
- [10] Vichi, S., Romero, A. Tous, J., López Tamames, E., Buxaderas, S. (2008). Determination of volatile phenols in virgin olive oils and their sensory significance. *Journal of Chromatography A*, 1211, 1-7.
- [11] Sales, C., Cervera, M.I., Gil, R., Portolés, T., Pitarch, E., Beltran, J. (2017). Quality classification of Spanish olive oils by untargeted gas chromatography coupled to hybrid quadrupole-time of flight mass spectrometry with atmospheric pressure chemical ionization and metabolomics-based statistical approach. *Food Chemistry*, 216, 365-373.
- [12] Reboredo-Rodríguez, P., González-Barreiro, C., Cancho-Grande, B., Valli, E., Bendini, A., Gallina Toschi, T., Simal-Gandara, J. (2016). Characterization of virgin olive oils produced with autochthonous Galician varieties. *Food Chemistry*, 212, 162-171.
- [13] Zhu, H., Wang, S.C., Shoemaker, C.F. (2016). Volatile constituents in sensory defective virgin olive oils. *Flavour and Fragrance Journal*, 31, 22-30.
- [14] Vichi, S. (2010) Extraction techniques for the analysis of virgin olive oil aroma. In V.R. Preedy, R.R. Watson (Eds), *Olives and olive oil in health and disease prevention* (pp. 615-623). London: AP.
- [15] van Ruth, S.M., (2011). Techniques for sampling and identification of volatile compounds contributing to sensory perception. In L.M.L. Nollet, F. Toldrá (Eds), *Sensory analysis of foods of animal origin* (pp.39-48). New York: CRC.
- [16] Inarejos-García, A.M., Santacatterina, M., Salvador, M.D., Fregapane, G., Gómez-Alonso, S. (2010). PDO virgin olive oil quality-Minor components and organoleptic evaluation, *Food Research International*, 43, 2138-2146.

- [17] Kalua, C.M., Allen, M.S., Bedgood, Jr., Bishop, A.G., Prenzler, P.D., Robards, K. (2007). Olive oil volatile compounds, flavour development and quality: A critical review, *Food Chemistry*, 100, 273-286.
- [18] Gomes da Silva, M.D.R., Costa Freitas, A.M., Cabrita, M.J.B., Garcia, R. (2011). Olive oil composition: volatile compounds. In B. Dimitrios (Ed), *Olive oil - constituents, quality, health properties and bioconversions* (pp.17-46). Rijeka: InTech.
- [19] Kalua, C., Prenzler, P., Ryan, D., Robards, K. (2010). Volatile compounds in Australian olive oils: how different are they from other oils. In V.R. Preedy, R.R. Watson (Eds), *Olives and olive oil in health and disease prevention* (pp. 201-209). London: AP.
- [20] Gómez-Caravaca, A.M., Lozano-Sánchez, J., Contreras Gámez, M.d.M., Segura Carretero, A. (2015). Bioactive phenolic compounds from *olea europaea*: a challenge for analytical chemistry. In D. Boskou (Ed.), *Olive and olive oil bioactive constituents* (pp. 123-145). Urbana: AOCS.
- [21] Segura-Carretero, A., Carrasco-Pancorbo, A., Bendini, A., Cerretani, L., Fernández-Gutiérrez, A. (2010). Analytical determination of polyphenol in olive oil. In V.R. Preedy, R.R. Watson (Eds), *Olives and olive oil in health and disease prevention* (pp. 509-523). London: AP.
- [22] Dais, P., Boskou, D. (2009). Detection and quantification of phenolic compounds in olive oil, olives, and biological fluids. In D. Boskou (Ed), *Olive oil, minor constituents and health* (pp. 55-107). Boca Raton: CRC.
- [23] Checchi, T., Alfei, B. (2013). Volatile profiles of Italian monovarietal extra virgin olive oils via HS-SPME-GC-MS: Newly identified compounds, flavors molecular markers, and terpenic profile. *Food Chemistry*, 141, 2025-2035.
- [24] Boskou, D., Blekas, G., Tsimidou (2006). Olive oil composition. In D. Boskou (Ed), *Olive oil chemistry and technology* (pp. 41-72). 2nd ed. Champaign: AOCS.

- [25] Boskou, D. (2009). Phenolic compounds in olives and olive oil. In D. Boskou (Ed), *Olive oil, minor constituents and health* (pp. 11-44). Boca Raton: CRC.
- [26] Boskou, D., Tsimidou, M., Blekas, G. (2009). Polar phenolic compounds. In D. Boskou (Ed), *Olive oil chemistry and technology* (pp. 73-92). 2nd ed. Champaign: AOCS.
- [27] Cuadros-Rodríguez, L., Ruiz-Samblás, C., Valverde-Som, L., Pérez-Castaño, E., González-Casado, A. (2016). Chromatographic fingerprinting: An innovate approach for food "identification" and food authentication - A tutorial. *Analytica Chimica Acta*, 909, 9-23.
- [28] Hirri, A., Bassbasi, M., Platikanov, S., Tauler, R., Oussama, A. (2016). FTIR spectroscopy and PLS-DA classification and prediction of four commercial grade virgin olive oils from Morocco. *Food Analytical Methods*, 9, 974-981.
- [29] Garrido-Delgado, R., Dobao-Prieto, M.M., Arce, L., Valcárcel, M. (2015). Determination of volatile compounds by GC-IMS to assign the quality of virgin olive oil. *Food Chemistry*, 187, 572-579.
- [30] Cavalli, J.F., Fernandez, X., Lizzani-Cuverlier, L., Loiseau, A.M. (2013). Comparison of static headspace, headspace solid phase microextraction, headspace sorptive extraction, and direct thermal desorption techniques on chemical composition of French olive oils. *Journal of Agricultural Food Chemistry*, 51, 7709-7716.
- [31] Pérez-Castaño, E., Ruiz-Samblás, C., Medina-Rodríguez, S., Quirós-Rodríguez, V., Jiménez-Carvelo, A.M., Valverde-Som, L., González-Casado, A., Cuadros-Rodríguez, L. (2015). Comparison of different analytical classification scenarios: application for the geographical origin of edible palm oil by sterolic (NP) HPLC fingerprinting. *Analytical Methods*, 7, 4192-4201.

- [32] Tomasi, G., Savorani, F., Engelsen, S.B. (2011). Icoshift: An effective tool for the alignment of chromatographic data. *Journal of Chromatography A*, 1218, 7832-7840.
- [33] ISO 7870-2:2013, Control chart. Part 2: Shewhart control charts, International Organization for Standardization (ISO). Geneva.
- [34] T.28/Doc. No 1/Rev. 1 (2017), Guidelines for the accomplishment of the requirements of the norm ISO 17025 by the laboratories of sensory analysis of virgin olive oil, International Olive Council. Madrid.

3.7. CONCLUSIONES

A la vista de los resultados obtenidos, podemos concluir que se ha desarrollado un **método para la obtención del perfil de compuestos orgánicos volátiles** en muestras de aceite de oliva virgen a partir de la extracción con fibra de **SPME** y con el equipo **GC-FID**. Además, con este método es posible **monitorizar la estabilidad organoléptica en muestras de patrones de aceite de oliva virgen** a partir de los datos multivariantes instrumentales obtenidos de la "huella dactilar" cromatográfica de la fracción de VOCs.

Han sido descritos diferentes índices de similitud que se pueden usar para representar los datos multivariantes en gráficos de control. En este capítulo se explica una **forma aproximada de calcular la distancia de Mahalanobis** a partir del cálculo con réplicas, en el caso en el que el número de muestras es menor al número de variables. Además, se propone un **nuevo parámetro** que normaliza dicha distancia entre los valores cero y uno y que se le ha dado el nombre de "**índice de manejabilidad**", con las siglas: **HAND**.

También se propone la **media cuadrática, QM**, a partir del T^2 -Hotelling reducida y Q^2 -residual reducida para representar el **gráfico de control multivariable**, en lugar de representar cada una de ellas por separado.

A partir de los ejemplos propuestos con las muestras de aceite de oliva virgen, se puede confirmar el uso de este tipo de gráficos de control para monitorizar la estabilidad organoléptica.

Esto es de gran importancia en el sector del aceite de oliva. Hasta la fecha no se había elaborado antes este tipo de propuesta de control instrumental para la monitorización de la estabilidad y que permite un control complementario al "panel test".

Por último señalar que, esta metodología puede aplicarse también al control de la estabilidad organoléptica en otro tipo de muestras.

3.8. COMUNICACIONES A CONGRESOS

- ❖ L. Valverde-Som, C. Ruiz-Samblás, A. González-Casado, A. Carrasco-Pancorbo, F.P. Rodríguez-García, L. Cuadros-Rodríguez. **Verificación de la estabilidad organoléptica del aceite de oliva virgen mediante monitorización de la "huella dactilar" cromatográfica de los compuestos orgánicos volátiles.** XVIII Reunión de la Sociedad Española de Química Analítica (SEQA). **Póster.**
- ❖ L. Valverde-Som, C. Ruiz-Samblás, L. Cuadros-Rodríguez. **Estudio de la estabilidad organoléptica de aceites de oliva utilizando cromatografía y quimiometría.** Jornadas sobre Avances en Calidad y Tecnología Alimentaria. **Póster.**
- ❖ L. Valverde-Som, C. Ruiz-Samblás, A. González-Casado, A. Carrasco-Pancorbo, F.P. Rodríguez-García, L. Cuadros-Rodríguez. **Evaluación de la similaridad entre cromatogramas: Monitorización de la estabilidad del aceite de oliva.** V Workshop de Quimiometría para Jóvenes Investigadores. **Oral-flash y póster.**
- ❖ L. Valverde-Som, A.M. Jiménez-Carvelo, C. Ruiz-Samblás, A. González-Casado, L. Cuadros-Rodríguez. **Metodologías para determinar la similitud de cromatogramas.** I Congreso Nacional Multidisciplinar de Jóvenes Investigadores. **Oral-flash y póster.**
- ❖ L. Valverde-Som, C. Ruiz-Samblás, A. González-Casado, L. Cuadros-Rodríguez. **Procedimiento para la determinar la similitud de cromatogramas.** Congreso Virtual Iberoamericano sobre Gestión de Calidad en Laboratorios (IBEROLAB). **Publicación en revista no indexada.**
- ❖ L. Valverde-Som, C. Ruiz-Samblás, L. Cuadros-Rodríguez. **Control instrumental mediante datos multivariable de la estabilidad organoléptica de materiales de referencia certificados de aceites de**

olive virgin mediante señales cromatográficas. XVIII Simposio Científico-Técnico de Expoliva 2017. Oral y póster.

- ❖ **L. Valverde-Som, C. Ruiz-Samblás, L. Cuadros-Rodríguez. Control instrumental de la estabilidad de los compuestos orgánicos volátiles en muestras de aceite de oliva virgen mediante gráficos de control de datos multivariable. II Jornadas de Investigadores en Formación Fomentando la Interdisciplinariedad. Póster.**
- ❖ **L. Valverde-Som, C. Ruiz-Samblás, L. Cuadros-Rodríguez. Multivariate control chart for monitoring the stability of the organoleptic quality of reference materials by chromatographic data. IX Colloquium Chemiometricum Mediterraneum. Oral.**
- ❖ **L. Valverde-Som, A.M. Jiménez-Carvelo, M. Delgado-Aguilar, C. Ruiz-Samblás, F.P. Rodríguez-García, A. Carrasco-Pancorbo, A. González-Casado, L. Cuadros-Rodríguez. The production of olive oil certified reference materials for both sensory and physicochemical analysis in the University of Granada, Spain. 15th Euro Fed Lipid Congress: Oil, Fats and Lipids: New Technologies and Applications for Healthier Life. Poster Oral Pitch.**

ANEXOS II. ESTUDIOS COMPLEMENTARIOS NO PUBLICADOS

II.1. HOMOGENEIDAD DE LOS MRC

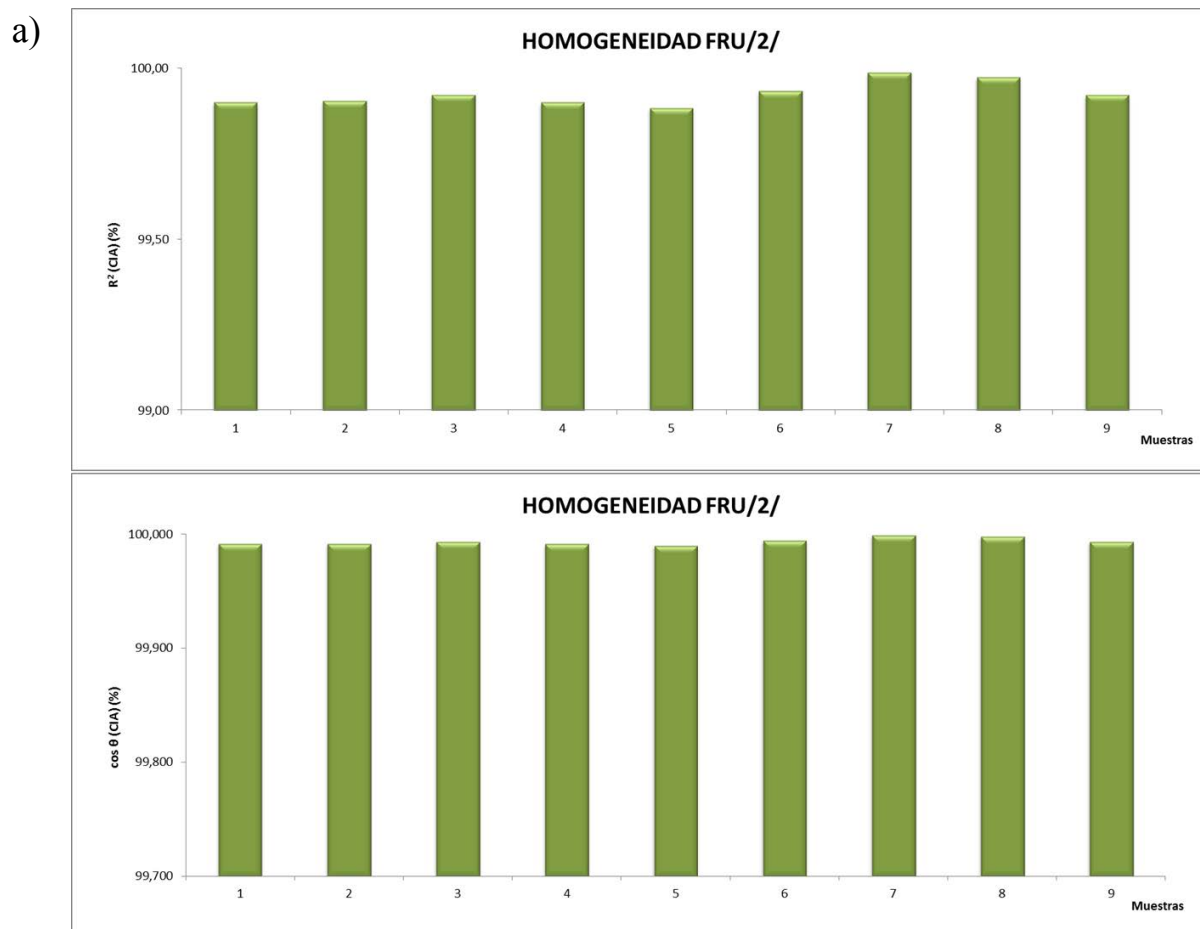
Como parte del proyecto SensOLEO-MRC 2015, se llevó a cabo el estudio de homogeneidad mediante el control instrumental de cada una de las unidades que compone cada lote de los diferentes materiales de referencia.

La homogeneidad de cada MRC, se verifica en base a la comparación de los cromatogramas correspondientes a las fracciones de compuestos volátiles (cromatografía de gases, SPME-GC/FID). El método analítico usado es el que se describe en las tablas 3.3 y 3.4.

En la realización de este estudio, se seleccionaron sistemáticamente diez unidades de cada material, distribuidas homogéneamente en el rango de 010-390, correspondiente al número de lote de cada material, y se inyectaron en el equipo de forma aleatoria.

De los diez cromatogramas obtenidos por material, se selecciona el primero como referencia y se compara con cada uno de los nueve restantes. La comparación de los cromatogramas se llevó a cabo calculando los índices de similitud (R^2 y COS, ecuaciones 12 y 13, respectivamente, que aparecen en el apartado 3.4. *proceso experimental*). Así como la tendencia (o no) de los diez cromatogramas mediante un estudio de análisis de componentes principales.

Los materiales fueron declarados homogéneos para valores de los índices de similitud R^2 y COS iguales o superiores al 90%. En la figura II.1.1 se observa: (a) un ejemplo del gráfico de barras, para el material FRU/2/, en el que se representa el índice de similitud respecto a cada uno de los 9 cromatogramas comparados con el cromatograma seleccionado de referencia; y (b) la tabla con los datos de los índices de similitud de la fracción de compuestos orgánicos volátiles del material BRA/1/.



b)

MATERIAL :	BRA/1/		
FRACCIÓN DE COMPUESTOS ORGÁNICOS VOLÁTILES			
Código de las unidades	Índices de similitud		Declaración
	R² (%)	Cosθ (%)	
082	99.96	100.00	HOMOGÉNEO
112	99.98	100.00	HOMOGÉNEO
142	99.96	100.00	HOMOGÉNEO
172	99.97	100.00	HOMOGÉNEO
202	99.95	100.00	HOMOGÉNEO
231	99.96	100.00	HOMOGÉNEO
262	99.93	99.99	HOMOGÉNEO
292	99.99	100.00	HOMOGÉNEO
322	99.98	100.00	HOMOGÉNEO

Figura II.1.1. Índices de similitud R² y COS, representado como: a) gráfico de barras; (b) tabla de datos numérica.

Los valores de similitud obtenidos para cada uno de los materiales, se encuentran por encima del 90% y el gráfico de PCA no mostró ninguna tendencia. Por lo que todos los MRCs fueron declarados CONFORMES y HOMOGÉNEOS.

II.2. SELECCIÓN DE PICOS CARACTERÍSTICOS (MARCADORES)

Con el objetivo de no trabajar con toda la señal cromatográfica completa, se plantea la reducción de la misma a través de la selección de picos característicos para cada material. Así, se podrá calcular el área del pico seleccionado y a partir de ella el índice de proximidad, reduciendo el tiempo de tratamiento de los datos.

Para llevar a cabo la selección de cada uno de los picos característicos para cada material, se trabajó con la herramienta quimiométrica de PCA.

Primeramente se seleccionaron grupos de muestras que tenían características comunes entre sí. Así, por ejemplo, se hizo un grupo con los materiales FRU/1/, BRA/3/, BRA/2/ y BRA/1/. Este grupo de muestras presentan el atributo certificado frutado, para la muestra FRU/1/ y atributo certificado atrojado/borras para la muestra BRA/3/. Ambos atributos están presentes en las muestras BRA/2/ y BRA/1/, puesto que son una mezcla de los materiales FRU/1/ y BRA/3/.

Todo este estudio quedó recogido en un informe, que se llevó a cabo en el seno del proyecto SensOLEO-MRC 2015. En este informe se indica cómo se realizó la búsqueda y selección de cada uno de los picos característicos (o marcadores) para cada MRC.

Una vez seleccionados los picos característicos, únicamente se calcula el área del mismo para obtener el valor de NEAR y hacer el gráfico de control a partir de su valor (se muestra en el ANEXO II.3). Para ello se define el valor de la variable correspondiente al valor máximo del pico seleccionado y se relaciona con el tiempo de retención.

A continuación se presenta el informe elaborado para la selección de picos característicos que sirvan de marcadores para la monitorización de estabilidad de los materiales de referencia de aceite de oliva.



PROYECTO SENSOLEO-MRC 2015
SELECCIÓN DE PICOS CARACTERÍSTICOS QUE
SIRVAN DE MARCADORES

Código: SPM
Revisión: 2.0
Página 1 de 38

PROYECTO SENSOLEO-MRC 2015

**PCAs PARA LA SELECCIÓN DE PICOS
CARACTERÍSTICOS QUE SIRVAN DE
MARCADORES
EN LA ESTABILIDAD DE
LOS MATERIALES DE REFERENCIA**

Elaborado por:	Lucía Valverde Som
Firma:	
Revisado por:	
Firma:	
Fecha de implantación:	05 Noviembre 2015

Control de revisiones

Revisión	Fecha	Modificaciones introducidas
1.0	05 Nov 2015	Primera edición del documento
2.0	13 Nov 2015	<p>Modificación respecto al primer documento, se incluyen los siguientes puntos:</p> <ol style="list-style-type: none"> 1. Objetivo y campo de aplicación, ha sido resumido. 2. Resultados del PCA de los siguientes grupos: <ul style="list-style-type: none"> - Grupo 3. FRU/1/ y FRU/2/ - Grupo 4. FRU/1/ y COC/2/. Corroborar con FRU/2/ y COC/2/ - Grupo 5. FRU/1/ e HIG/2/. Corroborar con FRU/2/ e HIG/2/ 3. Una nueva columna en la Tabla 1, en la que se indica la intensidad relativa de cada uno de los picos cromatográficos seleccionados como marcadores de cada MRC. 4. Una nueva tabla (Tabla 2), en la que se incluyen los MRC BRA/2/, BRA/1/ y MAH/1/, que al ser diluciones de BRA/3/ y MAH/2/ con FRU/1/, se han incluido los picos característicos de estos tres últimos materiales en función de la dilución. 5. Se incluye el Anexo 2: Cromatogramas de los materiales de referencia de aceite de oliva.



ÍNDICE

1. OBJETO Y CAMPO DE APLICACIÓN	4
2. DOCUMENTACIÓN DE REFERENCIA	4
3. FUNDAMENTO	4
4. EQUIPOS	5
5. PERSONAL	5
6. RESULTADOS	5
6.1. Grupo 1. FRU/1/ y BRA/1/, /2/ y /3/	7
6.2. Grupo 2. FRU/1/ y MAH/1/ y /2/	12
6.3. Grupo 3. FRU/1/ y FRU/2/	15
6.4. Grupo 4. FRU/1/ y COC/2/. Corroborar con FRU/2/ y COC/2/	19
6.5. Grupo 5. FRU/1/ e HIG/2/. Corroborar con FRU/2/ e HIG/2/	22
7. CONCLUSIONES	25
ANEXOS	
1. Descripción de los materiales de referencia de aceite de oliva	29
2. Cromatogramas de los materiales de referencia de aceite de oliva	30



1. OBJETO Y CAMPO DE APLICACIÓN

El presente documento recoge los resultados para la selección de picos cromatográficos característicos de los materiales de referencia certificados (MRC) a partir del análisis de componentes principales (PCA).

Para ello se han seleccionado diferentes grupos que engloban los 9 MRC del proyecto SensOLEO-MRC 2015 adquiridos por la Consejería de Agricultura, Pesca y Desarrollo Rural (CAPDR) de la Junta de Andalucía y coordinado por la Unidad de Cualimetría y Metrología Química (CMQ) del Departamento de Química Analítica de la Universidad de Granada (UGR). Dichos materiales han sido preparados, envasados y etiquetados en unidades independientes en las instalaciones del Centro "Venta del Llano" de Mengíbar (Jaén), perteneciente al Instituto de Formación Agraria y Pesquera, Alimentaria y de la Producción Ecológica (IFAPA) dependiente de la CAPDR de la Junta de Andalucía.

2. DOCUMENTACIÓN DE REFERENCIA

- ISO Guide 35:2006, *Reference materials: general and statistical principles for certification*.
- UNE 82130:2003 IN. *Términos y definiciones utilizados en relación con los materiales de referencia* (Equivale a ISO Guide 30:1992), Asociación Española de Normalización y Certificación (AENOR).
- UNE-ISO/IEC Guía 99:2012 IN. *Vocabulario Internacional de Metrología: conceptos fundamentales y generales y términos asociados (VIM)*. (Equivale a ISO Guide 99:2007), Asociación Española de Normalización y Certificación (AENOR).
- Unidad de Cualimetría y Metrología Química (CMQ), *Procedimiento para la Caracterización y Certificación de los Materiales de Referencia*, (documento interno).

3. FUNDAMENTO

El análisis de componentes principales es una de las técnicas principales del análisis exploratorio de datos, encuentra combinaciones de variables que describen las mayores tendencias en los datos. Si "X" es una matriz de datos con "m" filas (objetos o muestras) y "n" columnas (variables), PCA descompone "X" como suma de términos en vectores de menor o igual dimensión que "X":

- ↳ Los vectores "scores" (puntuaciones) contienen la información de cómo las muestras se relacionan unas con otras.
- ↳ Los vectores "loadings" (pesos o ponderales) contienen información de cómo las variables se relacionan unas con otras.

Cada componente principal (PC) consiste en un vector "scores" y un "loadings". PC1 tiene la varianza más alta posible. PC2 es ortogonal (en scores y loadings) y tienen la varianza siguiente más alta posible, y así sucesivamente.

La necesidad del PCA surge con el objetivo de seleccionar uno o varios picos cromatográficos



PROYECTO SENSOLEO-MRC 2015
SELECCIÓN DE PICOS CARACTERÍSTICOS QUE
SIRVAN DE MARCADORES

Código: SPM
Revisión: 2.0
Página 5 de 38

característicos de cada uno de los atributos positivos/negativos comparando los diferentes grupos de materiales de referencia, cuya descripción se encuentra en el Anexo 1.

Este estudio se lleva a cabo en el momento en que se estudia la homogeneidad de los materiales de referencia. Para el estudio de homogeneidad se analiza la fracción de compuestos orgánicos volátiles de 10 lotes de los 9 MRC que presentan 5 atributos diferentes. Además, se inyectan en el equipo dos réplicas instrumentales del primer lote de cada uno de los 9 MRC, por lo que se obtienen un total de 12 cromatogramas representativos de cada uno de los MRC.

Para el análisis de las componentes principales (PCAs) se lleva a cabo un pretratamiento de los datos con la función "Medina_14" implantada por el Doctor Santiago Medina en la CMQ de la UGR. Esta función se ejecuta en el entorno Matlab® y a través de la herramienta de PLS_Toolbox se lleva a cabo el estudio del PCA.

De esta manera se elige el modo en el que se va a estudiar la estabilidad (trimestral) de dichos materiales a través de los índices de similitud (mediante el coseno del ángulo y el coeficiente de determinación). De manera que se elige para el estudio de estabilidad: (i) la fracción completa de los compuestos orgánicos volátiles analizados; (ii) una selección de dicha fracción, eliminando la parte final del mismo para disminuir el número total de variables; y/o (iii) seleccionando uno o varios picos cromatográficos característico de cada material de referencia.

4. EQUIPOS

Para el análisis cromatográfico de la fracción de volátiles, se utiliza un cromatógrafo de gases acoplado a un detector convencional de ionización de llama (FID). El equipo dispone de un sistema de preparación de muestra automatizado, capaz de realizar: (1) una microextracción de los componentes volátiles sobre una fibra absorbente sólida (SPME); y (2) la posterior inserción de la fibra en el inyector del cromatógrafo.

Para el tratamiento quimiométrico de los datos, se utilizará el programa Matlab® con la herramienta PLS_Toolbox para la elaboración del PCA.

5. PERSONAL

El análisis cromatográfico para la monitorización de los compuestos orgánicos volátiles de los materiales de referencia así como su tratamiento de datos se ha realizado por personal cualificado del equipo CMQ

6. RESULTADOS

En este apartado se muestran los "scores" y "loadings" de diferentes grupos de materiales de referencia seleccionados para el estudio de selección de uno/varios pico(s) característico(s) que sirvan como marcadores de algún(os) atributo(s) dentro de los 9 MRC. La selección de dichos grupos de MRC se ha llevado a cabo en función de la composición/mezcla de los diferentes tipos de aceites que se explica en detalle en el Anexo I.

Realmente, la selección de los picos característicos no sirven de marcadores de atributos sensoriales, sirven de marcadores del MRC en estudio. Por ello, existen dos grupos claros a estudiar:

- ↳ Grupo 1. Selección de FRU/1/, BRA/1/, BRA/2/ y BRA/3/.
- ↳ Grupo 2. Selección de FRU/1/, MAH/1/ y MAH/2/.

Con el estudio de estos dos grupos de materiales de referencia y teniendo en cuenta la descripción de los mismos (Anexo I), se seleccionan los picos cromatográficos característicos del material FRU/1, BRA/3 y MAH/2.

En el caso de los materiales: FRU/2/, COC/2/ e HIG/2/ se estudian con respecto al material FRU/1/ por ser un material que no presenta atributos negativos y que es un cromatograma más plano que el FRU/2/, por lo que permite mostrar las diferencias entre los picos cromatográficos que separan y distribuyen ambos materiales al hacer el PCA. Además, y con el objetivo de corroborar la selección de dichos picos cromatográficos característicos de cada uno de los MRC a comparar, se estudian COC/2/ e HIG/2/ con respecto al material FRU/2/.

- ↳ Grupo 3. Selección de FRU/1/ y FRU/2/.
- ↳ Grupo 4. Selección de FRU/1/ y COC/2/. Corroborar con FRU/2/ y COC/2/.
- ↳ Grupo 5. Selección de FRU/1/ e HIG/2/. Corroborar con FRU/2/ e HIG/2/.

Para llevar a cabo el PCA de cada uno de esos grupos, se ha tratado cada grupo individualmente mediante la función "Medina_14" en la que se lleva a cabo los siguientes pasos:

1. Unión de los cromatogramas. Se unen todos los cromatogramas del estudio de homogeneidad (12 cromatogramas de cada MRC), para cada uno de los grupos a estudiar de forma independiente para cada PCA.
2. Selección de la matriz de datos. Con el objetivo de reducir el tamaño de la matriz de datos, se selecciona una parte del cromatograma de la fracción de compuestos orgánicos volátiles comprendida en el tiempo: 1.27 y 5.97 minutos. Del cromatograma completo que llega hasta el minuto 10.
De esta manera, se reduce el tamaño de la matriz de datos de: 5994 a 2820 variables.
3. Suavizado de los datos. Tiene como objetivo eliminar el ruido que pueda presentar la señal cromatográfica.
4. Corrección de la línea base. Con el objetivo de eliminar las variaciones en la línea base que provocan los pequeños cambios en las condiciones experimentales.
5. Alineamiento de los datos. Para llevar a cabo el alineamiento del cromatograma y que los cromatogramas a estudiar coincidan al superponerlos, se utiliza el algoritmo "icoshift", de manera que los datos se desplazan de izquierda a derecha en el tiempo de retención.



Tras el procesado de los datos y mediante la herramienta PLS_Toolbox se lleva a cabo un centrado en la media para fijarse en las variaciones entre objetos (muestras). Se calcula el vector medio de los datos en cada una de las columnas y sustrayéndolo punto a punto de cada vector de esa columna. Después de esto, cada fila de datos incluye cómo las filas difieren de la muestra media en la matriz original de los datos.

6.1. Grupo 1. FRU/1/ y BRA/1/, /2/ y /3/

Al estudiar este grupo, se pueden buscar los picos característicos del material FRU/1/ y BRA/3/, porque el material BRA/3/ es un aceite de oliva que presenta el defecto atrojado/borras sin estar mezclado con ningún otro aceite de oliva virgen. Por tanto, al hacer el PCA de este grupo, se va poder seleccionar y corroborar el(los) pico(s) característico(s) de cada uno de ellos al ser el BRA/2/ y BRA/1/ diluciones de la mezcla: BRA/3/ con FRU/1/.

En este caso, se escogen dos componentes principales, que explican el 98.01% de la varianza.

A continuación se muestran los gráficos de "scores" del PC2 vs PC1 junto con los de "loadings" de cada uno de ellos.

Gráfico de "scores" PC2 vs PC1 y de "loadings" PC1 vs variables

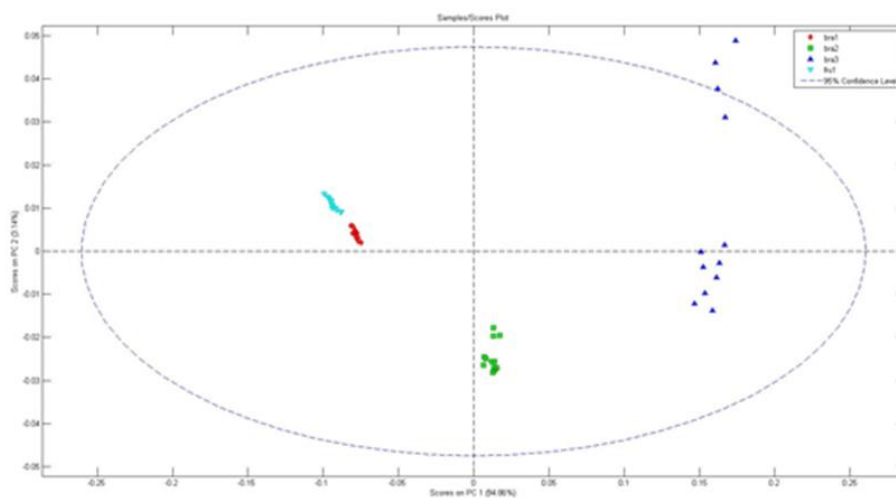
En la Figura 1 se muestran los "scores" (PC2 frente a PC1) y los "loadings" de la PC1.

A través de la primera componente principal se diferencia los materiales que contienen frutado (a la izquierda del gráfico de "scores") del material que contiene el defecto atrojado/borras. Por lo que parece bastante interesante estudiar los "loadings" de esta componente principal. Además con ella se explica la toda la distribución, puesto que presenta una varianza del 94.86%.

La distribución de los "scores" es, de derecha a izquierda: BRA/3/, cuya muestra contiene el 100% de aceite de oliva virgen con defecto atrojado/borras, que no ha sido mezclado con ningún otro tipo de aceite y que es de la variedad picual; va seguido del material BRA/2/, puesto que este material está formado por el material BRA/3/ y FRU/1/ en una proporción 12/88, es por ello que dicho material se encuentra cercano al eje de coordenadas dentro del gráfico de "scores"; a continuación le sigue el BRA/1/ que se encuentra muy próximo al FRU/1/ debido a que contiene una mezcla de BRA/3/ y FRU/1/ con una proporción 2.5/97.5. Al tener tanta cantidad del material FRU/1/ se encuentra muy próximo a él en la distribución de sus componentes.

Por lo tanto, es muy revelador el estudio de la primera componente principal frente a las variables, para así observar el o los picos cromatográficos responsables del atributo positivo frutado y el negativo de atrojado/borras, aunque realmente se puede decir que son picos cromatográficos marcadores del material de referencia FRU/1/ y BRA/3/. Por ello, se estudia el gráfico de "loadings" de la PC1 vs variable, en él se observa que los picos representativos del BRA/3/ son los picos de las variables 466 y 1787, que corresponden con los tiempos de retención aproximados de 2.0 y 4.2 minutos, respectivamente.

"Scores PC2 vs PC1"



"Loadings PC1"

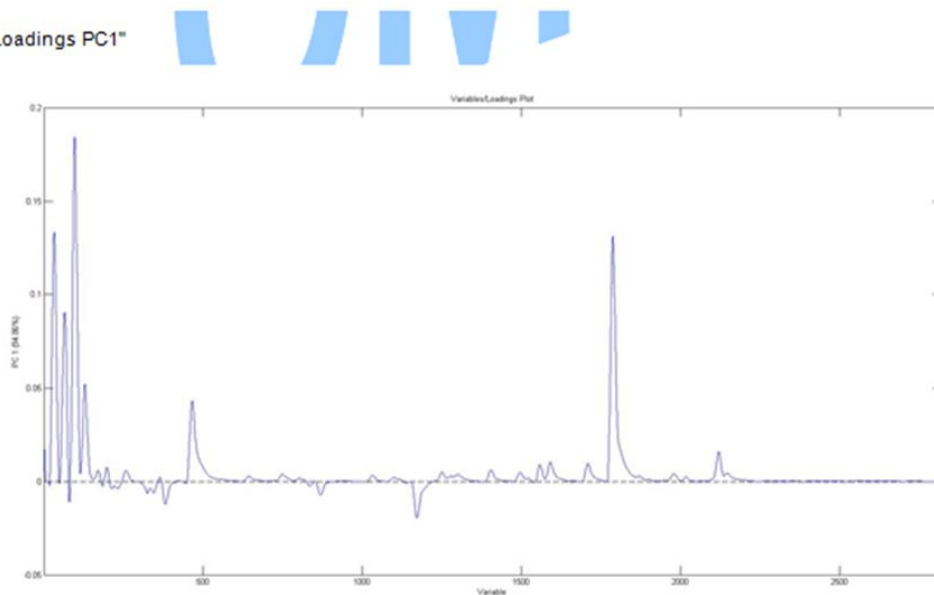


Figura 1. Grupo 1. FRU/1/ y BRA/1/, 1/2/ y 1/3/.

Ambos picos cromatográficos se encuentran en la parte superior (o positiva) del gráfico de "loadings" y por tanto, representa la parte derecha del gráfico de "scores" (parte positiva). El pico representativo del FRU/1/ aparece en la variable 1172, que corresponde aproximadamente con el tiempo de retención a 3.2 minutos. Este pico, aparece en la parte negativa del gráfico de "loadings" puesto que aparece en la parte izquierda del gráfico de "scores" si se observa la PC1 (parte negativa).

Gráfico de "scores" PC2 vs PC1 y de "loadings" PC2 vs variables

En la Figura 2 se muestran los "scores" (PC2 frente a PC1) y los "loadings" de la PC2.

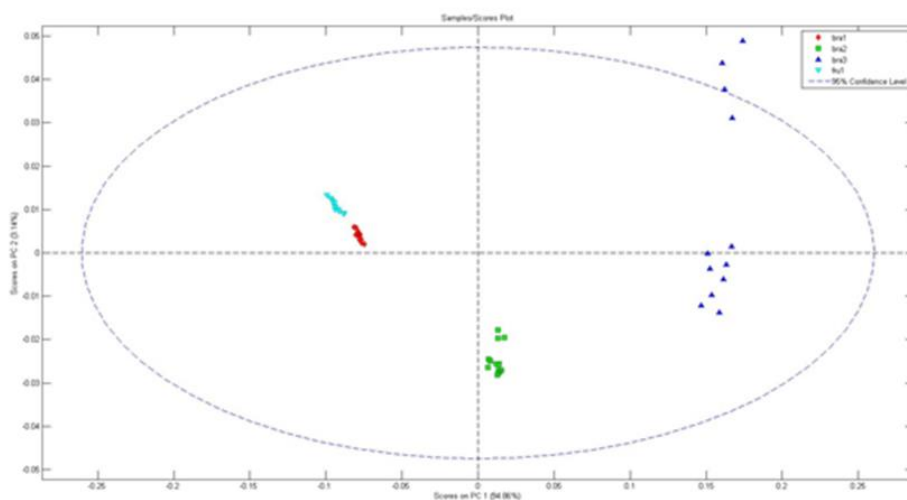
A través de la segunda componente principal no se diferencian los materiales puesto que en la parte de arriba de la gráfica de "scores" (que corresponde con la parte positiva) aparece además del frutado (en la parte izquierda) un grupo de BRA/3/ en el que aparecen cuatro muestras demasiado altas en el análisis de componentes principales debido a que presentan el pico con tiempo de retención alrededor del minuto 4.2 mucho más elevado que el resto de muestras, por lo que no se pueden diferenciar ninguna característica que diferencie los materiales FRU/1/ y BRA/3/.

Las cuatro muestras corresponden con los lotes: BRA/3/142; BRA/3/322; BRA/3/262 y BRA/3/202 (enumeradas desde arriba hacia abajo, desde la parte más alta de la PC2 hacia el centro del PCA). Por lo que se eliminan dichas muestras del análisis para ver cómo se comporta la distribución en el PCA.

Para eliminar dichas muestras se ha llevado a cabo dentro de la herramienta PLS_Toolbox. En este caso, se han escogido dos componentes principales, para crear el modelo, que explican el 98.35% de la varianza. En la Figura3 se muestran los "scores" (PC2 frente a PC1) y los "loadings" de la PC2 para este modelo en el que se han eliminado cuatro muestras del material BRA/3/.

Al estudiar el PCA eliminando estas cuatro muestras, se sigue observando la misma distribución a través de la primera componente principal. Es la segunda componente principal la que varía, aun así, no es posible diferenciar entre materiales, puesto que en la parte de arriba (positiva) de la PC2 aparecen los materiales que contienen el mayor porcentaje de atributo frutado (arriba a la izquierda), pero arriba a la derecha aparece el material BRA/3/, por lo que no es posible estudiar sus diferencias ni el pico característico de cada material a través de esta componente principal. Si se observa el gráfico de "loadings" la la PC2 frente a las variables, se observan dos picos muy intensos que corresponden con las variables 466 y 1787 (en la parte positiva del gráfico, parte de arriba), que coinciden con los que se han seleccionado con los picos del BRA/3/ puesto que este material es el que más peso tiene en este PCA.

"Scores PC2 vs PC1"



"Loadings PC2"

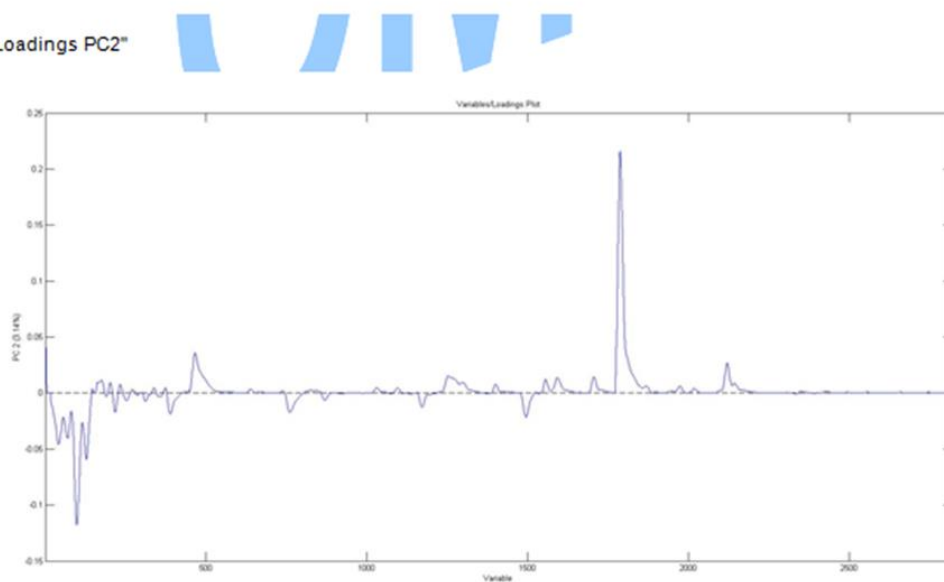
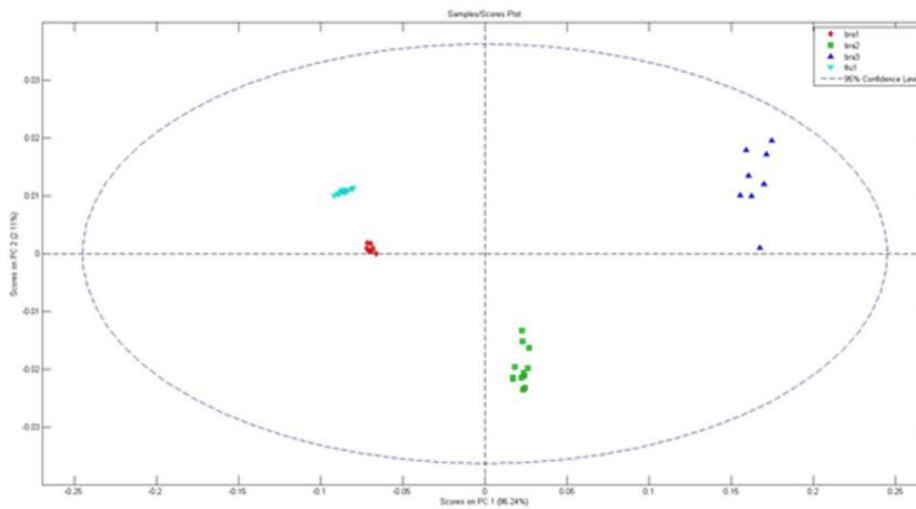


Figura 2. Grupo 1. FRU/1/ y BRA/1/, /2/ y /3/.



"Scores PC2 vs PC1"



"Loadings PC2"

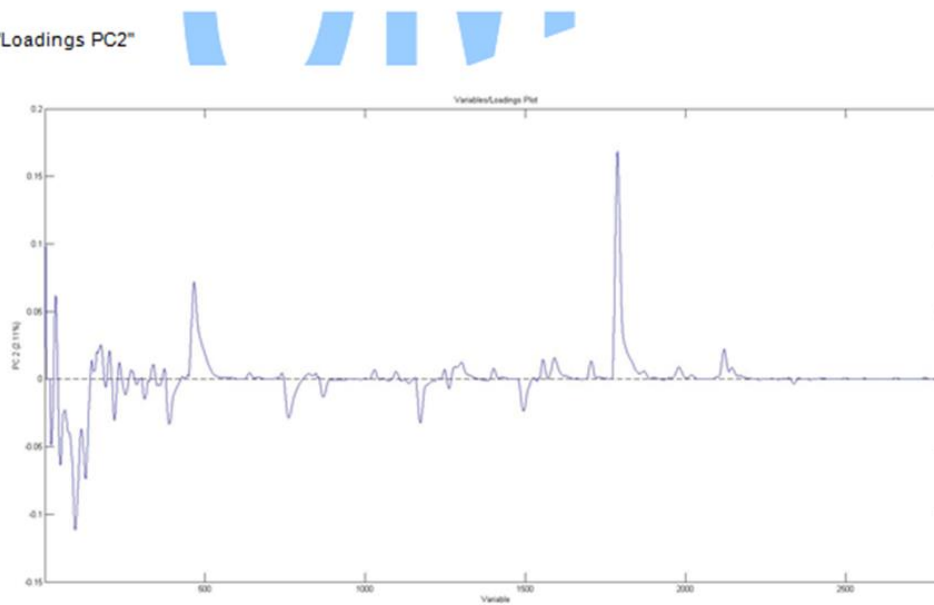


Figura 3. Grupo 1. FRU/1/ y BRA/1/, /2/ y /3/. Eliminando cuatro muestras del material BRA/3/.



6.2. Grupo 2. FRU/1/ y MAH/1/ y /2/

Al hacer el tratamiento de los datos de este grupo se pasa a estudiar los componentes principales. En este caso, se escogen dos componentes principales, que explican el 99.61% de la varianza.

A continuación se muestran los "scores" y "loadings" de las dos componentes principales.

Gráfico de "scores" PC2 vs PC1 y de "loadings" PC1 vs variables

En la Figura 4 se muestran los "scores" (PC2 frente a PC1) y los "loadings" de la PC1.

A través de la primera componente principal se diferencia los materiales que contienen frutado (a la derecha del gráfico de "scores") del material que contiene el defecto madera húmeda. Por lo que parece bastante interesante estudiar los "loadings" de esta componente principal. Además con ella se explica la toda la distribución, puesto que presenta una varianza del 99.33%.

La distribución de los "scores" es, de izquierda a derecha: MAH/2/, cuya muestra contiene aceite de oliva virgen con el defecto madera húmeda, que no ha sido mezclado con ningún otro tipo de aceite y que es de la variedad picual procedente de aceitunas que han sufrido heladas; va seguido del material MAH/1/, puesto que este material es una mezcla del material MAH/2/ y FRU/1/ en una proporción 40/60, es por ello que dicho material se encuentra cercano al eje de coordenadas dentro del gráfico de "scores" y en una zona más cercana al grupo del FRU/1/ puesto que contiene un 60% de mezcla de este material.

Por lo tanto, es muy interesante estudiar la primera componente principal frente a las variables, para así seleccionar los picos cromatográficos responsables del atributo positivo frutado y el negativo de madera húmeda, o al menos los picos cromatográficos marcadores del material de referencia FRU/1/ y MAH/2/. Por ello, se estudia el gráfico de "loadings" de la PC1 vs variable, en él se observa que los picos representativos de la MAH/2/ son los picos de las variables 764 y 863, que corresponden con los tiempos de retención aproximados de 2.5 y 2.7 minutos, respectivamente. Puesto que ambos picos se encuentran en la parte inferior (o negativa) de la gráfica de "loadings" y por tanto, representa la parte izquierda del gráfico de "scores". El pico representativo del FRU/1/ aparece en la variable 1172, que corresponde aproximadamente con el tiempo de retención a 3.2 minutos y que coincide con el pico cromatográfico detectado como FRU/1/ en el grupo anterior en el que se comparan los distintos tipos de BRA y FRU/1/.

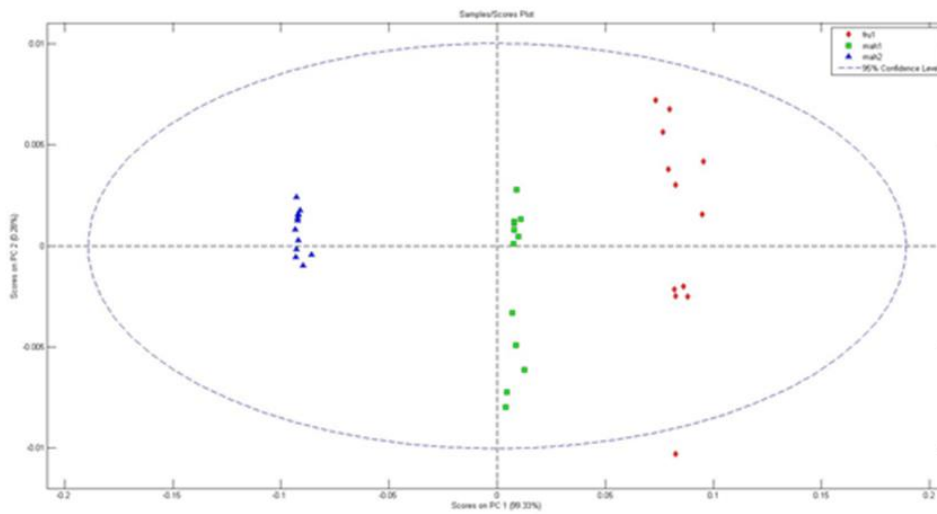
Gráfico de "scores" PC2 vs PC1 y de "loadings" PC2 vs variables

En la Figura 5 se muestran los "scores" (PC2 frente a PC1) y los "loadings" de la PC2.

A través de la segunda componente principal no se diferencian los materiales puesto que en la parte de arriba de la gráfica de "scores" (que corresponde con la parte positiva) aparece además del frutado (en la parte derecha) un grupo de MAH/2/ que se encuentra en la parte izquierda y superior del gráfico de "scores". Aun así, aparece un pico cromatográfico en la variable 1496.



"Scores PC2 vs PC1"



Loadings PC1"

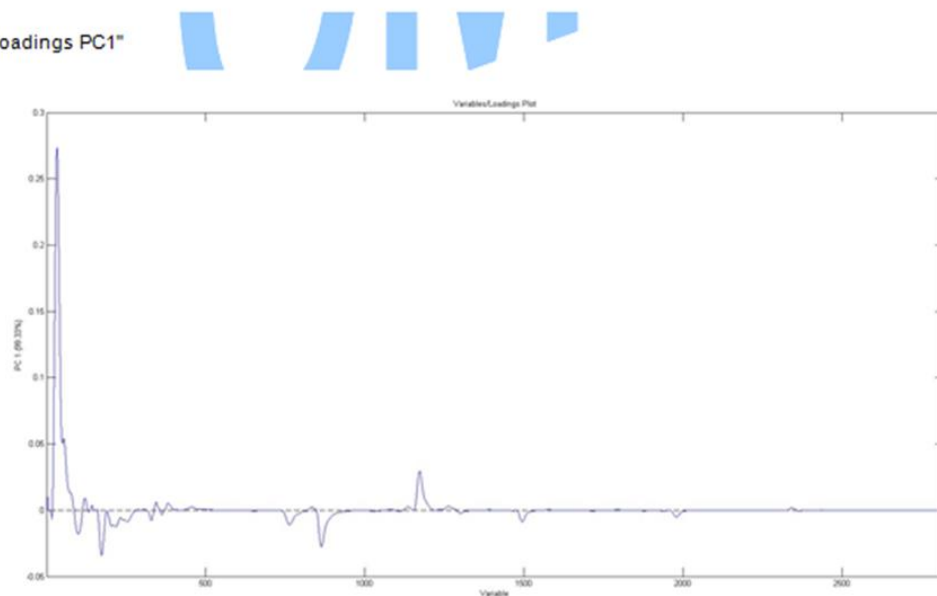
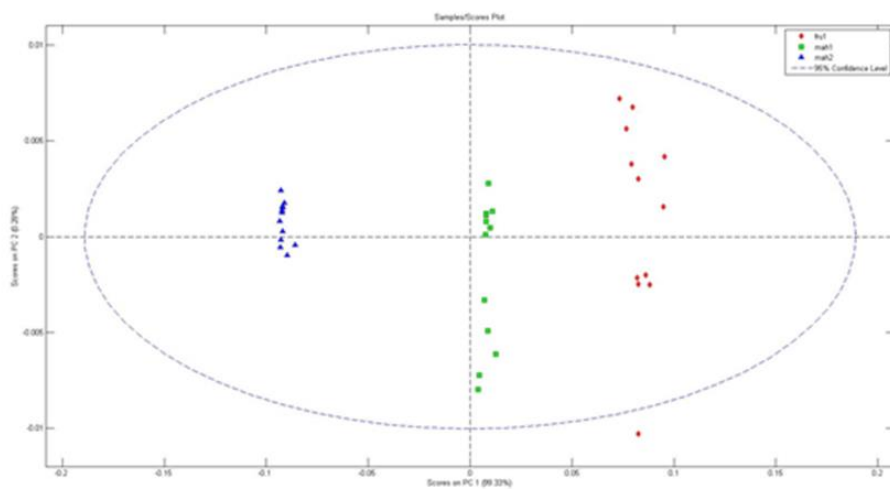


Figura 4. Grupo 2. FRU/1/ y MAH/1/ y /2/.



"Scores PC2 vs PC1"



"Loadings PC2"

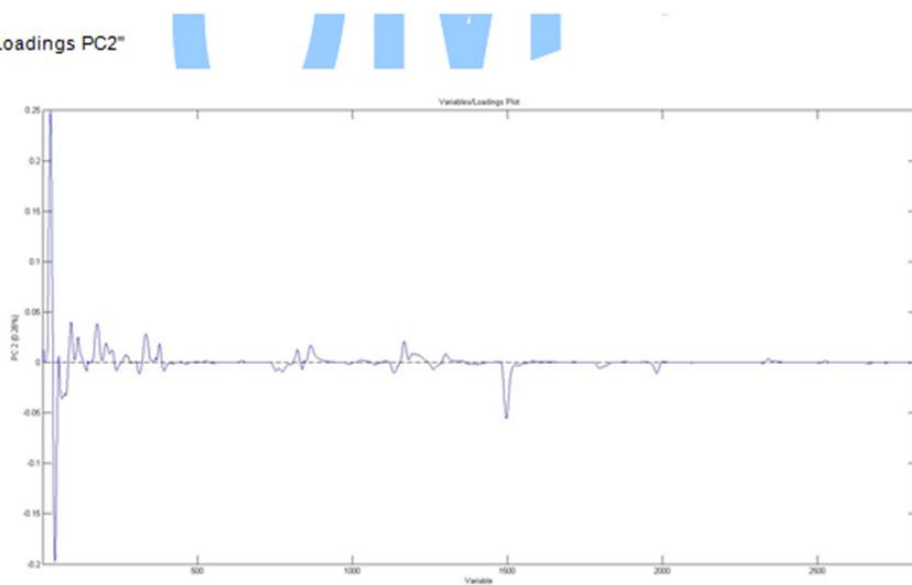


Figura 5. Grupo 2. FRU/1/ y MAH/1/ y /2/.

6.3. Grupo 3. FRU/1/ y FRU/2/

Se hace el PCA con estos materiales, con el objetivo de corroborar el pico cromatográfico característico del material FRU/1/, así como de estudiar el(los) pico(s) característicos del material FRU/2/.

Para este grupo se seleccionan dos componentes principales que explican el 99.73% de la varianza.

Gráfico de "scores" PC2 vs PC1 y de "loadings" PC1 vs variables

En la Figura 6 se muestra el gráfico de "scores" de PC2 frente a PC1 y el gráfico de "loadings" en el que se representa la PC1 frente a las variables.

Como se observa en el gráfico de "scores", a partir de la primera componente principal se explica todo el modelo y en ella se muestran dos divisiones, a la derecha aparece el material FRU/1/ y a la izquierda el material FRU/2/, por lo que a través del gráfico de "loadings" de esta componente principal, se pueden seleccionar los picos característicos de cada uno de ellos. Para el material FRU/1/, que se encuentra en la parte positiva de la PC1 (a la derecha) aparece el pico 1171, que aparece en la parte positiva de dicho gráfico. Este pico cromatográfico corrobora lo que se ha observado en el resto de grupos en el que se ha comparado este material.

Para el material FRU/2/ se podrían seleccionar los picos cromatográficos que aparecen en las variables 234, 328 y 381, que corresponden con los tiempos de retención a: 1.6, 1.8 y 1.9 minutos, respectivamente.

Gráfico de "scores" PC2 vs PC1 y de "loadings" PC2 vs variables

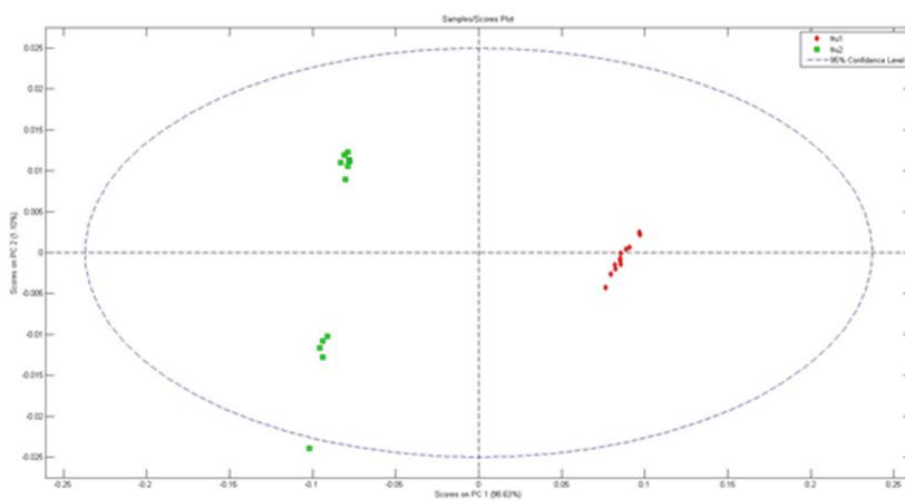
En la Figura 7 se muestran el gráfico de "scores" y "loadings" del PC2. Con esta componente no se diferencian los MRC de este grupo en estudio, debido a que el material FRU/2/ aparece tanto en la parte positiva, como en la parte negativa de dicha componente.

Aparecen cinco muestras en la parte negativa (parte de abajo) a la izquierda del material FRU/2/, cuyos lotes son: FRU/2/052 (y sus dos réplicas instrumentales), FRU/2/262 y FRU/2/292. Al eliminar estas cinco muestras todo el modelo se explica con una única componente principal, puesto que esta explica el 99.61% de la varianza. Por tanto, este grupo de muestras se explica únicamente con la primera componente principal, tal y como se ha explicado anteriormente.

En la Figura 8, se muestran el gráfico de "scores" y de "loadings" de esta segunda componente principal al eliminar esas cinco muestras.



"Scores PC2 vs PC1"



"Loadings PC1"

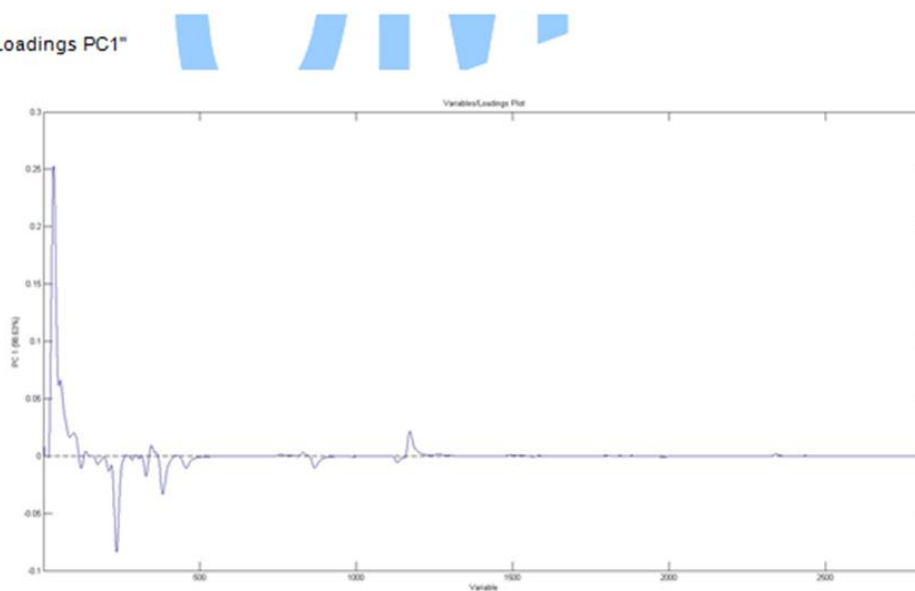
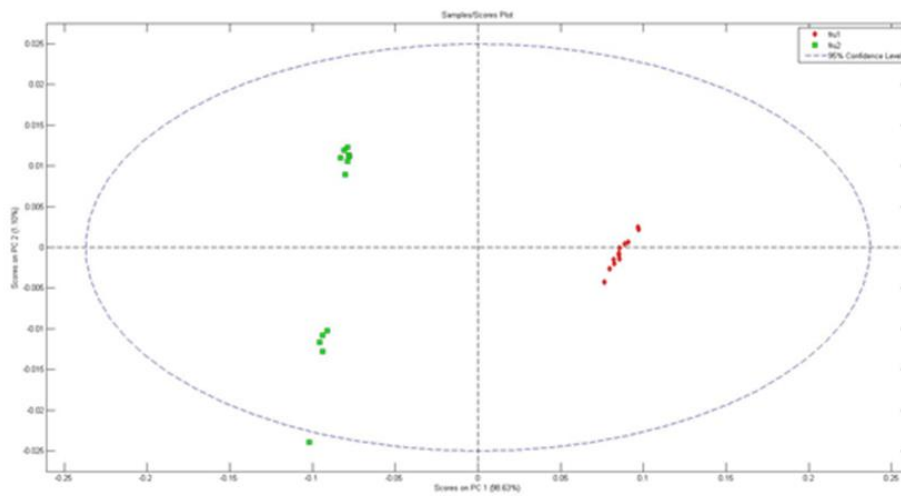


Figura 6. Grupo 3. FRU/1/ y FRU/2/.

"Scores PC2 vs PC1"



"Loadings PC2"

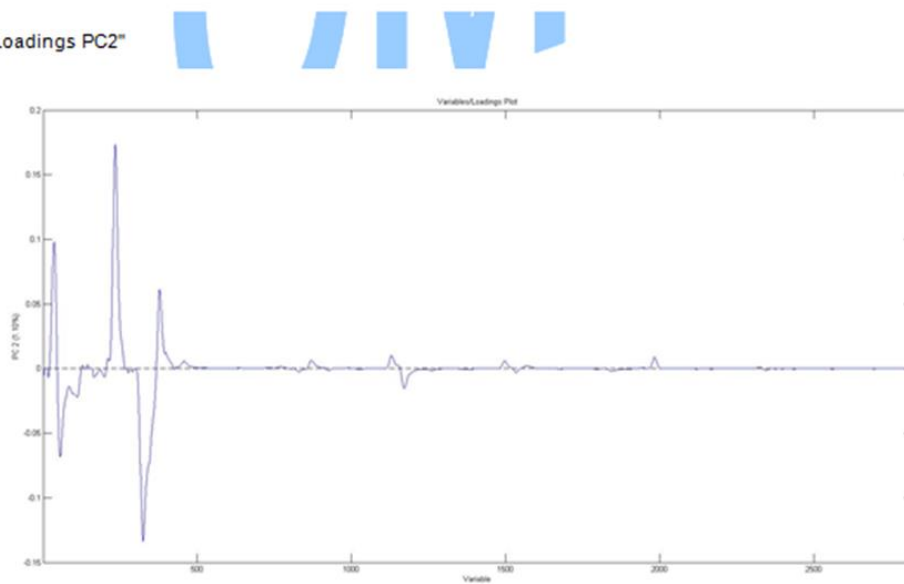
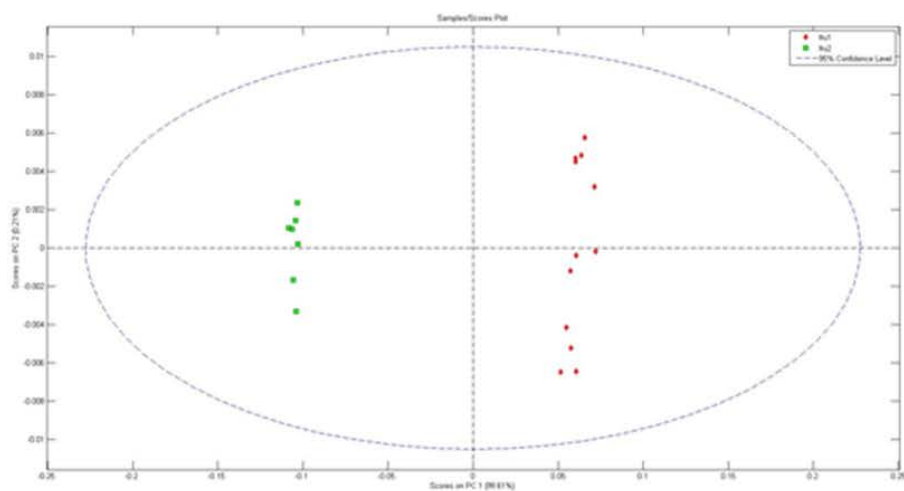


Figura 7. Grupo 3. FRU/1/ y FRU/2/.

"Scores PC2 vs PC1"



"Loadings PC2"

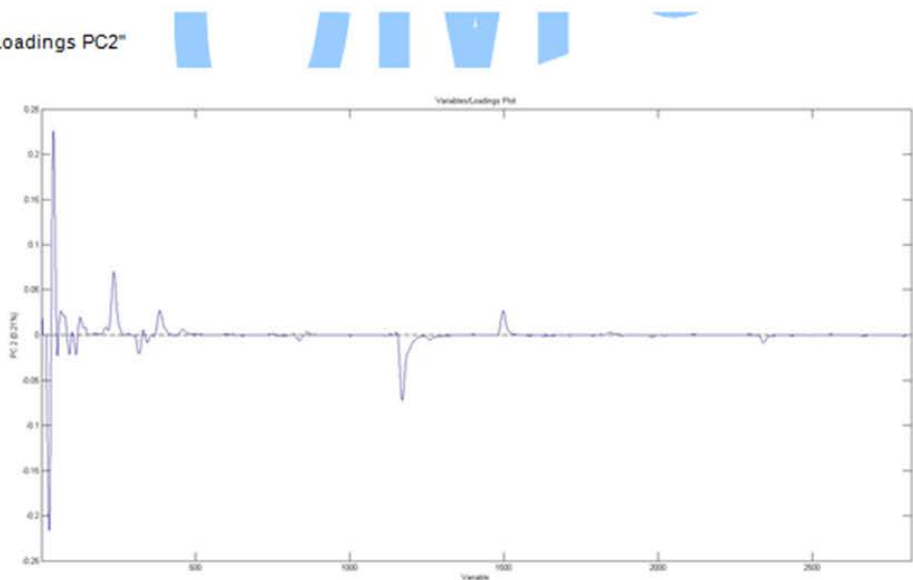


Figura 8. Grupo 3. FRU/1/ y FRU/2/. Eliminando cinco muestras del material FRU/2/.



6.4. Grupo 4. FRU/1/ y COC/2/. Corroborar con FRU/2/ y COC/2/

Se hace el PCA con los materiales FRU/1/ y COC/2/ con el objetivo de corroborar el pico cromatográfico característico del material FRU/1/, así como de estudiar el(los) pico(s) característicos del material COC/2/. De la misma forma, se hace con los materiales FRU/2/ y COC/2/.

Grupo FRU/1/ y COC/2/

Para este grupo se seleccionan dos componentes principales que explican el 99.92% de la varianza. Toda la varianza queda explicada con la primera componente principal (99.04%), por lo que únicamente se muestra el gráfico de "scores" y "loadings" para esta componente, que se muestra en la Figura 9.

A través de la primera componente, quedan separados los materiales FRU/1/ a la izquierda (parte negativa) y COC/2/ a la derecha (parte positiva). En este caso, para el COC/2/ su pico cromatográfico característico es el que aparece en la variable 33, que corresponde con el tiempo de retención a 1.3 minutos. Al estudiar los cromatogramas de dicho material con respecto a FRU/1/, se observa claramente este pico cromatográfico mucho más elevado que en el FRU/1/ y el resto del cromatograma es bastante plano.

Para el caso del material FRU/1/ vuelve a aparecer como pico característico el pico que aparece en la variable 1171.

Grupo FRU/2/ y COC/2/

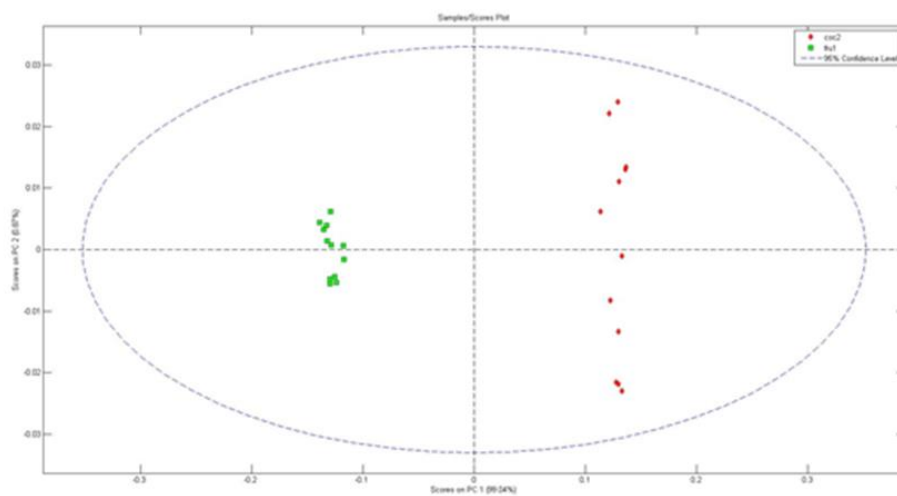
Para este grupo se seleccionan dos componentes principales que explican el 99.83% de la varianza. Toda la varianza queda explicada con la primera componente principal, que explica el 99.52% de la varianza. Por este motivo, únicamente se explica el gráfico de los "scores" PC2 frente a PC1 y los "loadings" de la PC1 para estudiar el(los) pico(s) cromatográficos a seleccionar para cada uno de estos materiales.

En la Figura 10 aparecen los gráficos de "scores" y "loadings" de la PC1.

A través de la primera componente principal se separan los materiales FRU/2/, que aparece a la izquierda (parte negativa) y COC/2/, que aparece en la parte positiva de dicha componente (parte derecha). En este caso, para el COC/2/ se corrobora el pico característico que se ha elegido al estudiar los materiales FRU/1/ y COC/2/.

Para el material FRU/2/ aparecen tres picos cromatográficos característicos, que corresponden con las variables: 234, 326 y 380. Que corrobora lo que se observó al comparar los materiales FRU/1/ y FRU/2/ en el grupo 3 del presente documento.

"Scores PC2 vs PC1"



"Loadings PC1"

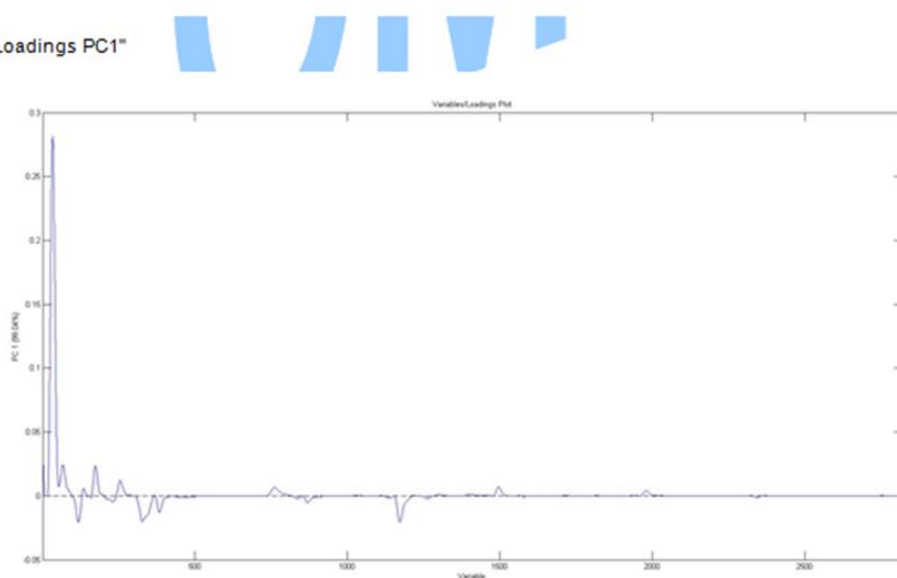
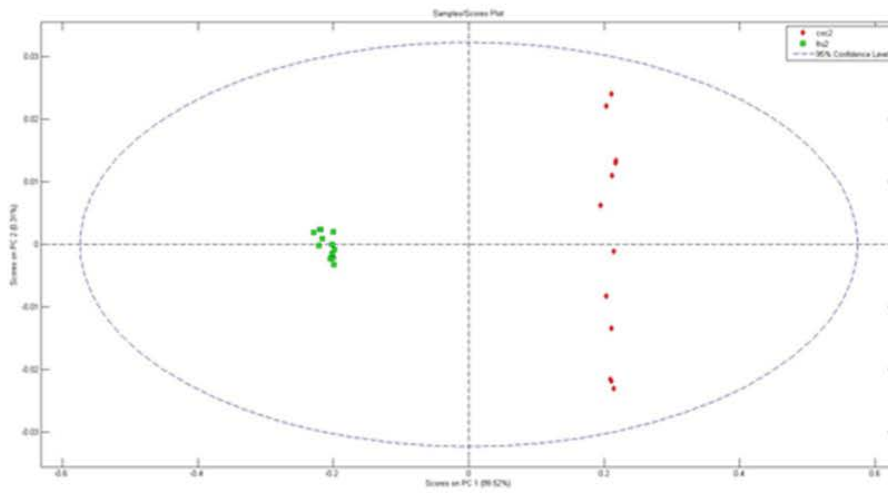


Figura 9. Grupo 4. FRU/1/ y COC/2/.

"Scores PC2 vs PC1"



"Loadings PC1"

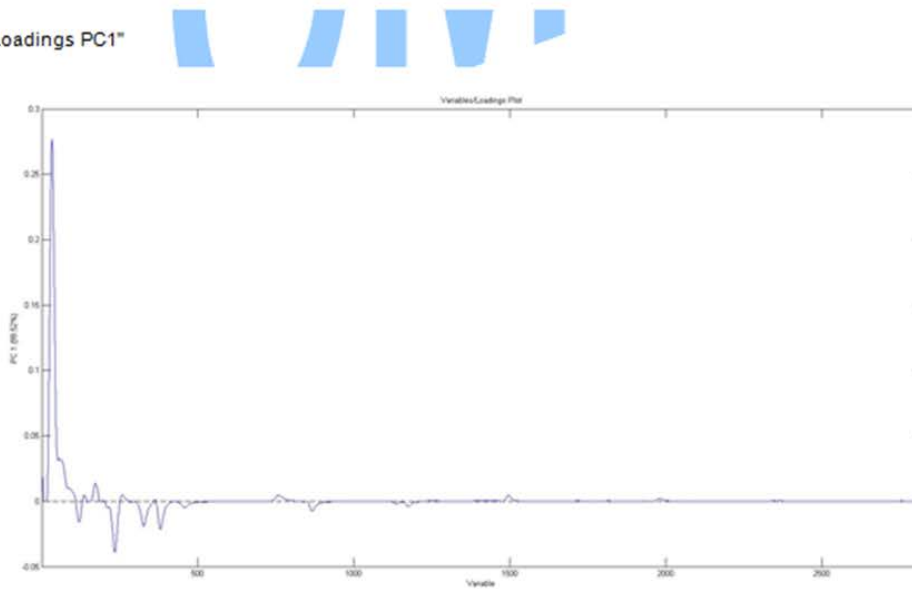


Figura 10. Grupo 4. FRU/2/ y COC/2/. Para corroborar los picos cromatográficos seleccionados.

6.5. Grupo 5. FRU/1/ e HIG/2/. Corroborar con FRU/2/ e HIG/2/

Se hace el PCA con los materiales FRU/1/ e HIG/2/ con el objetivo de corroborar el pico cromatográfico característico del material FRU/1/, así como de estudiar el(los) pico(s) característicos del material HIG/2/. De la misma forma, se hace con los materiales FRU/2/ e HIG/2/.

Grupo FRU/1/ e HIG/2/

Para este grupo se seleccionan dos componentes principales que explican el 99.80% de la varianza. Toda la varianza queda explicada con la primera componente principal (99.56%), por lo que únicamente se muestra el gráfico de "scores" y "loadings" para esta componente, que se muestra en la Figura 11.

A través de la primera componente, quedan separados los materiales FRU/1/ a la derecha (parte positiva) e HIG/2/ a la izquierda (parte negativa). En este caso, para la HIG/2/ su pico cromatográfico característico es el que aparece en la variable 1495, que corresponde con el tiempo de retención a los 3.7 minutos.

Para el caso del material FRU/1/ vuelve a aparecer como pico característico el pico que aparece en la variable 1171. Además, se destaca otro pico característico, el que aparece en la variable 326, pero este no va a ser tenido en cuenta como característico de este material, puesto que sólo se ha puesto presente al hacer este PCA y ya aparece el pico cromatográfico en la variable 328 para el FRU/2/, por lo que únicamente se considera el pico cromatográfico de la variable 1171 como característico del FRU/1/.

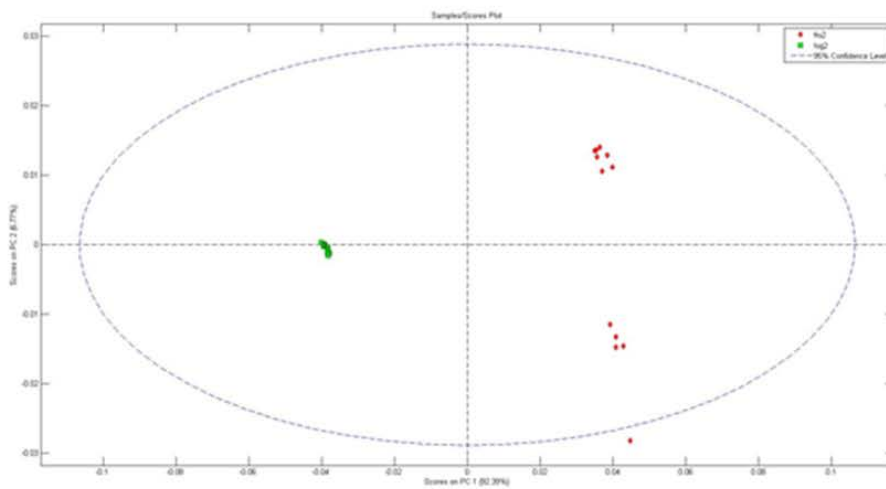
Grupo FRU/2/ e HIG/2/

Para este grupo se seleccionan dos componentes principales que explican el 99.16% de la varianza. Toda la varianza queda explicada con la primera componente principal, que explica el 92.39% de la varianza. Por este motivo, únicamente se explica el gráfico de los "scores" PC2 frente a PC1 y los "loadings" de la PC1 para estudiar el(los) pico(s) cromatográficos a seleccionar para cada uno de estos materiales.

En la Figura 11 aparecen los gráficos de "scores" y "loadings" de la PC1.

A través de la primera componente principal se separan los materiales FRU/2/, que aparece a la derecha (parte positiva) e HIG/2/, que aparece en la parte negativa de dicha componente (parte izquierda). En este caso, para la HIG/2/ se corrobora el pico característico que se ha elegido a estudiar los materiales FRU/1/ e HIG/2/. Además, aparecen otros dos picos característicos, que corresponden con las variables 757 y 1172, esta última variable es la que hasta ahora se ha relacionado con el material FRU/1/. Tiene sentido, puesto que el material con defecto higuera, corresponde con un material frutado que tiene defecto moho, por lo tanto, no es extraño que aparezca el pico cromatográfico correspondiente con el atributo frutado. Por lo que, estos últimos picos cromatográficos, no se van a considerar como característicos del material HIG/2/.

"Scores PC2 vs PC1"



"Loadings PC1"

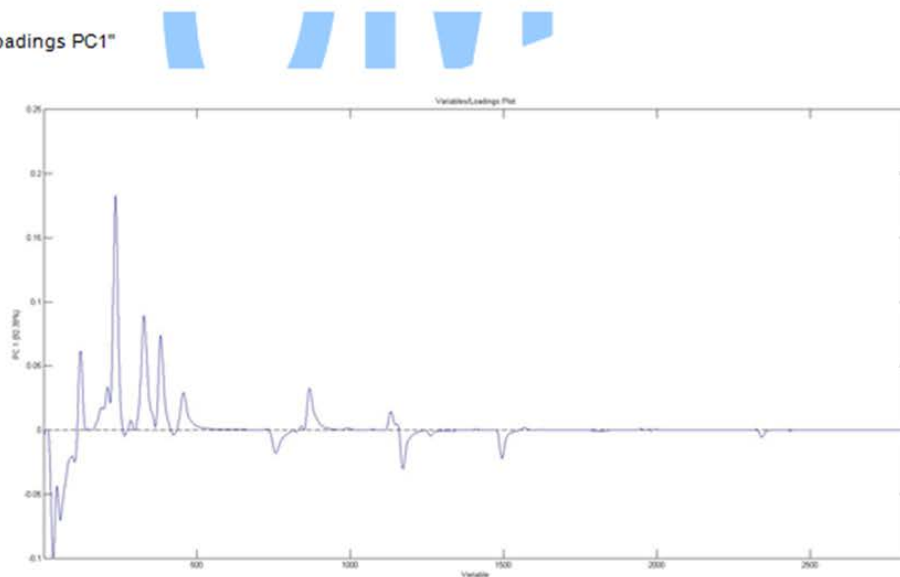
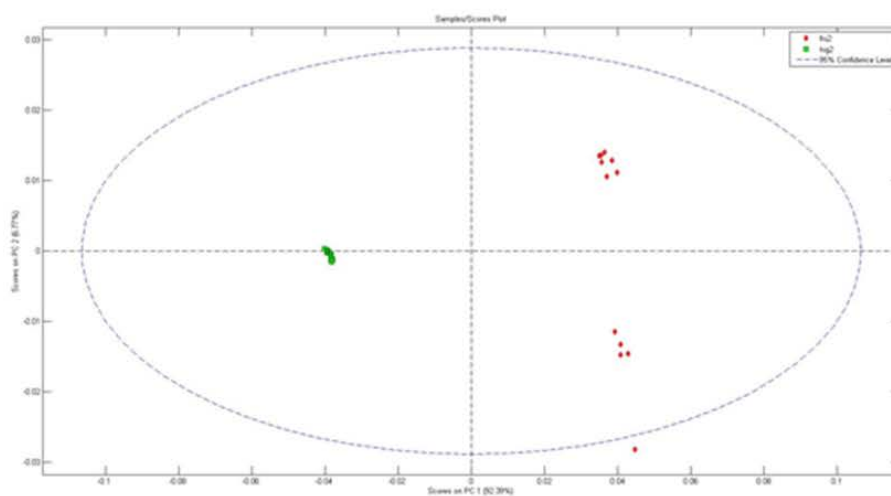


Figura 11. Grupo 4. FRU/1/ e HIG/2/.

"Scores PC2 vs PC1"



"Loadings PC1"

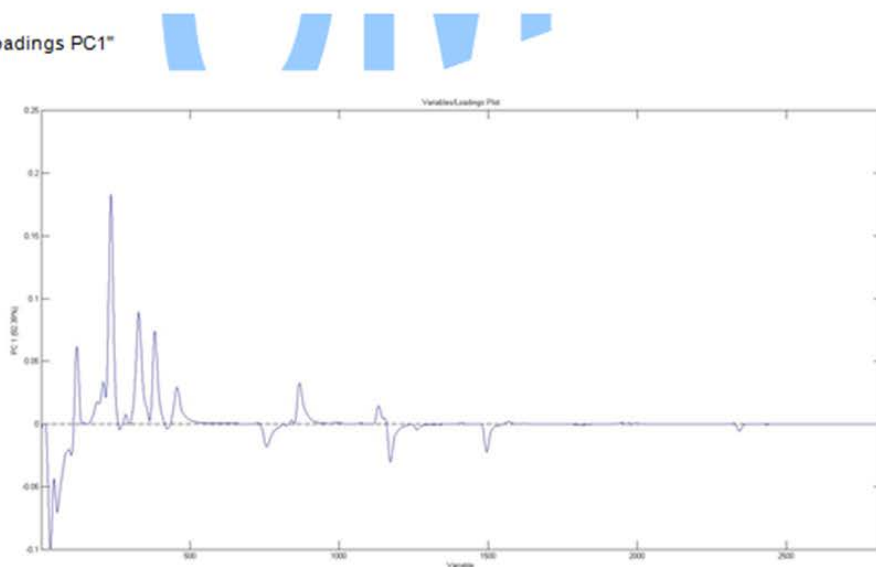
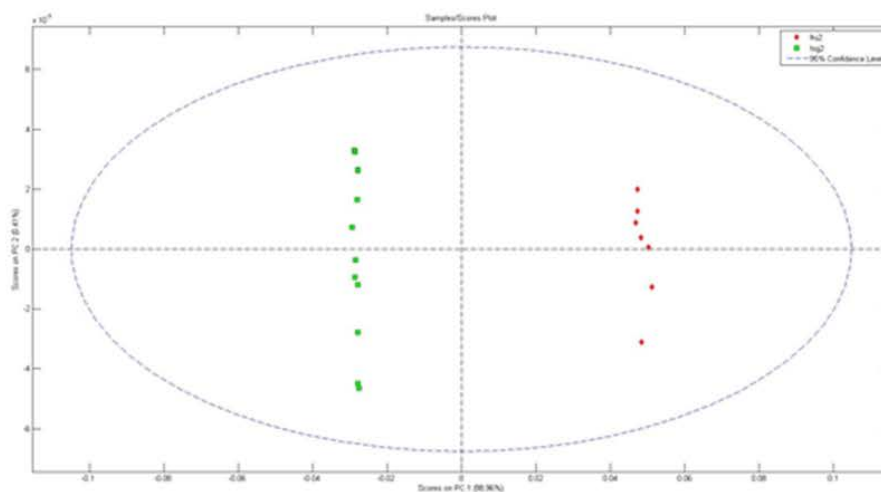


Figura 12. Grupo 4. FRU2/ e HIG2/. Para corroborar los picos cromatográficos seleccionados.

"Scores PC2 vs PC1"



"Loadings PC1"

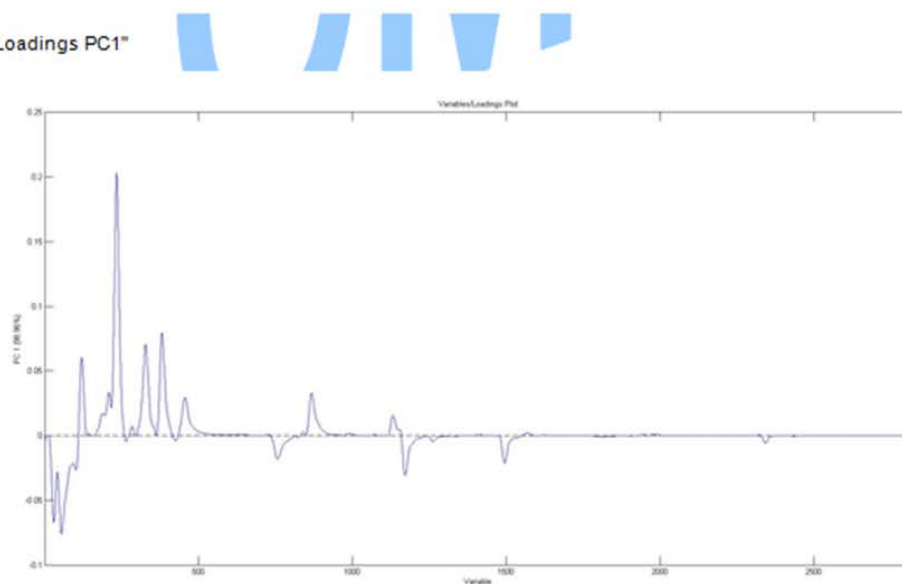


Figura 13. Grupo 4. FRU/2/ e HIG/2/. Para corroborar los picos cromatográficos seleccionados. Eliminando cinco muestras del material FRU/2/.

Para el material FRU/2/ aparecen tres picos cromatográficos característicos, que corresponden con las variables: 234, 327 y 382. Que corrobora lo que se ha observado hasta ahora. Además aparecen otros dos picos cromatográficos más, que corresponden con las variables: 456 y 867. Dado que estas variables no se han detectado al realizar el PCA para el resto de los grupos estudiados, no se incluyen como característicos del material FRU/2/.

En el gráfico de "scores" PC2 frente a PC1 de la Figura 12 se observan cinco muestras del material FRU/2/ que se separan del resto a través de la segunda componente principal. Los códigos del lote de dichas muestras son: FRU/2/052 (y sus dos réplicas instrumentales), FRU/2/262 y FRU/2/292. Corresponde con las mismas muestras que en el grupo 3 donde se hace el PCA con los materiales FRU/1/ y FRU/2/.

Al eliminar dichas muestras lo que ocurre es que cada uno de los grupos de los materiales en estudio, se escogen dos componentes principales, que explican el 99.37% de la varianza, la primera componente principal explica el 98.96%. En la Figura 13 se observa la distribución de ambos materiales de referencia de forma ortogonal a la componente PC1, por lo que no modifica los resultados estudiados hasta ahora.

7. CONCLUSIONES

Con el estudio llevado a cabo, se seleccionan una serie de picos cromatográficos marcadores de cada uno de los MRC, como se muestran en la Tabla 1.

Tabla 1. Picos cromatográficos característicos de cada uno de los MRC que sirven de marcadores de dichos materiales.

Material de referencia	Variable	Tiempo de retención aprox. (min)	Intensidad relativa aprox. (V)
FRU/2/	234	1.6	0.0671
	328	1.8	0.0427
	381	1.9	0.0400
FRU/1/	1172	3.2	0.0228
BRA/3/	465	2.0	0.0528
	1785	4.2	0.1493
MAH/2/	764	2.5	0.0096
	863	2.7	0.0294
COC/2/	33	1.3	0.5584
HIG/2/	1495	3.7	0.0108

Nota: el "arprox." que aparece en las columnas de "tiempo de retención" y de "intensidad relativa", se refiere a valores aproximados, puesto que en función del cromatograma, este tiempo de retención puede variar en función del máximo al que se encuentre el pico cromatográfico (que se indica en la columna de "variable" que puede variar ± 5 unidades).

El tiempo de retención de cada una de las variables, se obtiene al comparar el cromatograma medio de todos los materiales ya alineados (el archivo que se denomina target con la función "Medina_14") con el cromatograma de ese material del lote 052, que es el lote al que se le hace réplica instrumental y que se considera de referencia en el estudio de homogeneidad de los MRC.

La intensidad relativa se obtiene al dividir el valor de intensidad promedio de las 12 muestras de homogeneidad entre la intensidad promedio de las réplicas de los patrones control (3) que se encuentran antes, durante y tras el análisis de cada uno de los MRC.

En la Tabla 2 se han incluido los materiales de referencia diluidos, estos son: BRA/2/, BRA/1/ y MAH/1/ en los que se han seleccionado los picos característicos del MRC del que procede su dilución.

Tabla 2. Picos cromatográficos de los MRC diluidos a partir de cada uno de los MRC que sirven de marcadores de dichos materiales.

Material de referencia	Pico seleccionado	Variable	Intensidad relativa aprox. (V)
BRA/2/	BRA/3/	465	0.0146
		1785	0.0332
	FRU/1/	1172	0.0221
BRA/1/	BRA/3/	465	0.0059
		1785	0.0061
	FRU/1/	1172	0.0262
MAH/1/	MAH/1/	764	0.0062
		863	0.0182
	FRU/1/	1172	0.0174

Nota: el "arprox." que aparece en la columna de "intensidad relativa", se refiere a valores aproximados, puesto que en función del cromatograma, el tiempo de retención puede variar en función del máximo al que se encuentre el pico cromatográfico (que se indica en la columna de "variable" que puede variar ± 5 unidades).

En el Anexo 2 se muestran los cromatogramas del lote 052 de cada uno de los MRC en los que se ha señalado el(los) pico(s) cromatográfico(s) característico(s) para cada uno de ellos. Las imágenes han sido obtenidas a través del software "MS Data Review".

Claramente, se observa una disminución en los picos cromatográficos seleccionados en el caso de que los materiales hayan sido diluidos. Como es el caso del material BRA, en la Figura A2.3 se observan los picos característicos del BRA/3/ y se observa como esos picos cromatográficos van disminuyendo a medida que dicho material se diluye con el material FRU/1/, como se observa en la Figura A.2.4 y 5 para los materiales BRA/2/ y BRA/1/, respectivamente. En el caso del material MAH, en la Figura A2.6 se observan los picos seleccionados y como estos disminuyen en la muestra diluida MAH/1/ con el material FRU/1/, que se muestra en la Figura A2.7.



PROYECTO SENSOLEO-MRC 2015
SELECCIÓN DE PICOS CARACTERÍSTICOS QUE
SIRVAN DE MARCADORES

Código: SPM
Revisión: 2.0

Página 28 de 38

Esto es también observable en la columna de "Intensidad relativa aprox." comparando los materiales de la Tabla 1 y 2, para los MRC diluidos.

Tras llevar a cabo este estudio y seleccionar los picos característicos para cada uno de los MRC se ha decidido llevar a cabo el estudio de estabilidad (trimestal) a través de los índices de similitud (coseno del ángulo y coeficiente de correlación) de: (i) una zona del cromatograma comprendida entre el minuto 1.27 y 5.97, con el objetivo de disminuir el número de variables y (ii) la selección del(o los) pico(s) característico(s). Además se va a monitorizar la intensidad relativa de dichos picos, de manera que se compruebe si esta varía en el tiempo en el que se lleva a cabo el estudio.

CMQ



ANEXO 1:

Descripción de los materiales de referencia de aceite de oliva.

Código Material	Descripción
FRU/2/	Aceite de oliva obtenido de aceitunas de vuelo maduras, de la variedad hojiblanca, con intensidad media de frutado.
FRU/1/	Aceite de oliva obtenido de aceitunas de vuelo maduras, de la variedad hojiblanca, con intensidad baja de frutado.
BRA/3/	Aceite de oliva picual, con una intensidad alta del defecto atrojado/borras y media del defecto avinado, que ha sufrido un proceso extensivo de fermentación.
BRA/2/	Aceite de oliva, con una intensidad alta/media del defecto atrojado/borras, obtenido por dilución del material BRA/3/ con el material FRU/1/, en una proporción 12/88.
BRA/1/	Aceite de oliva, con una intensidad media del defecto atrojado/borras, obtenido por dilución del material BRA/3/ con el material FRU/1/, en una proporción 2,5/97,5.
MAH/2/	Aceite de oliva picual con defecto aceituna helada obtenido de aceitunas que han sufrido una helada.
MAH/1/	Aceite de oliva, con una intensidad alta/media del defecto aceituna helada, obtenido por dilución del material MAH/2/ con el material FRU/1/, en una proporción 40/60.
COC/2/	Aceite de oliva hojiblanca que ha sufrido un excesivo calentamiento durante la obtención del aceite, durante el termo-batido de la pasta, y que incorpora el defecto de cocido o quemado. Complementariamente, contiene también el defecto de avinado/avinagrado.
HIG/2/	Aceite de oliva picual con frutado maduro persistente que se asigna a defecto tierra.

ANEXO 2:

Cromatogramas de los materiales de referencia de aceite de oliva.

En este Anexo se presentan los cromatogramas de los 9 materiales de referencia certificados con lote 052 que han sido analizados para el estudio de homogeneidad de los mismos. En cada una de las figuras, se indica el(los) pico(s) cromatográfico(s) seleccionado(s) como marcador de dicho material.

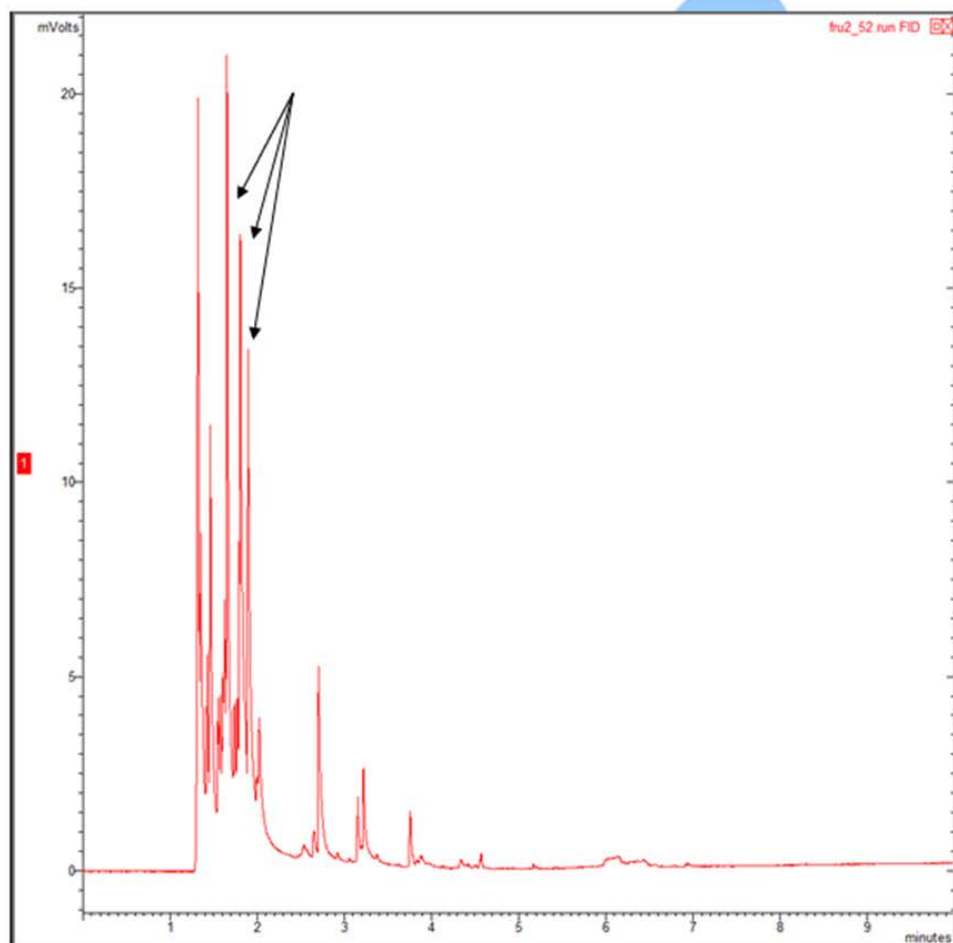


Figura A2.1. Cromatograma FRU/2/052



PROYECTO SENSOLEO-MRC 2015
SELECCIÓN DE PICOS CARACTERÍSTICOS QUE
SIRVAN DE MARCADORES

Código: SPM
Revisión: 2.0
Página 31 de 38

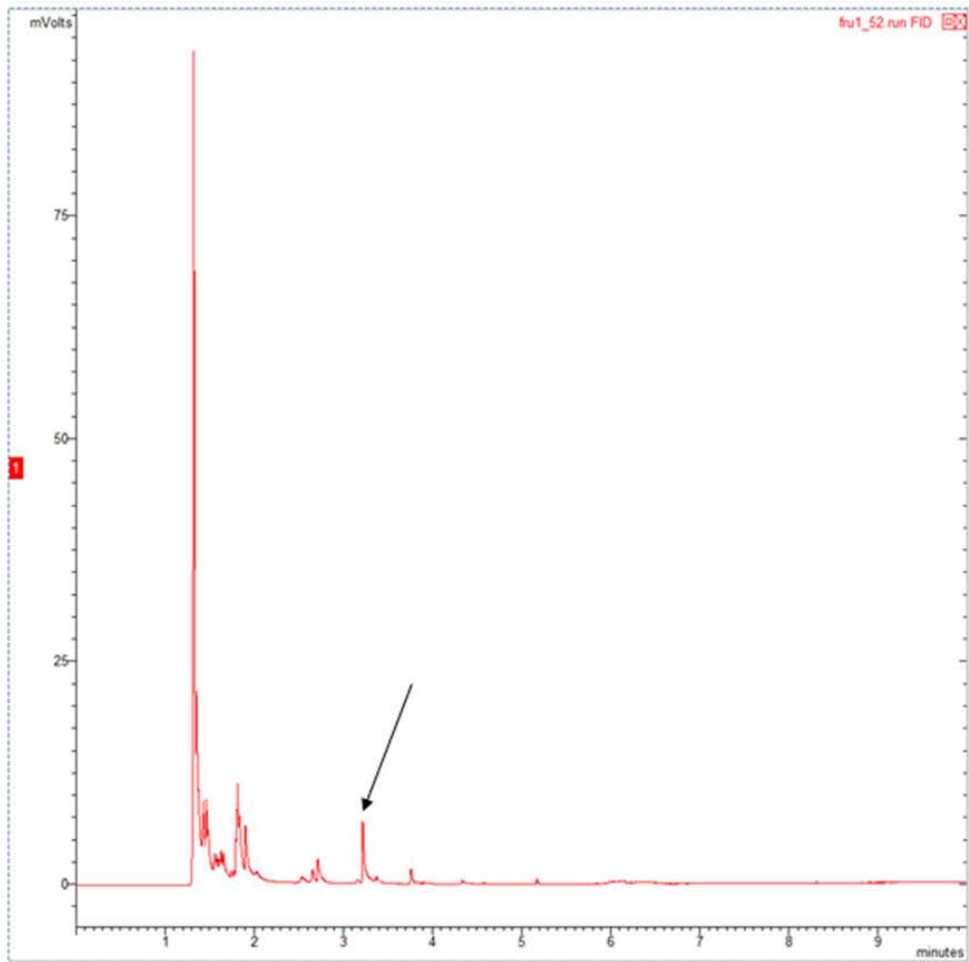


Figura A2.2. Cromatograma FRU/1/052

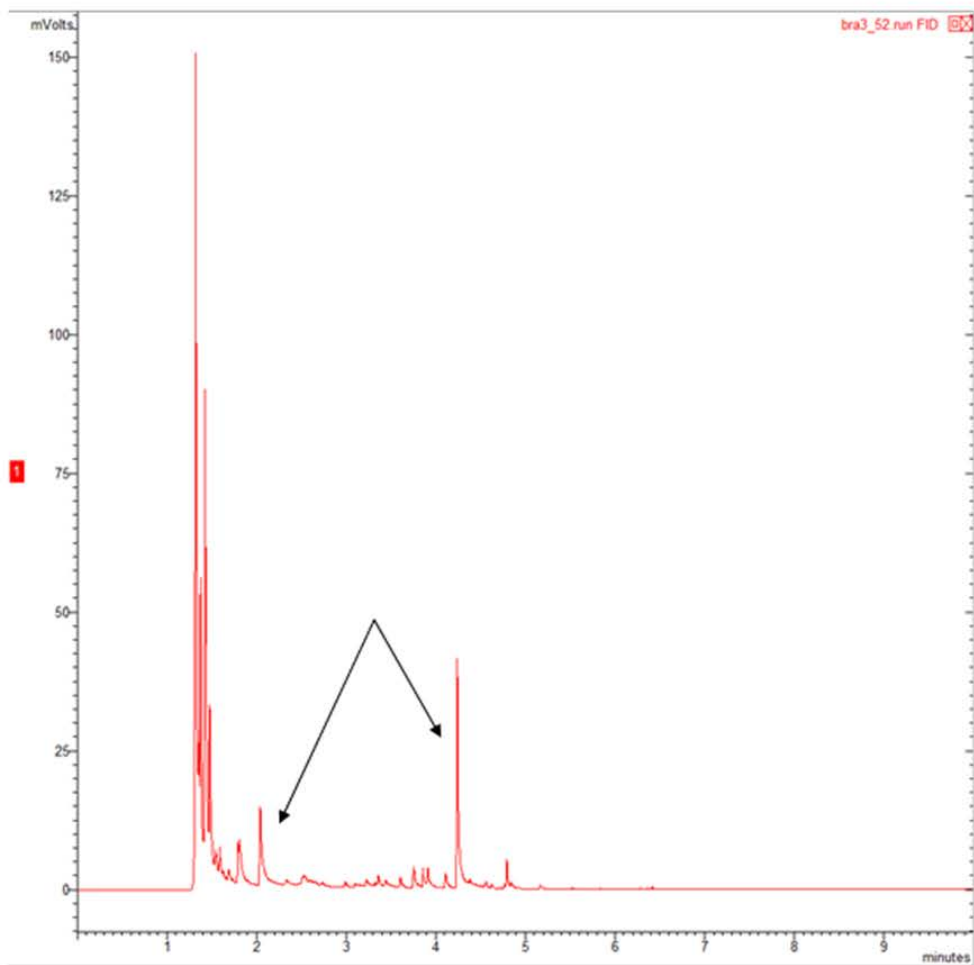


Figura A2.3. Cromatograma BRA/3/052



PROYECTO SENSOLEO-MRC 2015
SELECCIÓN DE PICOS CARACTERÍSTICOS QUE
SIRVAN DE MARCADORES

Código: SPM
Revisión: 2.0
Página 33 de 38

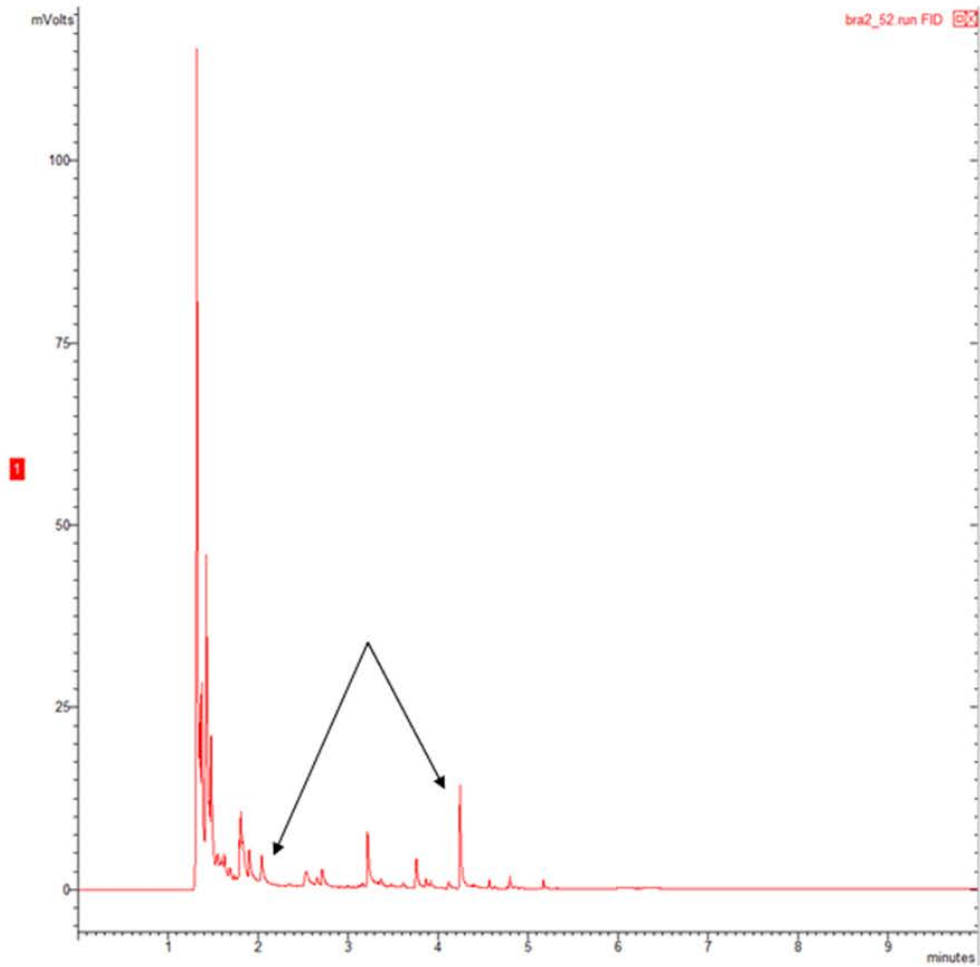


Figura A2.4. Cromatograma BRA/2/052

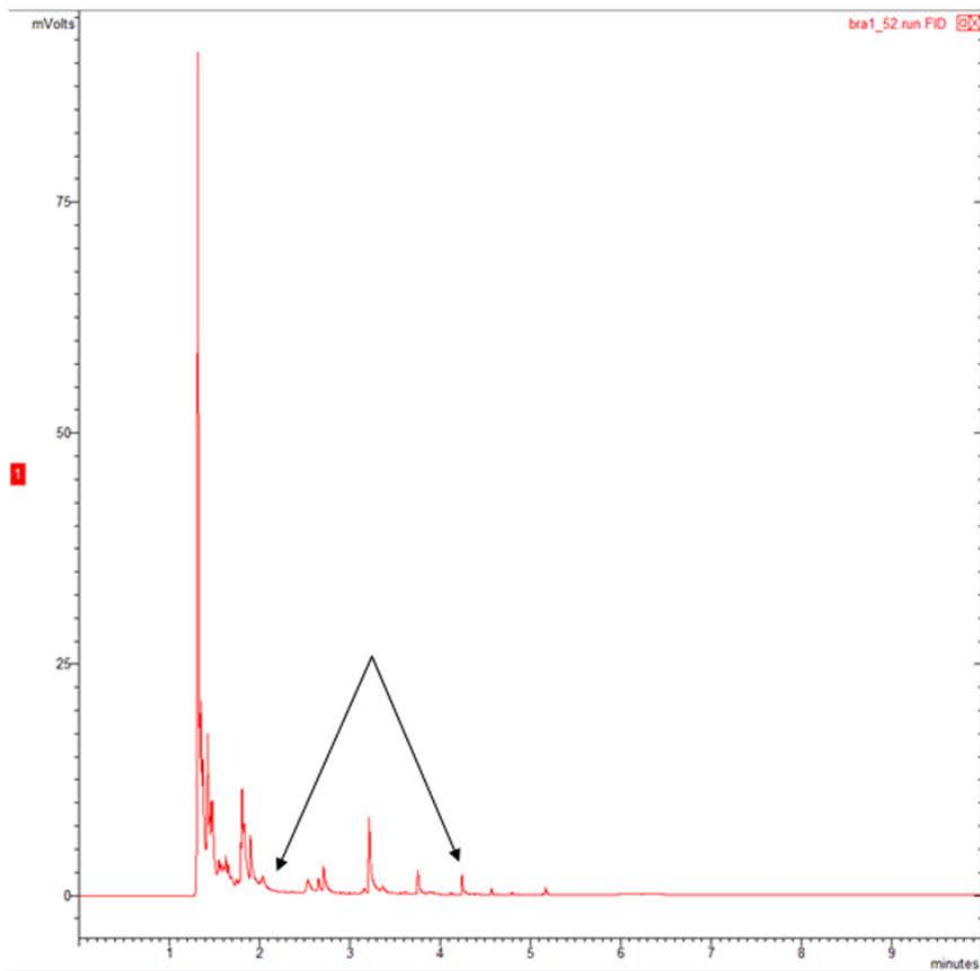


Figura A2.5. Cromatograma BRA/1/052



PROYECTO SENSELEO-MRC 2015
SELECCIÓN DE PICOS CARACTERÍSTICOS QUE
SIRVAN DE MARCADORES

Código: SPM
Revisión: 2.0

Página 35 de 38

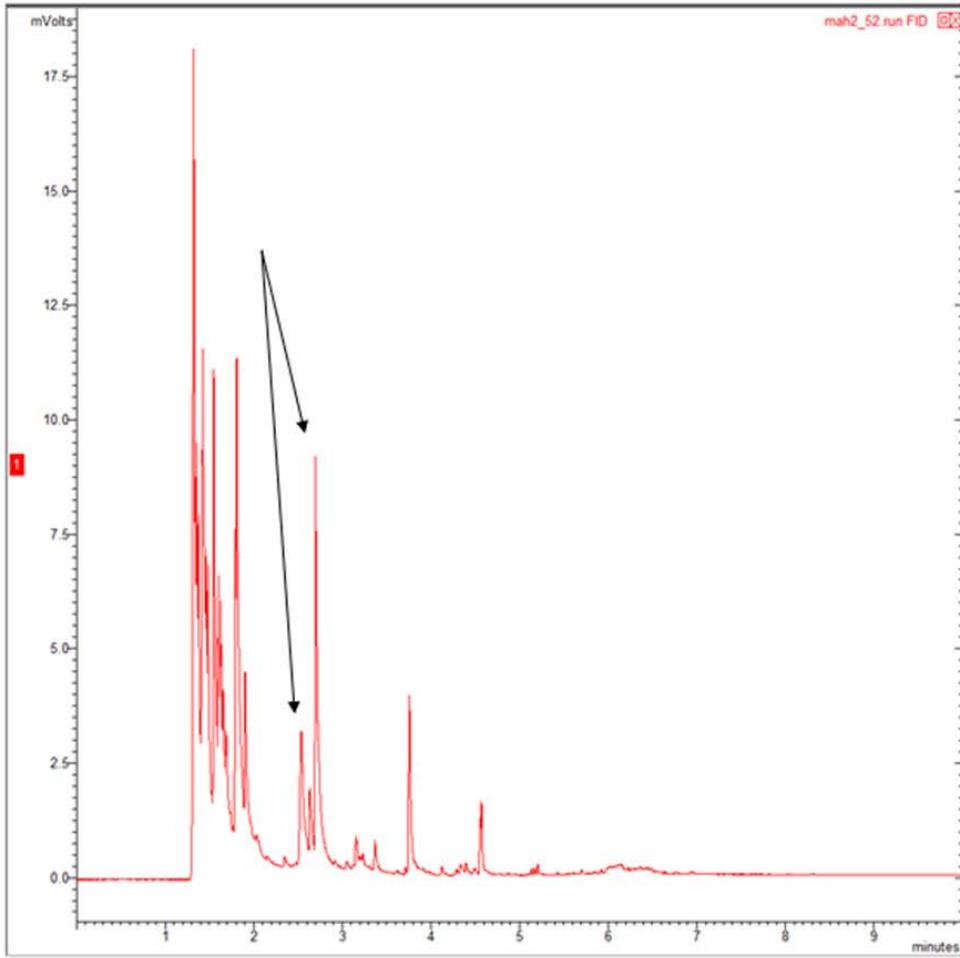


Figura A2.6. Cromatograma MAH/2/052



PROYECTO SENSOLEO-MRC 2015
SELECCIÓN DE PICOS CARACTERÍSTICOS QUE
SIRVAN DE MARCADORES

Código: SPM Revisión: 2.0
Página 36 de 38

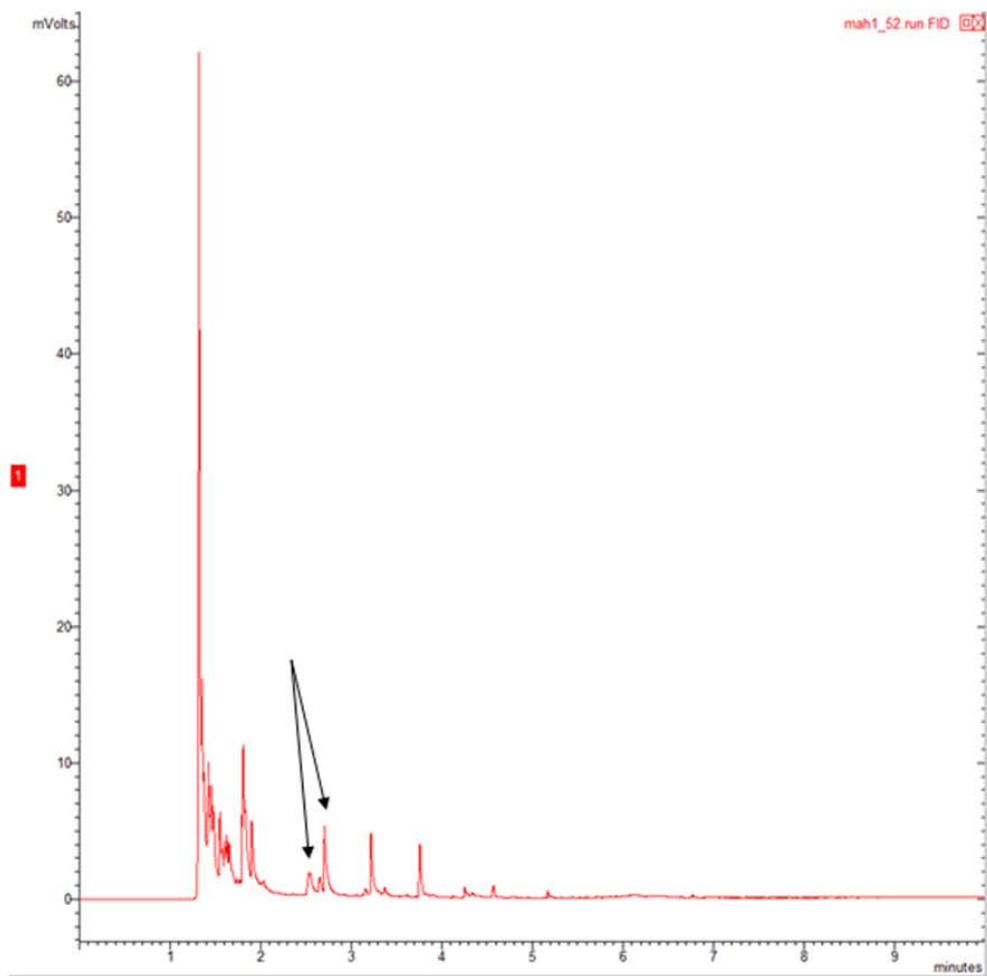


Figura A2.7. Cromatograma MAH/1/052



PROYECTO SENSOLEO-MRC 2015
SELECCIÓN DE PICOS CARACTERÍSTICOS QUE
SIRVAN DE MARCADORES

Código: SPM Revisión: 2.0
Página 37 de 38

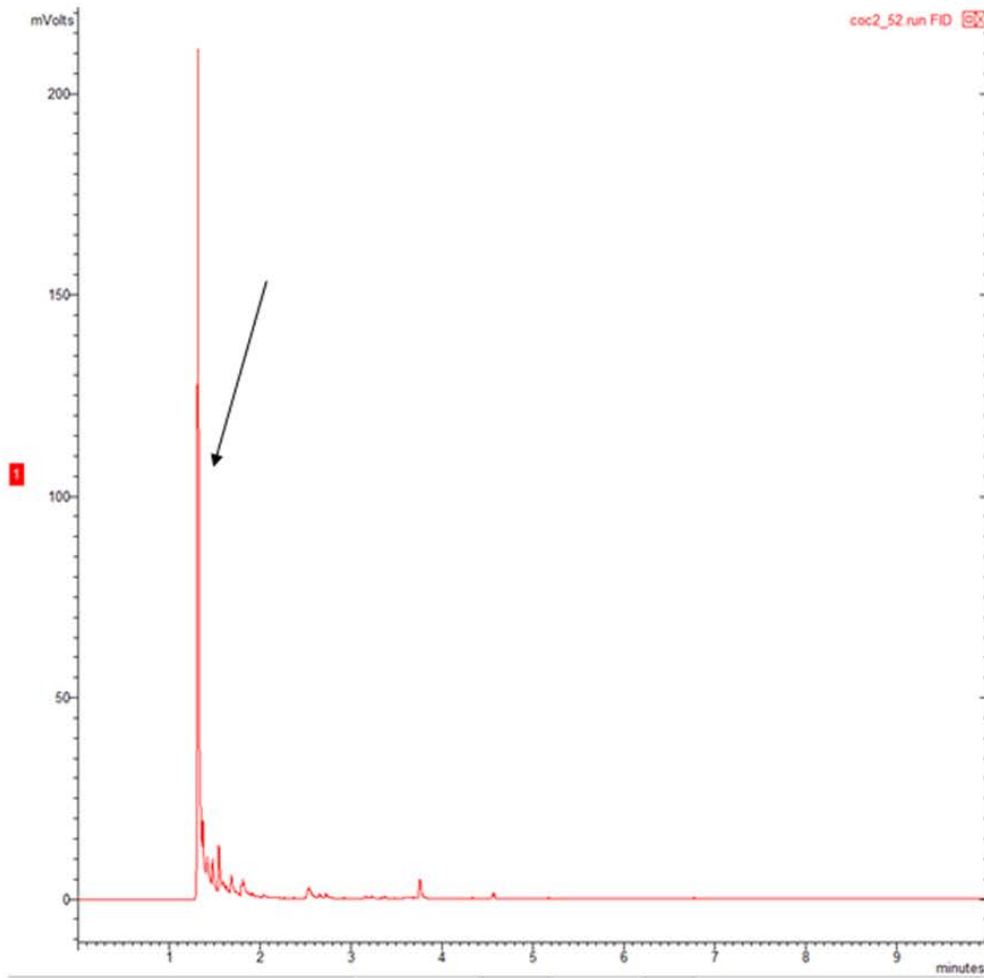


Figura A2.8. Cromatograma COC/2/052



PROYECTO SENSOLEO-MRC 2015
SELECCIÓN DE PICOS CARACTERÍSTICOS QUE
SIRVAN DE MARCADORES

Código: SPM
Revisión: 2.0
Página 38 de 38

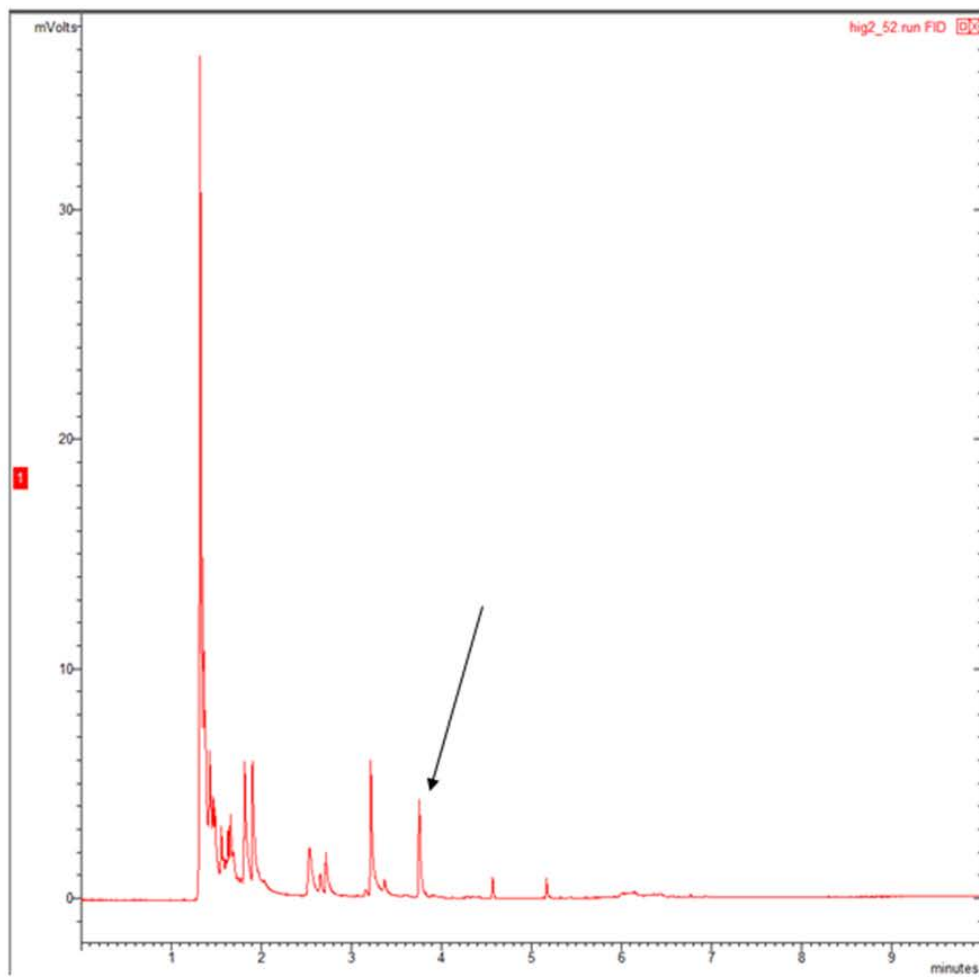


Figura A2.9. Cromatograma HIG/2/052

II.3. MONITORIZACIÓN DE LA ESTABILIDAD

En este ANEXO se muestran cada uno de los gráficos de control obtenidos, para cada uno de los índices de similitud contemplados en la investigación, con el objetivo de monitorizar la estabilidad organoléptica de cada material de referencia analizado.

En todos ellos aparecen cuatro puntos indicativos de los cuatro meses en los que se realizó el análisis: diciembre 2015 (mes 3), marzo 2016 (mes 6), junio 2016 (mes 9) y septiembre 2016 (mes 12).

Siempre aparecen en los gráficos de control un cambio en la tendencia de los datos a partir del mes 9, que genera dos grupos, correspondiente a los meses 3 y 6, con una tendencia diferente a los meses 9 y 12. Este cambio en la monitorización de la señal se debe a que en el mes 9 se hizo un cambio en la columna cromatográfica. Por ello la importancia en los estudios de estabilidad mantener el equipo con las menores alteraciones posibles con respecto a las condiciones iniciales, de referencia.

El límite de aceptación para todos los gráficos de control se selecciona bajo un criterio determinista teniendo en cuenta un estudio que se realizó al principio de la Tesis Doctoral con otros materiales de referencia certificados de aceite de oliva.

Los índices de similitud: $d(\text{eucl})$, T^2 de Hotelling, Q^2 -residual y QM, el valor cero es el que indica que los dos vectores de datos comparados son iguales. Por lo tanto, el eje "x" del gráfico de control empieza en el valor cero y los datos obtenidos de cada muestra para indicar que el material permanece estable se deberían de encontrar por debajo de la línea roja (límite de aceptación).

En el caso de los índices de similitud: NEAR, NEAR (calculado a partir del área de pico seleccionado como marcador del material), R^2 y COS, el valor uno es el que indica que los dos vectores de datos comparados son iguales. Por lo que el eje "x" del gráfico de control empieza en el valor uno. Por lo tanto, y para considerar el material estable, los valores obtenidos para cada muestra han de encontrarse entre ese valor uno y la línea roja (límite de aceptación).

Si comentamos cada una de los gráficos de control en función del índice de similitud usado para la monitorización de la estabilidad:

Para la distancia Euclídea. En todos los casos, salvo en los materiales FRU/1/ y BRA/3/ se observan que los dos primeros datos (mes 3 y mes 6) indican la estabilidad de la muestra y los dos últimos, después del cambio de columna en el cromatógrafo de gases (meses 9 y 12) aparecen por encima del límite de aceptación. Aun así, no podemos decir que el material no sea estable, puesto que no se observa una tendencia en los datos y el valor del mes 9 aparece más alto que el mes 12.

El gráfico en el que se representa el NEAR del cromatograma completo. Todas las muestras se pueden decir que son estables, puesto que el valor del índice de similitud se encuentra dentro del límite de aceptación. Salvo para los materiales FRU/2/, donde no se puede decir que haya perdido estabilidad, puesto que se va fuera sólo un dato (mes 9) y no se muestra una tendencia que indique la pérdida de estabilidad. Así como el material MAH/2/, en el que sería necesario analizar una muestra más, en el mes 13 para corroborar si existe una tendencia que indique la pérdida de estabilidad. Pero esta investigación se contempló durante 12 meses de duración, por lo que no se pudo llevar a cabo más análisis.

Cuando se lleva a cabo una reducción de variables y se hace el cálculo del T^2 de Hotelling y el Q^2 -residual, así como el QM, obtenidos a partir del uso de una herramienta quimiométrica de PCA que agrupa las muestras en función de la similitud entre las señales cromatográficas, se observa claramente la agrupación de los dos primeros meses, con respecto al cambio de columna en los dos últimos meses.

Finalmente, los gráficos de R^2 y COS muestran todos los datos obtenidos dentro del límite de aceptación establecido.

MATERIAL FRU/1/

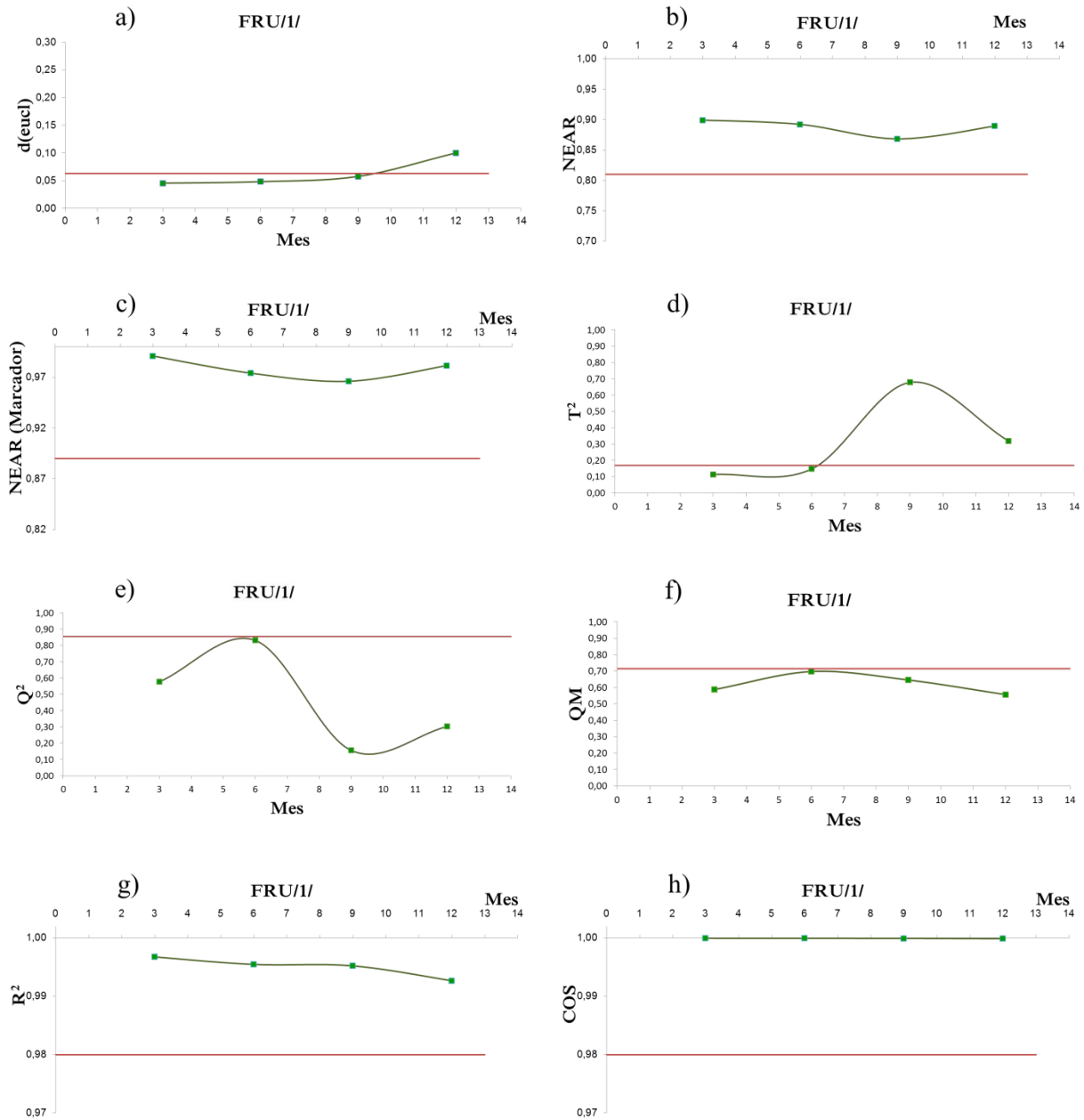


Figura II.3.1. Gráficos de control multivariable con los índices de similitud: a) d(eucl), b) NEAR, c) NEAR (marcadores), d) T²-Hotelling, e) Q²-residual, f) QM, g) R², h) COS; del material FRU/1/.

MATERIAL FRU/2/

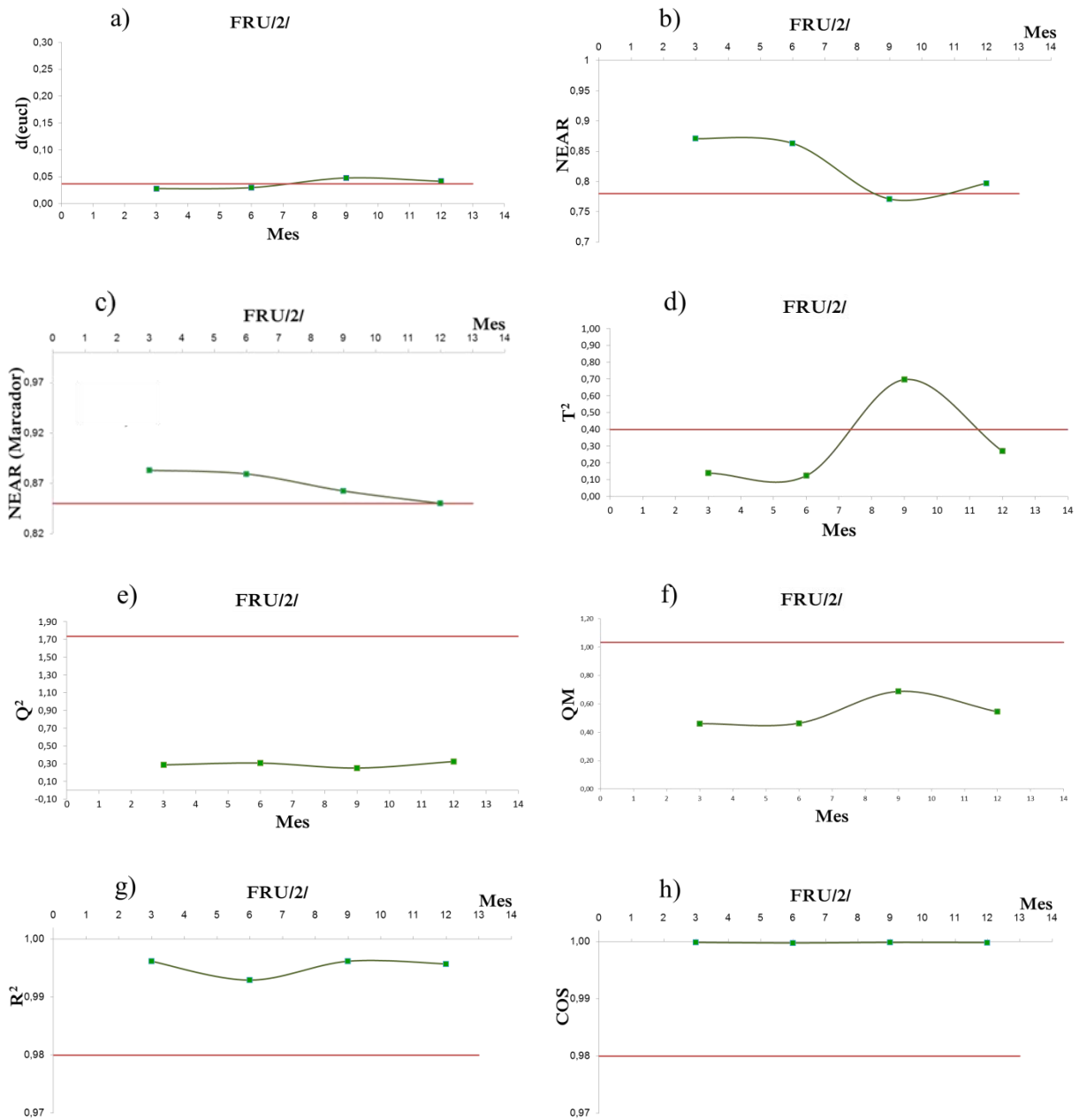


Figura II.3.2. Gráficos de control multivariable con los índices de similitud: a) d(eucl), b) NEAR, c) NEAR (marcadores), d) T²-Hotelling, e) Q²-residual, f) QM, g) R², h) COS; del material FRU/2/.

MATERIAL BRA/1/

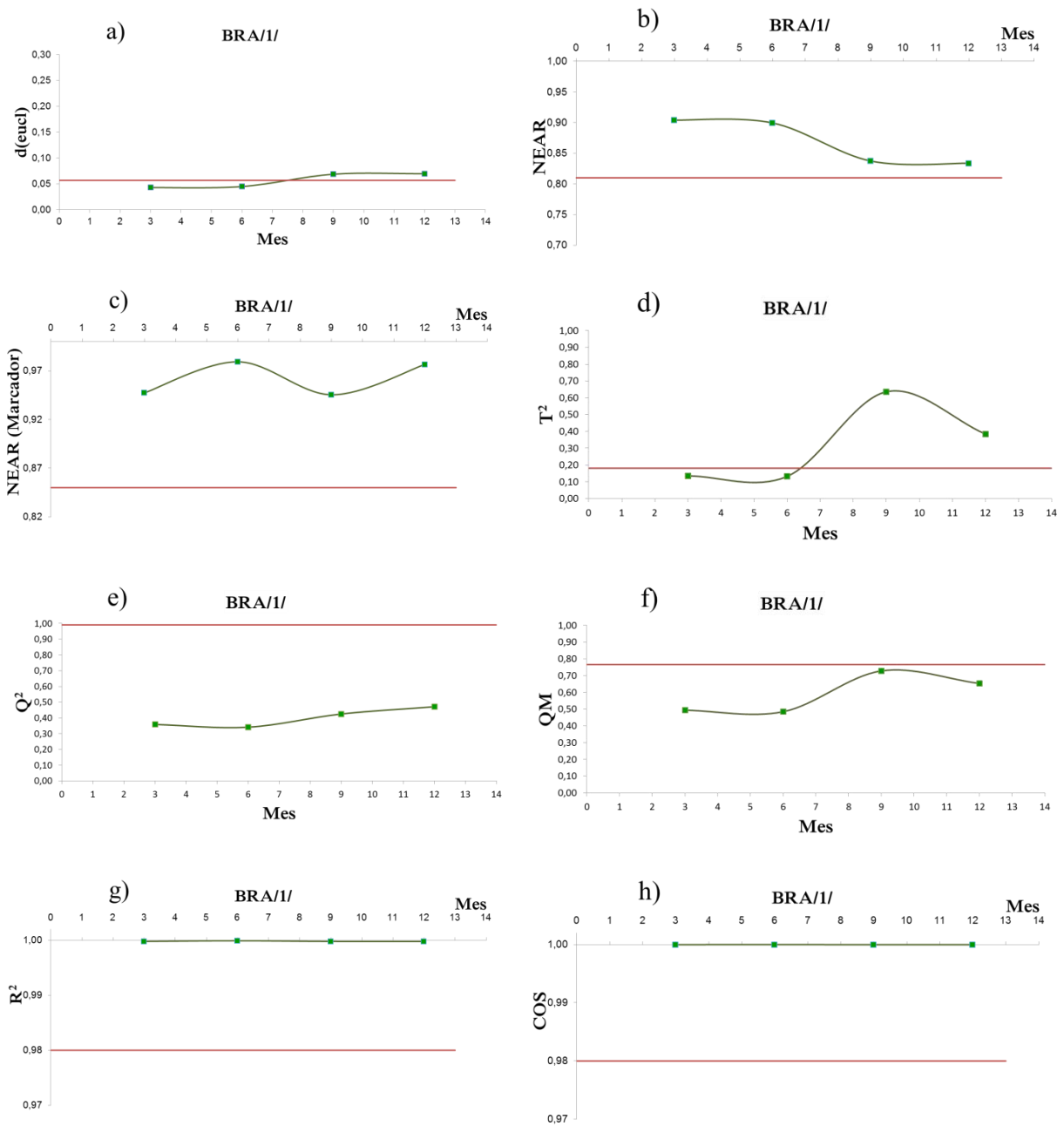


Figura II.3.3. Gráficos de control multivariable con los índices de similitud: a) d(eucl), b) NEAR, c) NEAR (marcadores), d) T²-Hotelling, e) Q²-residual, f) QM, g) R², h) COS; del material BRA/1/.

MATERIAL BRA/2/

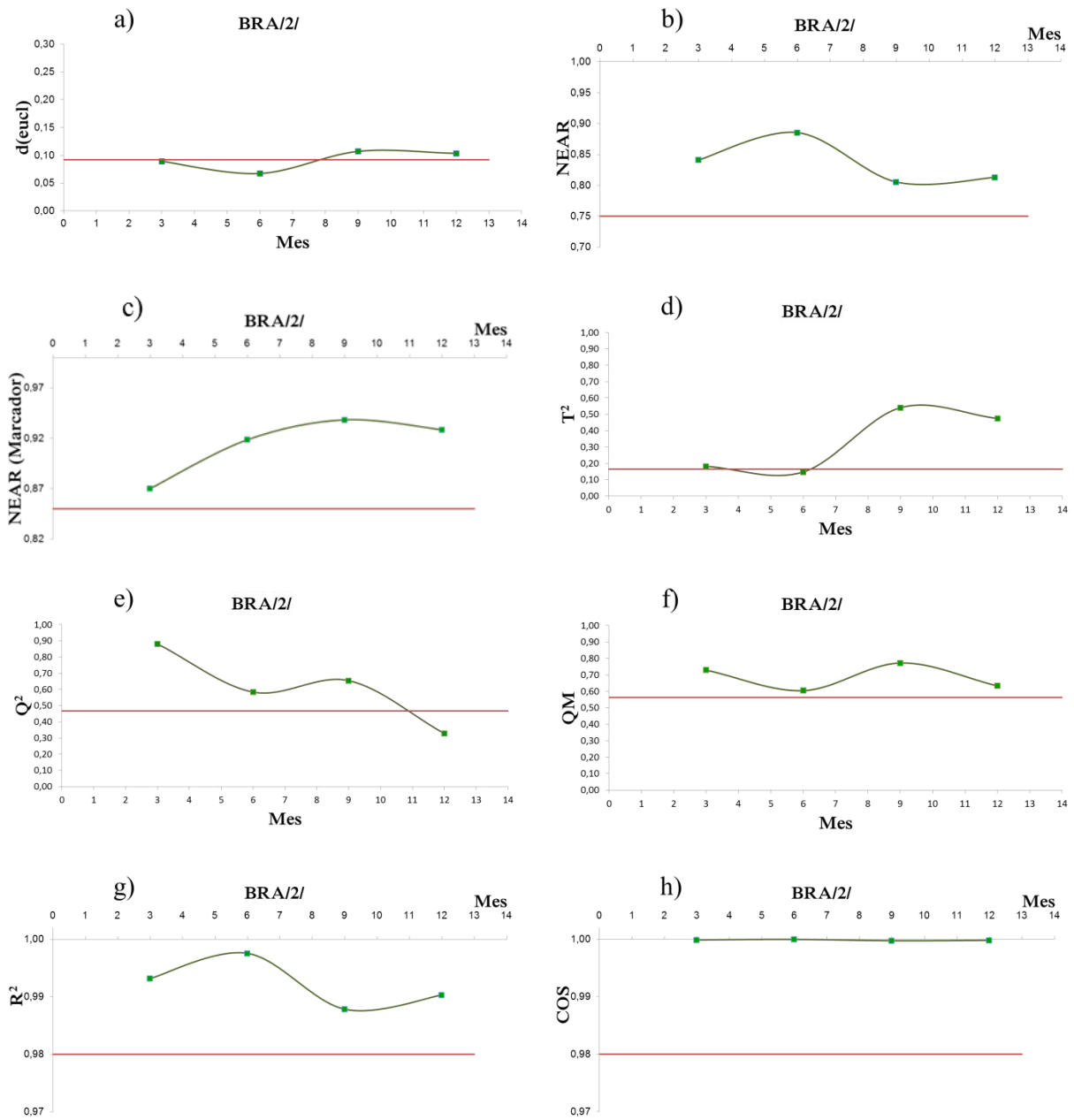


Figura II.3.4. Gráficos de control multivariable con los índices de similitud: a) d(eucl), b) NEAR, c) NEAR (marcadores), d) T²-Hotelling, e) Q²-residual, f) QM, g) R², h) COS; del material BRA/2/.

MATERIAL BRA/3/

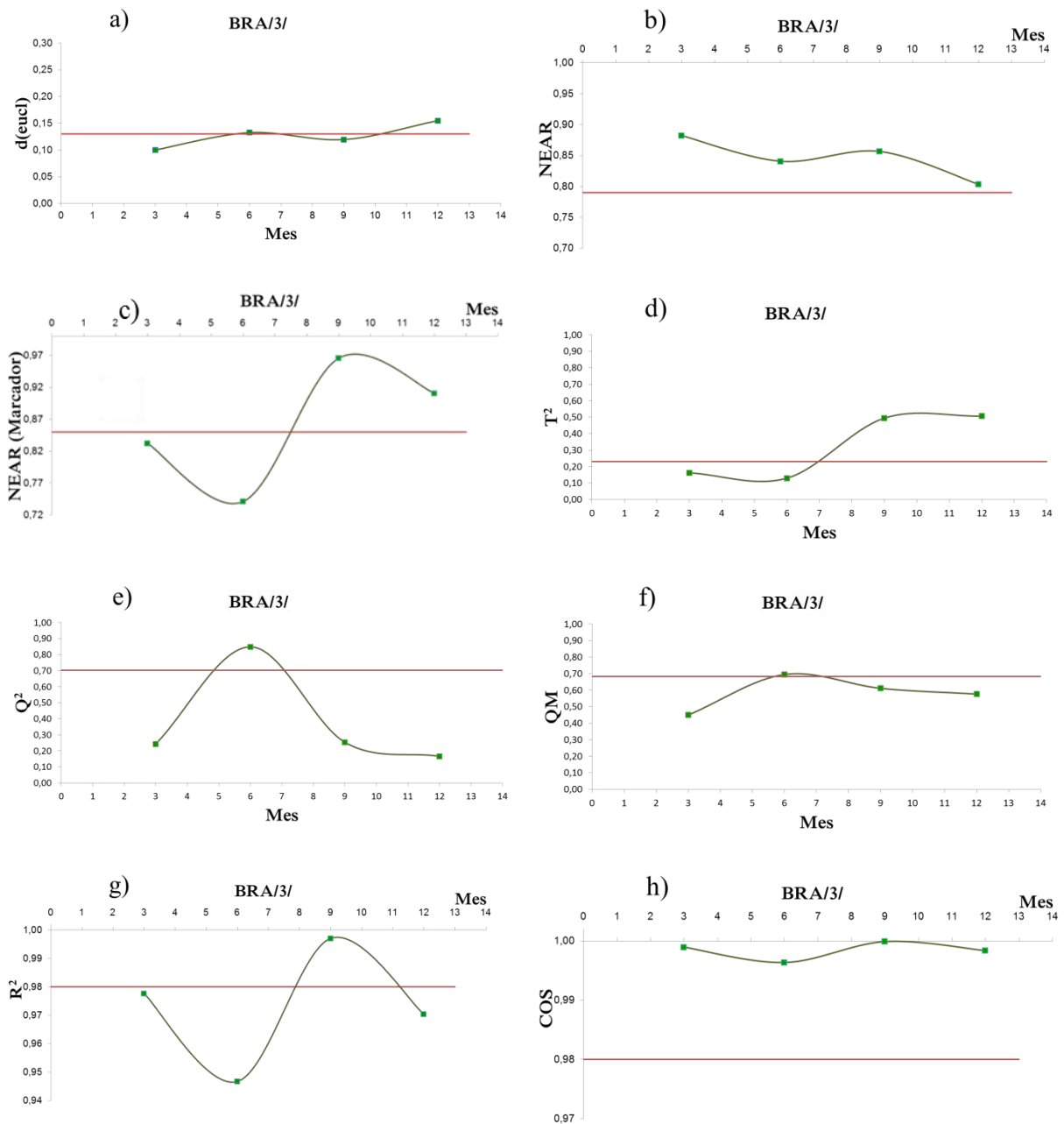


Figura II.3.5. Gráficos de control multivariable con los índices de similitud: a) d(eucl), b) NEAR, c) NEAR (marcadores), d) T²-Hotelling, e) Q²-residual, f) QM, g) R², h) COS; del material BRA/3/.

MATERIAL MAH/1/

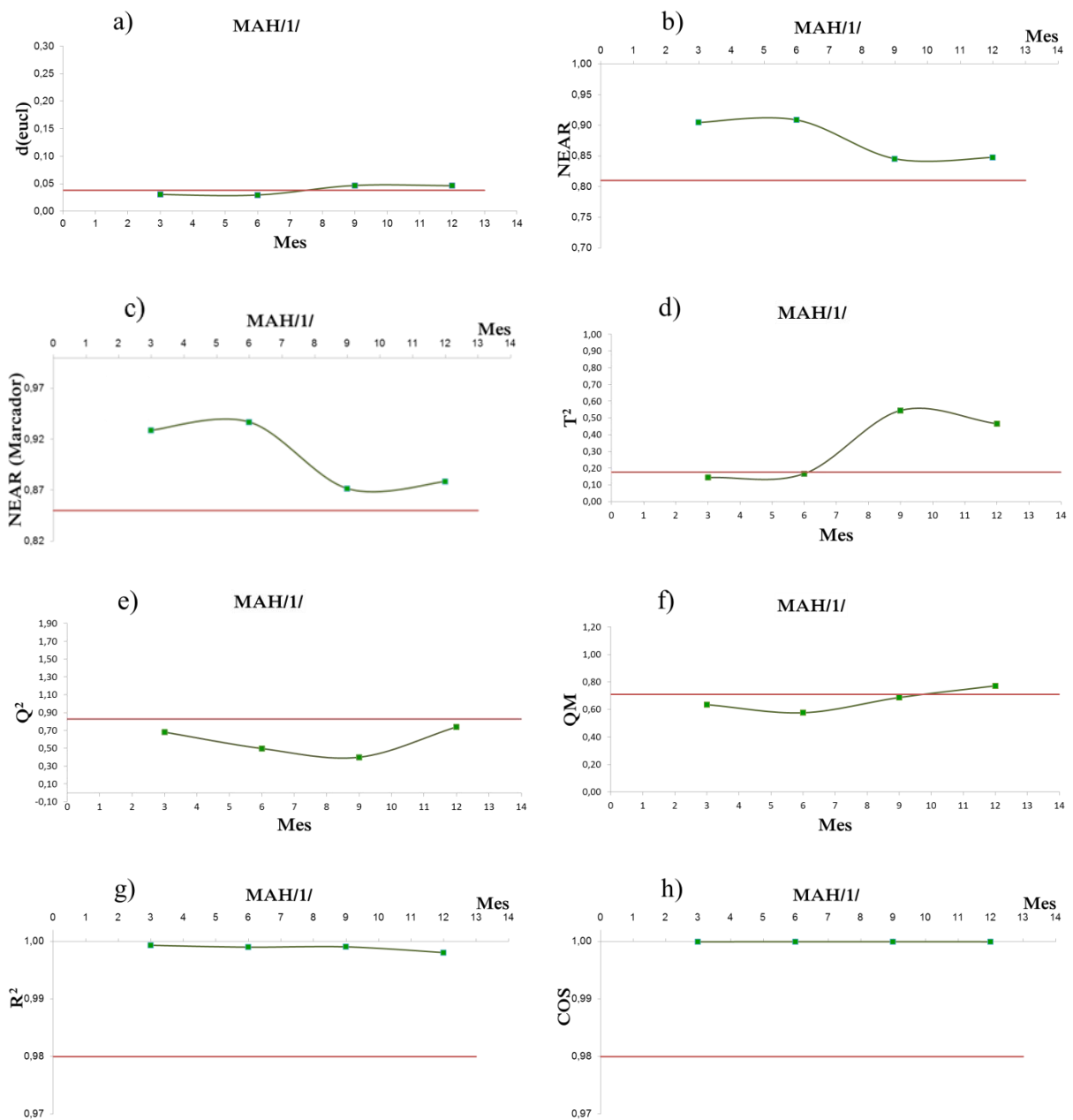


Figura II.3.6. Gráficos de control multivariable con los índices de similitud: a) d(eucl), b) NEAR, c) NEAR (marcadores), d) T²-Hotelling, e) Q²-residual, f) QM, g) R², h) COS; del material MAH/1/.

MATERIAL MAH/2/

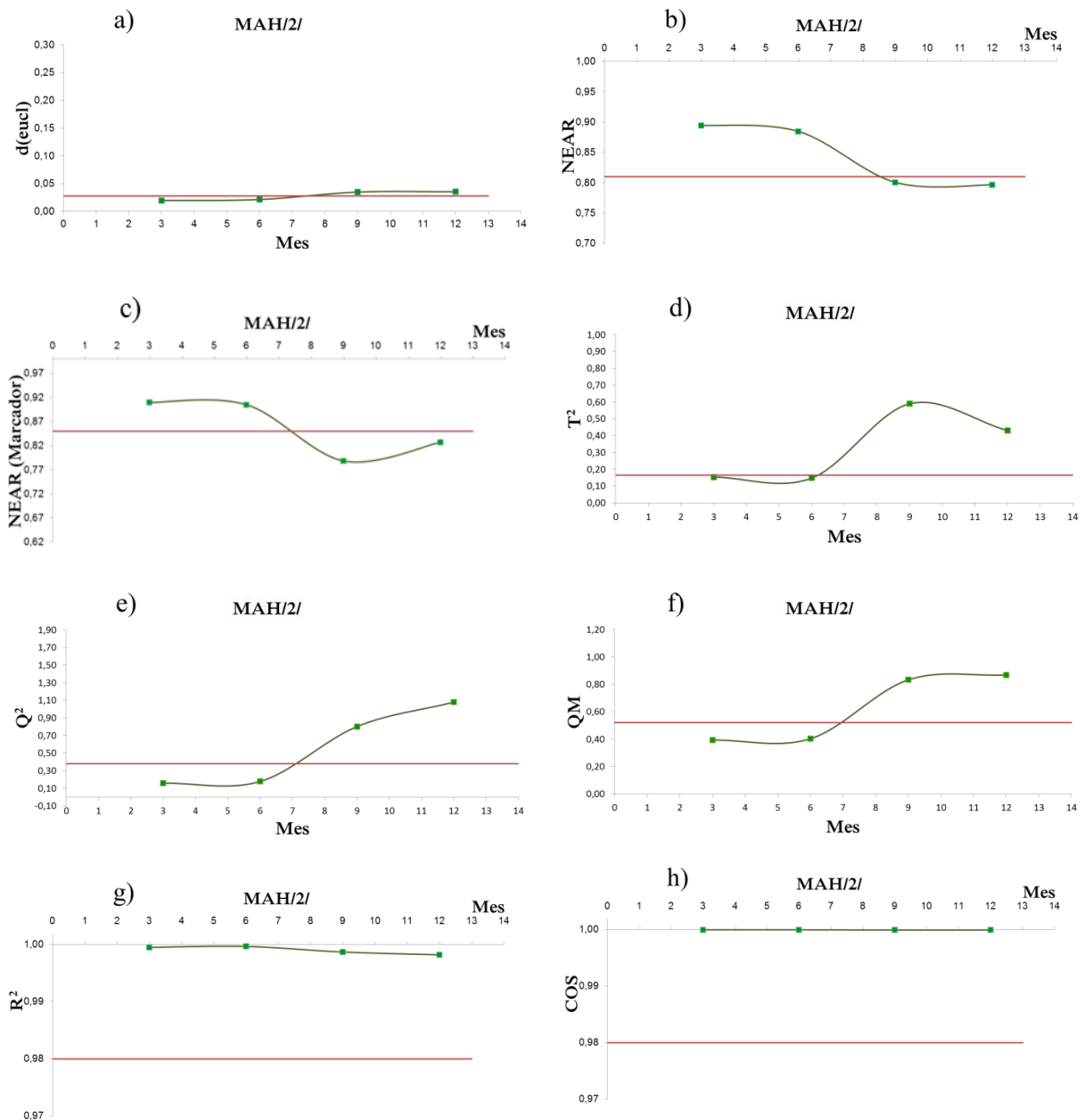


Figura II.3.7. Gráficos de control multivariable con los índices de similitud: a) d(eucl), b) NEAR, c) NEAR (marcadores), d) T²-Hotelling, e) Q²-residual, f) QM, g) R², h) COS; del material MAH/2/.

MATERIAL COC/2/

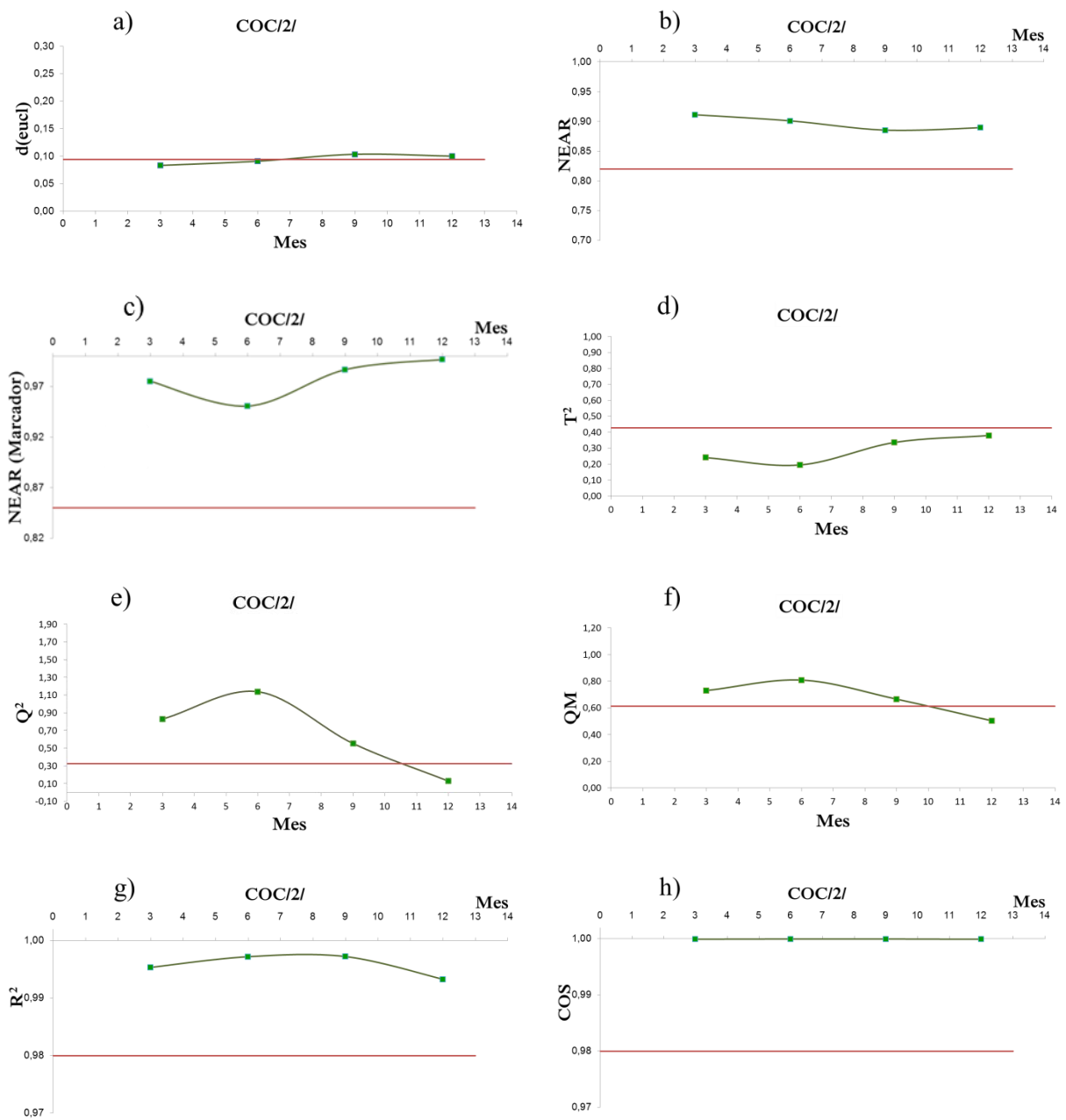


Figura II.3.8. Gráficos de control multivariable con los índices de similitud: a) d(eucl), b) NEAR, c) NEAR (marcadores), d) T²-Hotelling, e) Q²-residual, f) QM, g) R², h) COS; del material COC/2/.

MATERIAL HIG/2/

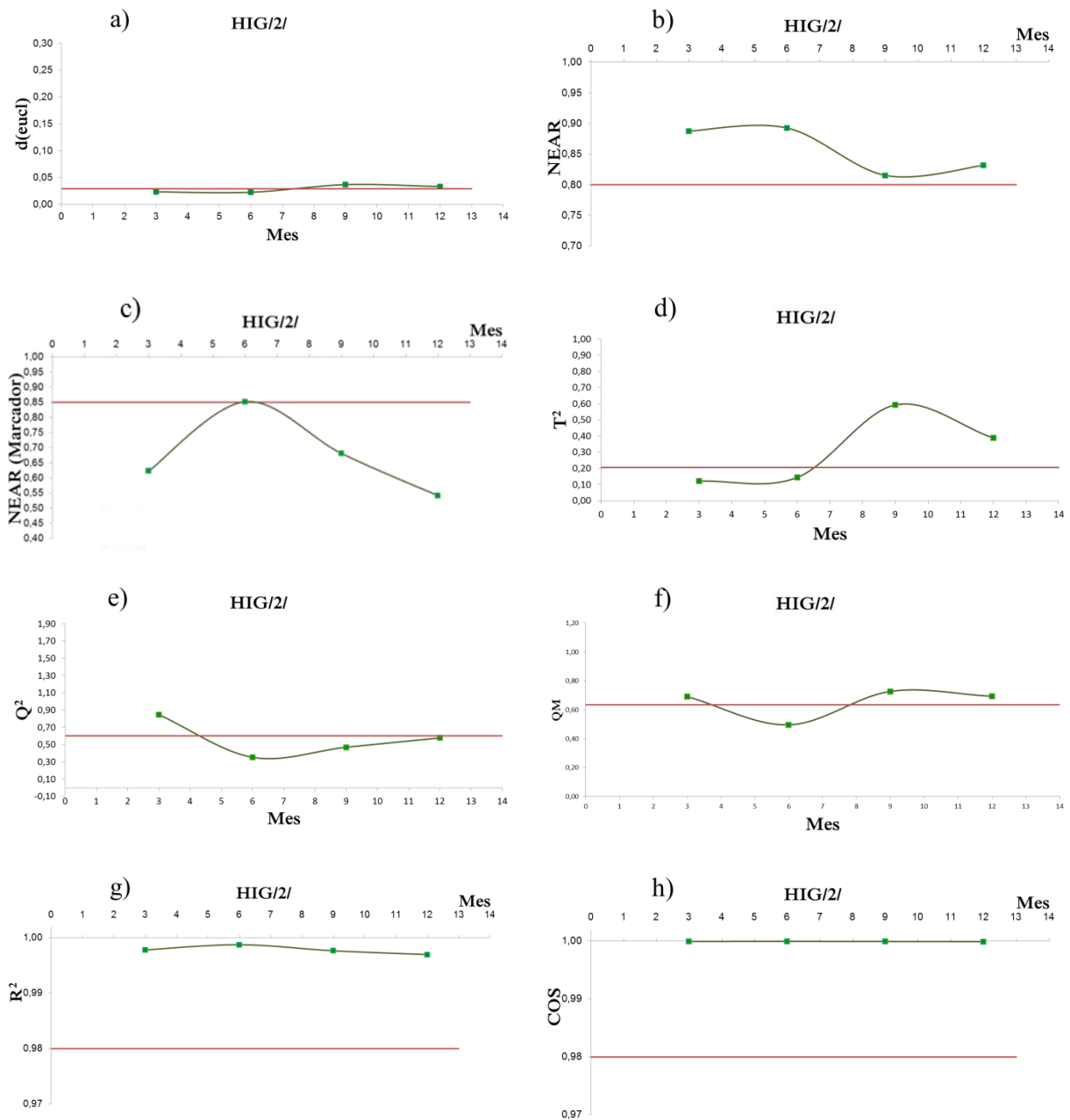


Figura II.3.9. Gráficos de control multivariable con los índices de similitud: a) d(eucl), b) NEAR, c) NEAR (marcadores), d) T²-Hotelling, e) Q²-residual, f) QM, g) R², h) COS; del material HIG/2/.



CAPÍTULO 4

Evolución de la madurez de la aceituna

CAPÍTULO 4

4.1. PRESENTACIÓN

Este capítulo recoge los resultados obtenidos de la identificación de los **compuestos orgánicos volátiles** relacionados con el aroma del aceite de oliva y la **selección** de algunos de ellos como indicadores del estado de **maduración** de la aceituna.

Las muestras utilizadas en el estudio han sido suministradas por el "Grupo de Desarrollo Rural del Altiplano de Granada", fruto de una colaboración entre el centro IFAPA de Cabra (Córdoba) y el grupo de investigación AnAMA, de la Universidad de Granada. El objetivo era estudiar la evolución de la calidad del aceite en función de la maduración de la aceituna. Para ello se seleccionaron tres parcelas situadas en tres localidades del norte de la provincia de Granada: Caniles, Baza y Benamaurel, de las que se recogió la aceituna y se obtuvo el correspondiente aceite. Las muestras de aceite de oliva suministradas contenían información acerca de la calidad: valores de acidez, índice de peróxidos, absorción ultravioleta y análisis sensorial.

Además, esta investigación se realizó en colaboración con el "Dipartimento di Scienza e Tecnologia del Farmaco", de la Universidad de Turín (Italia), durante una estancia pre-doctoral de tres meses y bajo la supervisión y codirección de la Dra. Chiara E. Cordero. Gracias a la realización de esta estancia fue posible el uso de la cromatografía de gases bidimensional exhaustiva (*comprehensive bidimensional gas chromatography, 2D-GC or GC×GC*) acoplada a un detector de espectrometría de masas con analizador de cuadrupolo simple, para el análisis de las muestras.

Los **objetivos** de la investigación fueron:

- ✓ Caracterizar el perfil analítico de los metabolitos que determinan la fracción de compuestos orgánicos volátiles de los aceites de oliva virgen.

- ✓ Estudiar los resultados que se obtienen al trabajar con la identificación dirigida (*target*) o sin ella (*untargeted*), a partir de la "huella dactilar" cromatográfica de los compuestos orgánicos volátiles y que den información de los atributos sensoriales y el estado de madurez de la aceituna.
- ✓ Seleccionar diferentes indicadores de madurez para obtener aceites de oliva virgen de calidad (entre los meses de noviembre a enero).

4.2. INTRODUCCIÓN

Formación de compuestos orgánicos volátiles

El aceite de oliva virgen se obtiene de aceitunas, por métodos físicos que no alteren el producto. La calidad y su clasificación en las tres categorías: virgen extra, virgen y lampante vienen dadas por determinados parámetros físico-químicos de acidez, índice de peróxidos, absorción ultravioleta, y ésteres etílicos de los ácidos grasos; y por la presencia/ausencia de ciertos atributos organolépticos determinados mediante análisis sensorial [1]. Como se comentó en el capítulo anterior, el aroma del aceite de oliva está formado por una mezcla compleja de compuestos orgánicos volátiles. La formación de dichos compuestos se produce a través de la ruta de la lipoxigenasa, por la acción de una serie de enzimas que oxidan los ácidos grasos poliinsaturados para producir aldehídos, alcoholes y ésteres (ver figura 4.1) [2].

Los **aldehídos, alcoholes, cetonas y ésteres** son los **principales compuestos**, pero también está presentes los **furanos, hidrocarburos, ácidos y compuestos aromáticos**, aunque en **cantidades menores**. La influencia de

[1] Reglamento (CEE) N° 2568/91 de la comisión de 11 de julio de 1991 relativo a las características de los aceites de oliva y de los aceites de orujo de oliva y sobre sus métodos de análisis (versión consolidada 01991R2568), Diario Oficial de la Unión Europea (2016), 031.001, 1-128.

[2] J. Sanchez, J.J. Salas (2003). "Biogénesis del aroma del aceite de oliva". En: R. Aparicio, J. Harwood (Eds), Manual del aceite de olive, AMV Ediciones, pp.89-107.

cada volátil en la calidad sensorial depende de las relaciones entre diferentes compuestos, su concentración, así como el umbral de olor [3].

Los aldehídos C6 y sus derivados son responsables del olor "verde" característico de las hojas verdes; los aldehídos C9 están presentes en el aroma de frutos tales como el pepino y la pera [2]. El (E)-2-hexenal, hexanal y (Z)-3-hexenal son los principales aldehídos que se encuentran en los aceites de oliva virgen. El (Z)-3-hexenol, (E)-2-hexenol y hexanol, se encuentran en altas concentraciones y la diferencia de su proporción depende de la variedad de la aceituna así como de la etapa de maduración de la misma [3].

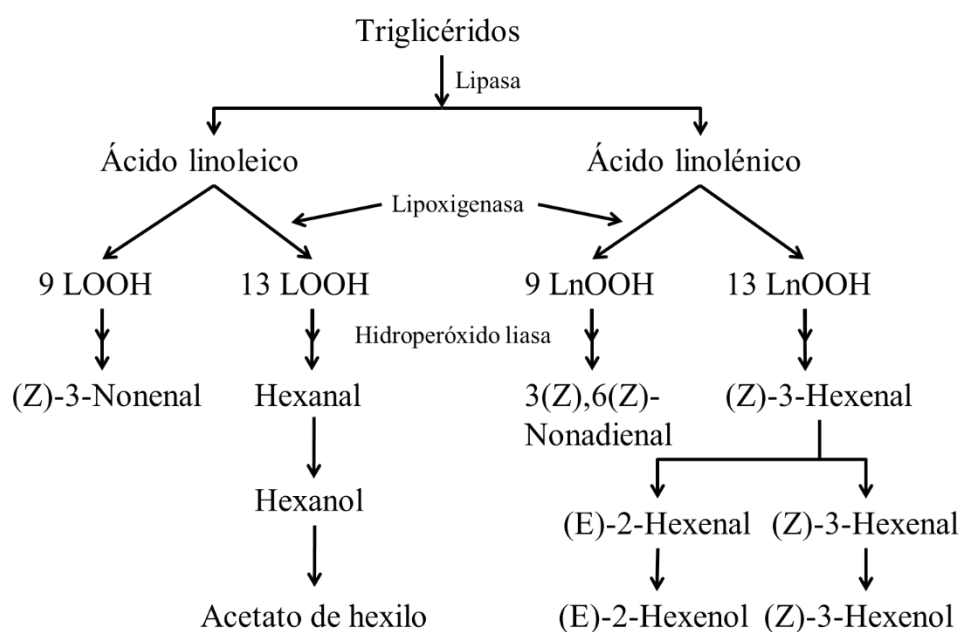


Figura 4.1. Esquema de los principales compuestos orgánicos volátiles generados en la ruta de la lipoxigenasa que dan las características aromáticas al aceite de oliva (esquema extraído de [2,4], donde LOOH: ácido hidroxiperoxilinoleico; LnOOH: ácido hidroxiperoxilinoléico).

Los compuestos volátiles son capaces de caracterizar el aceite de oliva en función de la variedad, la madurez, el sistema de extracción y el origen geográfico. Los principales volátiles relacionados con la madurez son los

[3] D.L. García González, M.T. Morales, R. Aparicio (2010). "Olive and olive oil". En: F. Chen, L.M.L. Nolle (Eds.), Handbook of fruit and vegetable flavours, Wiley, pp. 821-847.

compuestos volátiles que se relacionan con las aceitunas verdes, y que varían de acuerdo con el grado de madurez de las aceitunas [4].

Los atributos negativos o defectos aparecen en el aceite de oliva por la mala calidad de las aceitunas, producidos por la sobre-maduración del fruto, oxidación de los ácidos grasos insaturados, y ataques de mohos y bacterias; así como de posibles fermentaciones y procesos oxidativos [3,4]. El aceite de oliva de baja calidad muestra un perfil de volátiles más complejo y con un número de compuestos volátiles más elevado.

Seguidamente se describe cómo se forman cada uno de los atributos negativos que aparecen en la ficha de cata [1] así como algunos de los compuestos volátiles responsables de cada uno de esos defectos.

El **defecto atrojado/borras** presenta un flavor producto de la fermentación anaerobia que se produce por el almacenamiento o amontonamiento prolongado de las aceitunas, previa a la extracción del aceite; o aceites que ha permanecido en contacto con lodos de decantación. El **defecto moho-humedad** presenta un flavor característico del aceite obtenido de aceitunas en las que se han desarrollado abundantes hongos y levaduras a causa de mantenerlas amontonadas con humedad durante varios días; o de aceitunas recogidas con tierra o barro y que no han sido lavadas. El **defecto avinado-avinagrado/ácido-agrio**, representa el flavor característico de aceites en los que las aceitunas han fermentado de forma aeróbica y que dan lugar a la formación de ácido acético, acetato de etilo y etanol. El **defecto rancio** presenta un flavor característico en los aceites procedentes de aceitunas que han sufrido un proceso oxidativo. Por último, el defecto de **aceituna helada (madera húmeda)** es el flavor característico de aceites que han sido extraídos de aceitunas que han sufrido un proceso de congelación en el árbol. Existen otros defectos como el **heno-madera** que es el flavor del aceite procedente de aceitunas secas; **salmuera** que es el flavor del aceite extraído de aceitunas conservadas en salmuera; **gusano**, flavor característico del aceite obtenido de aceitunas atacadas por larvas de mosca del olivo; entre otros.

[4] M.T. Morales, M. Tsimidou (2003). "El papel de los compuestos volátiles y polifenoles en la calidad sensorial del aceite de oliva". En: R. Aparicio, J. Harwood (Eds), Manual del aceite de olive, AMV Ediciones, pp.381-441.

Existen numerosas tablas que relacionan los compuestos orgánicos volátiles con uno o varios atributos de olor [4,5,6]. Los volátiles pertenecientes a aldehídos y alcoholes ramificados de C5 o dienas C6-C10 o aldehídos monoinsaturados C7-C11, así como cetonas C8, pueden contribuir a los defectos atrojado/borras, moho-humedad, avinado-avinagrado y rancio [6]. El defecto atrojado/borras se relaciona con octano, 2-metilbutanal, 3-metilbutanoico, 6-metil-5-hepten-2-ona. Los aceites avinados muestran un pico elevado correspondiente al ácido acético, además de acetato de etilo y 3-metil-1-butanol. En los aceites con defecto rancio, aparecen los compuestos: pentanal, hexanal, heptanal, octanal, nonanal, algunos de ellos también aparecen en el defecto cocido o quemado [4,6].

Maduración de la aceituna

Actualmente existe un gran número de variedades de aceitunas que producen aceites de oliva de calidad virgen extra. Una de las preocupaciones de los productores estriba en la necesidad de conocer el momento óptimo para recolectar las aceitunas y obtener aceites de oliva con las mejores propiedades sensoriales y fisicoquímicas que den lugar a aceites virgen extra, para lo que es necesario conocer la evolución de las características sensoriales en relación con la madurez de la aceituna. Esta investigación se centra en poner de manifiesto la relación que existe entre las características sensoriales y los compuestos orgánicos volátiles para poder relacionarlos con el estado de maduración de la aceituna.

El proceso de maduración es el resultado de la combinación de factores genéticos, ambientales y condiciones culturales. Los cambios más evidentes se producen en el tamaño de las aceitunas y en el color de la piel, aunque no necesariamente se produce en todas las aceitunas el paso del color verde al

-
- [5] A. Taticchi, S. Esposto, M. Sevilí (2014). "The basic of the sensory properties of virgin olive oil". En: E. Monteleone, S. Langstaff (Eds), Olive oil sensory science, Wiley Blackwell, pp.33-54.
- [6] M.T. Morales, R. Aparicio-Ruiz, R. Aparicio (2013). "Chromatographic methodologies: compounds for olive oil odor issues". En: R. Aparicio, J. Harwood (Eds), Handbook of olive oil, segunda edición, Springer, pp.261-309.

negro, existen variedades que se mantienen de color verde a lo largo de toda la maduración de la aceituna, como es el caso de la variedad Frantoio [7].

Cuando las aceitunas están maduras, la cantidad de compuestos volátiles aumenta hasta alcanzar una concentración máxima cuando la piel de la fruta cambia de amarillo verdoso a violeta, especialmente en los aldehídos C6 como el (E)-2-hexenal. Posteriormente, la concentración del compuesto disminuye debido a la menor actividad de las enzimas encargadas de su producción, debilitándose algunas notas sensoriales del atributo "verde" [5]. Además, todos los volátiles formados en la ruta de la lipoxigenasa disminuyen con la madurez del fruto [6].

La maduración del olivo es un proceso escalar en el que no todas las frutas de un olivo maduran al mismo tiempo. Esto depende de si las ramas se encuentran en una posición más alta o baja, en partes sombreadas o soleadas, en función del tamaño del árbol, así como su posición [7].

Los estudios de maduración de la aceituna se centran en evaluar el rendimiento en función del contenido de aceite que es posible extraer de las aceitunas, los compuestos fenólicos que presentan, así como los parámetros físico-químicos y de análisis sensorial que dan lugar a la calidad y caracterización de la aceituna [8,9,10]. Sin embargo, existen menos estudios basados en la relación del perfil de compuestos orgánicos volátiles con diferentes etapas de maduración de la aceituna [11,12,13].

-
- [7] L. Nasini, P. Proietti (2014), "Olive harvesting". En: C. Peri (Ed.), *The extra-virgin olive oil handbook*, Wiley Blackwell, pp. 89-105.
 - [8] M.N. Franco, J. Sánchez, C. de Miguel, M. Martínez, D. Martín-Vertedor (2015), "Influence of the fruit's ripeness on virgin olive oil quality", *Journal of Oleo Science*, 64, 263-273.
 - [9] B. Jiménez Herrera, A. Rivas Velasco, A. Sánchez-Ortiz, M.L. Lorenzo Tovar, M. Úbeda Muñoz, R.M. Callejón, E. Ortega Bernaldo de Quirós (2012), "Influencia del proceso de maduración del fruto en la calidad sensorial de aceites de oliva virgen de las variedades Picual, Hojiblanca y Picudo", *Grasas y Aceites*, 63, 403-410.
 - [10] L. Martínez Nieto, G. Hodaifa, J.L. Lozano Peña (2010), "Changes in phenolic compounds and rancimat stability of olive oils from varieties of olives at different stages of ripeness", *Journal of Science of Food and Agriculture*, 90, 2393-2398.
 - [11] S.G. Karagoz, M. Yilmazer, G. Ozkan, A.A. Carbonell-Barrachina, M. Kiralan, M.F. Ramadan (2017), "Effect of cultivar and harvest time on C₆ and C₅ volatile compounds of Turkish olive oils", *European Food Research and Technology*, 243, 1193-1200.
 - [12] R. Hachicha Hbaieb, F. Kotti, M. Gargouri, M. Msallem, S. Vichi (2016), "Ripening and storage conditions of Chétoui and Arbequina olives: Part I. Effect on olive oils volátiles profile", 203, 548-558.

Instrumentación

La **técnica de cromatografía de gases bidimensional** es de utilidad para profundizar en el conocimiento del perfil de los compuestos orgánicos volátiles (y semivolátiles) del aceite de oliva. Esta técnica permite identificar un mayor número de compuestos que la cromatografía de gases de una sola dimensión debido a que presenta una mayor resolución [6]. Por este motivo, en esta investigación se ha utilizado un cromatógrafo de gases bidimensional acoplado a un detector de espectrometría de masas con analizador de cuadrupolo (GC×GC-(Q)MS) (figura 4.2).



Figura 4.2. Equipo GC×GC-(Q)MS con modulador de nitrógeno.

En este tipo de equipo se dispone de un **horno cromatográfico** donde se encuentran alojadas **dos columnas** cromatográficas. La primera de ellas es una columna convencional como la que se usa en la cromatografía de gases unidimensional, que se encuentra unida a través de una interfase a una segunda columna mucho más corta, normalmente de un metro de longitud, y de menor diámetro de partícula. La interfase, o **modulador**, colecta el efluente de la primera columna y lo introduce periódicamente a la segunda columna. Existen diferentes tipos de moduladores, en este caso, el equipo utilizado dispone de un

[13] M.T. Morales, R. Aparicio, J.J. Calvente (1996). "Influence of olive ripeness on the concentration of green aroma compounds in virgin olive oil", *Flavour and fragrance journal*, 11, 171-178.

modulador que usa **nitrógeno** líquido para enfriar rápidamente la interfase y "congelar" los analitos antes de que eluyan a la segunda columna [14]. En la figura 4.3 se muestra un esquema de la configuración del equipo.

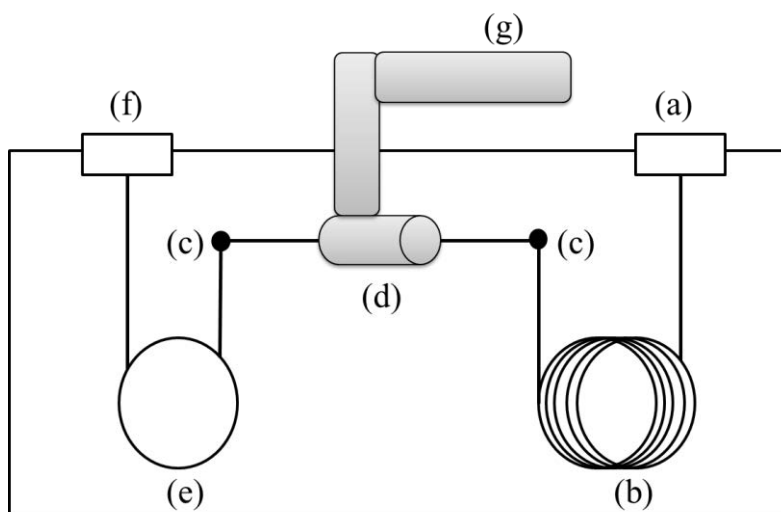


Figura 4.3. Configuración del equipo GC×GC: (a) inyector; (b) columna primaria; (c) conectores; (d) modulador; (e) columna secundaria; (f) detector; (g) nitrógeno líquido. (Figura adaptada [14]).

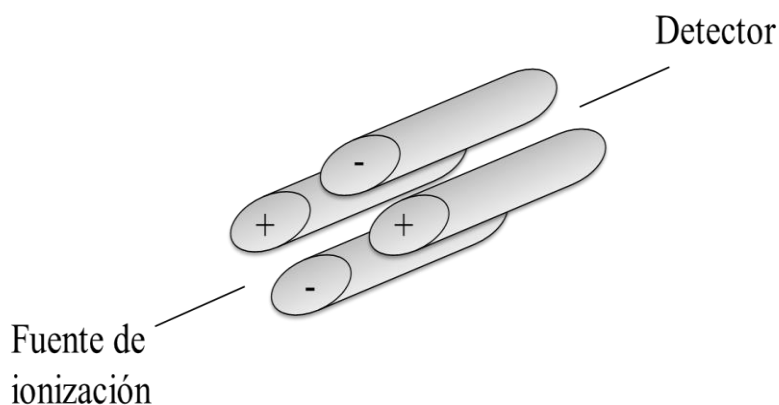


Figura 4.4. Configuración del analizador: cuadrupolo simple. (Figura adaptada [15]).

[14] A. Mostafa, T. Górecki, P.Q. Tranchida, L. Mondello (2011), "History, evolution, and optimization aspects of comprehensive two-dimensional gas chromatography". En: L. Mondello (Ed.), *Comprehensive chromatography in combination with mass spectrometry*, John Wiley and Sons, pp. 93-144.

Se ha utilizado un **espectrómetro de masas** con analizador tipo **cuadrupolo simple**, que consiste en cuatro varillas cilíndricas orientadas en una disposición cuadrada (figura 4.4) por donde pasan los iones cargados generados en la fuente de ionización con una trayectoria estable. En las barras cilíndricas se aplica un potencial constante y un potencial alterno de radiofrecuencia superpuesto que origina el movimiento de los iones, en función de su m/z , sin modificar su movimiento longitudinal [15].

Los **cromatogramas** obtenidos, en el caso de la cromatografía 2D-GC, se corresponden con cromatogramas tridimensionales que habitualmente se muestran como imágenes bidimensionales en forma de mapas de contorno o más generalmente en forma de **mapas de "calor"** donde la intensidad de los picos se representa por la intensidad del color [16]. En la figura 4.5 se muestra un ejemplo de las diferentes formas de representar un cromatograma 2D.

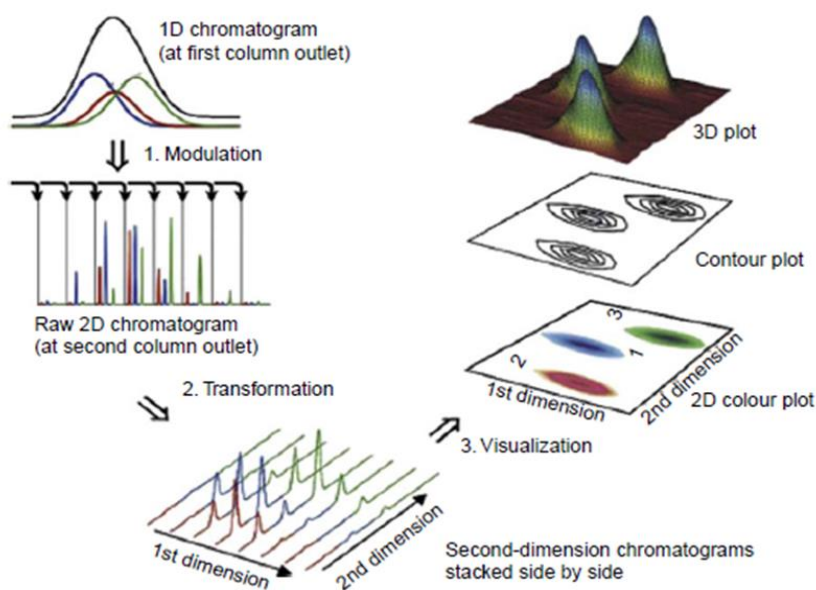


Figura 4.5. Generación de datos y visualización en GC×GC, tomado de la fuente Ramos *et al.* [16].

- [15] J.A. Masucci, G.W. Caldwell (2004), "Techniques for gas chromatography/mass spectrometry". En: R.L. Grob, E.F. Barry (Eds.), Modern practice of gas chromatography, cuarta edición, John Wiley and Sons, pp. 339-401.
- [16] L. Ramos, U.A.T. Brinkman (2009), "Multidimensionality in gas chromatography: general concepts". En: D. Barceló (Ed.) Comprehensive analytical chemistry, vol. 55, Comprehensive two dimensional gas chromatography (L.Ramos, Ed.), Elsevier, pp. 3-14.

Una vez obtenidos los cromatogramas bidimensionales en forma de "manchas", se puede trabajar con **análisis dirigido (*target*)**, cuyo objetivo es identificar cada uno de los picos del mapa de color. Un pico cromatográfico se identifica seleccionando la "mancha" y comparando la relación m/z de los iones que la forman en una librería de espectros, donde están almacenados un gran número de compuestos.

Ese pico cromatográfico seleccionado e identificado con un compuesto concreto mediante la librería de espectros, puede confirmarse con lo que se denomina "índice de retención lineal" (IST). Este índice indica un número que marca la posición del pico cromatográfico en las dos dimensiones, en función de la columna cromatográfica seleccionada. Para ello, previamente a la inyección de la muestra, se inyecta una mezcla de patrones de n-alcenos (C9-C25). El cromatograma de esta mezcla es muy sencillo porque aparecen únicamente cada uno de los patrones presentes en la mezcla, desde el C9 al C25. Entonces se identifican cada uno de ellos en el cromatograma seleccionando el pico del n-alceno correspondiente y dándole su nombre en el *software*. Tras esto, en el *software* se genera una tabla Excel con cada uno de los n-alcenos identificados, sus tiempos de retención en las dos dimensiones y el ITS de cada uno de ellos.

Una vez se obtiene el cromatograma bidimensional de la muestra, se carga en el *software* la tabla Excel generada con la información de la mezcla de patrones de n-alcenos. Así, a cada uno de los picos cromatográficos presente en la muestra, se le asignará automáticamente un ITS. Por lo que cada pico identificado con la librería de espectros en la muestra tendrá también un ITS que se podrá corroborar con los datos aportados en bibliografía. Confirmando así cada compuesto identificado en la muestra.

Otra forma de trabajar es con el **análisis no dirigido (*untargeted*)**. El objetivo es seleccionar todos los picos cromatográficos presentes en todas las muestras analizadas, aunque estos no sean identificados. Con este tipo de análisis se toman tanto los picos cromatográficos ya identificados mediante el análisis *targeted* y el resto de picos que no han sido identificados pero que aparecen en todas las muestras analizadas. En este caso, el tratamiento de datos es un paso crucial en el proceso analítico de esta técnica GC×GC, puesto que es posible obtener información de picos no identificados.

Ambas estrategias se basan en el enfoque conocido como "*template matching fingerprinting*" [17]. Este enfoque consiste en la selección de las regiones de cada uno de los picos cromatográficos (identificados o no), que se guardan como plantilla sobre la "huella dactilar" y que puede ejecutarse en el cromatograma bidimensional de cualquier tipo de muestra.

En esta investigación se ha trabajado con ambas estrategias y se explica con más detalle en los apartados 4.4 y 4.5 correspondiente con el proceso experimental y el artículo científico, respectivamente.

4.3. MUESTRAS

Las muestras utilizadas en esta investigación han sido aceites de oliva obtenidos directamente de aceitunas de la variedad Picual recolectadas en diferentes estados de maduración. Las aceitunas proceden de tres parcelas diferentes de la zona del Altiplano de Granada, en los municipios de Caniles, Baza y Benamaurel. Cada muestra de aceite fue obtenida por duplicado a partir de la selección de una mezcla de aceitunas procedentes 10 árboles distribuidos por la parcela de tal forma que se obtuvieron muestras distintivas de todas las partes de la finca, en cuatro etapas de maduración: 10-12 noviembre 2014; 24-28 noviembre 2014; 16-17 diciembre 2014 y 12-15 enero 2015.

El aceite de oliva se extrajo de las aceitunas recopiladas mediante un sistema ABENCOR. Este proceso consta de tres etapas. En primer lugar se lleva a cabo la molienda del fruto (con hueso incluido) con el objetivo de reducir su tamaño para facilitar la obtención del aceite. Posteriormente, la pasta recogida se somete a un batido a temperatura constante (38.5 °C) con el que se consigue la reducción de la viscosidad, la agregación de las gotas de aceite y la obtención de una fase homogénea donde comienza su separación de la fase acuosa. Ese proceso se realiza en una termobatidora. La pasta se pasa después a una centrifuga donde se obtienen dos fracciones: alpeorujo y aceite (con pequeñas impurezas debidas a la otra fase). Para terminar de obtener un aceite limpio, se

[17] S.E. Reichenbach, P.W. Carr, D.R. Stoll, Q. Tao (2009). "Smart templates for peak pattern matching with comprehensive two-dimensional liquid chromatography", *Journal of Chromatography A*, 1216, 3458-3466.

decanta la mezcla, de manera que el alpeorujo se deposita en el fondo del decantador y se puede retirar el aceite, que queda en la parte superior.

En la tabla 4.1 aparecen todos los parámetros de calidad según el Reglamento N° 2568/91 [1] de cada una de las muestras. La determinación de cada uno de esos parámetros se realiza de la siguiente manera:

- La acidez se produce por la liberación de los ácidos grasos al romperse las moléculas de triglicéridos y se produce por un mal estado de los frutos, una incorrecta elaboración o una mala conservación. La determinación consiste en disolver la muestra en 50-150 mL de una mezcla de éter dietílico con etanol previamente preparada y neutralizada para la posterior valoración de los ácidos grasos libres que contiene mediante una solución etanólica de hidróxido potásico. Su valor se expresa en porcentaje de ácido oleico.
- El grado de oxidación del aceite se puede determinar a partir del índice de peróxidos y su elevación depende de aquellos factores que favorecen la oxidación (luz, aireación, temperatura...) así como de las heladas. Se expresa en miliequivalentes de oxígeno activo por kilogramo de grasa y se determina tratando la muestra, disuelta en ácido acético y cloroformo, con una solución de yoduro potásico. El yodo liberado se valora con tiosulfato sódico.
- El K_{270} y K_{232} , son indicadores de la calidad de una materia grasa o de su estado de conservación y se determinan a partir de un examen espectrofotométrico en ultravioleta, cuando las longitudes de ondas son iguales a 270 y 232, respectivamente. K_{270} evalúa la existencia de dienos o trienos conjugados y K_{232} mide la oxidación primaria. Para obtenerlos se disuelve la materia grasa en ciclohexano y se determina la extinción de la disolución a las longitudes de onda seleccionadas, respecto al disolvente puro.
- El contenido de ésteres etílicos de los ácidos grasos. Los ésteres etílicos están asociados a la fermentación alcohólica de las aceitunas, la cual rinde etanol entre otros, que se encuentra disuelto en las pequeñas porciones de agua presentes en el aceite. El etanol en contacto con los ácidos grasos libres forma los ésteres etílicos, de modo que se puede deducir que el contenido de estos compuestos está relacionado con el grado de acidez del aceite virgen y con su contenido de humedad en los depósitos, así como

con el grado de fermentación que se haya podido producir en las aceitunas. Se determina mediante cromatografía de gases (GC-FID), en el mismo método que se aplica para la determinación de ceras.

- El análisis organoléptico se realiza mediante cata sensorial, que fue descrita en el capítulo anterior.

La clasificación en cada una de las categorías aceite de oliva virgen extra (AOVE), aceite de oliva virgen (AOV) y aceite de oliva lampante (AOL), se tienen en cuenta los parámetros de calidad indicados en la tabla 4.2, que es un extracto de la tabla que aparece en el Reglamento relativo a las características de los aceites de oliva y sus métodos de análisis [1].

Teniendo en cuenta los datos que se muestran en la tabla 4.1, todas las muestras se clasifican como AOVE, salvo las muestras correspondientes con las etapas 3 y 4 en la parcela de Baza (AOV y AOL, respectivamente) y en la etapa 4 para la parcela de Benamaurel (AOL). Por lo que se podría indicar que la etapa de maduración óptima es en el mes de noviembre.

La extracción del aceite de oliva a partir de la aceituna, así como los parámetros físico-químicos y sensoriales del aceite de oliva no fueron fruto de esta investigación, pero fueron proporcionados con cada una de las muestras de aceite de oliva.

Tabla 4.2. Parámetros físico-químicos y sensoriales para la clasificación del aceite de oliva en sus tres categorías [1].

Clasificación	Acidez (%)	Índice de peróxido (mEq O ₂ /kg)	K ₂₃₂	K ₂₇₀	Δk	Md ^a	Mf ^a
AOVE	≤ 0.8	≤ 20	≤ 2.50	≤ 0.22	≤ 0.01	0	> 0
AOV	≤ 2.0	≤ 20	≤ 2.60	≤ 0.25	≤ 0.01	≤ 3.5	> 0
AOL	> 2.0	-	-	-	-	> 3.5 ^b	-

^a Mediana de defectos (Md); Mediana de frutado (Mf)

^b La mediana del defecto puede ser inferior o igual a 3.5 cuando la mediana del frutado es igual a cero

Tabla 4.1. Datos de calidad y clasificación de las muestras de aceite de oliva procedentes de diferentes parcelas del Altiplano de Granada.

Parcela	Cosecha	Replica	Acidez (%)	Índice de peróxido (mEq O ₂ /kg)	K ₂₃₂	K ₂₇₀	Δk	Md*	Mf*	Clasificación
Caniles	1	1	0.2	4	2.02	0.2	0	0	7.9	AOVE
Caniles	2	1	0.1	4	1.61	0.2	0	0	3.5	AOVE
Caniles	3	1	0.3	4	1.41	0.2	0	0	3.6	AOVE
Caniles	4	1	0.2	4	1.40	0.1	0	0	2.8	AOVE
Caniles	1	2	0.2	4	2.23	0.2	0	0	6	AOVE
Caniles	2	2	0.2	4	1.71	0.2	0	0	3	AOVE
Caniles	3	2	0.3	4	1.38	0.2	0	0	3.4	AOVE
Caniles	4	2	0.3	4	1.41	0.2	0	0	2.8	AOVE
Baza	1	1	0.2	5	1.84	0.2	0	0	5	AOVE
Baza	2	1	0.2	3	1.60	0.2	0	0	4.1	AOVE
Baza	3	1	0.2	5	1.17	0.2	0	>0	1.3	AOV
Baza	4	1	0.4	11	1.11	0.1	0	>0	0	AOL
Baza	1	2	0.2	4	1.92	0.2	0	0	5.2	AOVE
Baza	2	2	0.1	3	1.65	0.2	0	0	3.8	AOVE
Baza	3	2	0.2	6	1.28	0.1	0	>0	1.7	AOV
Baza	4	2	0.4	13	1.12	0.1	0	>0	0	AOL
Benamaurel	1	1	0.2	5	1.61	0.2	0	0	4.4	AOVE
Benamaurel	2	1	0.2	4	1.53	0.2	0	0	4.3	AOVE
Benamaurel	3	1	0.2	8	1.19	0.1	0	0	3.1	AOVE
Benamaurel	4	1	0.4	19	1.05	0.1	0	>0	0	AOL
Benamaurel	1	2	0.1	4	1.64	0.2	0	0	4.2	AOVE
Benamaurel	2	2	0.2	3	1.48	0.2	0	0	4.3	AOVE
Benamaurel	3	2	0.2	6	1.51	0.1	0	0	2.9	AOVE
Benamaurel	4	2	0.2	14	1.05	0.1	0	>0	0	AOL

* Mediana de defectos (Md); Mediana de frutado (Mf)

4.4. PROCESO EXPERIMENTAL

La determinación de compuestos orgánicos volátiles (VOCs) se llevó a cabo con un equipo GC×GC-(Q)MS previa a su extracción por microextracción con fase sólida mediante fibra. El equipo GC×GC usado es de la marca Agilent 6869, y está acoplado a un detector de espectrometría de masas 5975C. El sistema está equipado con un modulador térmico KT 2004 que enfría con nitrógeno líquido.

Antes de la inyección de las muestras y siempre al comienzo del análisis, se inyecta una mezcla de patrones de n-alcános (C9 a C25), para obtener así los índices de retención lineal (ITS) y hacer una correlación con ellos cuando se seleccionen e identifiquen los compuestos presentes en la muestra.

La mezcla de n-alcános se prepara con una concentración de 100 mg/L en ciclohexano. Se inyecta en el equipo 1 μ L de esta mezcla de patrones, con una temperatura de inyección de 280°C y split (1:50). Las condiciones cromatográficas y del espectrómetro de masas se muestran en la tabla 4.3.

Antes de la extracción de los VOCs sobre la muestra de aceite de oliva, se prepara un patrón α -tujona a una concentración de 100 mg/L en dibutilftalato. Se toman 2 μ L de la disolución y se va depositando por las paredes de un vial de 20 mL con tapón de rosca magnético, como se muestra en la figura 4.6.



Figura 4.6. Distribución de 2 μ L de la disolución patrón de α -tujona en dibutilftalato (100 mg/L).

El vial, con la disolución del patrón interno, se calienta a una temperatura de 50°C. La fibra de extracción: 50/30 μ m DVB/CAR/PDMS, StableFlex 24 Ga (divinilbenceno/carboxen/polidimentilsiloxano), se expone en el vial durante 15

minutos y se retrae hasta la extracción de los VOCs de la muestra de aceite de oliva con esta misma fibra.

Tabla 4.3. Condiciones del cromatógrafo de gases bidimensional y del detector de espectrometría de masas.

Gas portador	Helio a 1.8 mL/min		
Columna cromatográfica 1	SolGel-Wax (30 m × 0.25 mm i.d. × 0.25 μm), 100% polietilenglicol		
Columna cromatográfica 2	OV1701 (1 m × 0.10 mm i.d. × 0.10 μm), 86% polidimetilsiloxano, 7% fenil, 7% cianopropil		
Rampa de temperatura	Temperatura (°C)	Rampa (°C/min)	Tiempo* (min)
	40	0.00	1.00
	200	3.00	0.00
	250	10.00	5.00
	Tiempo de análisis		64.33
Modulador	Tiempo de pulso (ms)	Tiempo de modulación (s)	
	250	4	
Sistema de ionización	Impacto electrónico, 70 eV		
Modo scan	Full		
Temperatura línea transferencia	270 °C		
Rango m/z	40-240		
Velocidad de escaneo (amu/s)	12500		
Frecuencia espectro (Hz)	30		

* *Tiempo que se mantiene la temperatura final a la que llega el gradiente*

Mientras se lleva a cabo la extracción del volátil de α -tujona sobre la fibra (durante esos 15 minutos), se preparan en viales de 20 mL con tapón de rosca magnética, 1.5 g de muestra de aceite de oliva. La muestra se calienta a una temperatura de 50°C para la formación de VOCs sobre el espacio de cabeza. A

continuación la fibra DVB/CAR/PDMS, que contiene el patrón, se expone en el vial de muestra para adsorber los VOCs, durante 40 minutos. En este tiempo, se produce un equilibrio entre los volátiles del patrón α -tujona adsorbido en la fibra y los nuevos volátiles que se encuentran en el espacio de cabeza de la muestra de aceite, adsorbiéndolos también en ella. La fibra se retrae y se lleva al equipo para su inyección.

La desorción de los volátiles contenidos en la fibra se produce a 250°C durante 5 minutos en el inyector, con un split (1:20). Las condiciones del GC×GC-(Q)MS se muestran en la tabla 4.3. Todo el proceso de extracción de los VOCs e inyección en el equipo se llevó a cabo manualmente. Las condiciones de extracción y cromatográficas fueron optimizadas a partir del estudio previo de Vaz-Freire et al. [18].

El número total de muestras de las que se disponía eran 24, y de cada una de ellas se analizó por duplicado, por lo que el número total de muestras analizadas fue 48.

Tratamiento de datos

Los datos son adquiridos por el software Agilent MSD ChemStation versión E.02.01.00 y procesados usando GC Image GC×GC Software versión 2.5.

El tratamiento de datos de la cromatografía bidimensional se lleva a cabo mediante análisis dirigido (*targeted*) y no dirigido (*untargeted*) y aplicando la estrategia de *template matching fingerprinting* [17]. Esto consiste en guardar una plantilla con una serie de picos cromatográficos ("manchas" en el mapa de calor bidimensional) identificados, de una de las muestras de aceite de oliva. En el cromatograma bidimensional de otra muestra de aceite de oliva, se aplica ese mismo *template*, de modo que se identifican los compuestos que aparezcan en el cromatograma y que presenten el mismo tiempo de retención en las dos dimensiones, así como la misma relación de iones m/z.

[18] L.T. Vaz-Freire, M.D.R.G. da Silva, A.M.C. Freitas (2009), "Comprehensive two-dimensional gas chromatography for fingerprint pattern recognition in olive oils produced by two different techniques in Portuguese olive varieties galega vulgar, cobracosa e carrasquenha", *Analytica Chimica Acta*, 633, 263-270.

Esa plantilla o *template* contiene información de los tiempos de retención en las dos dimensiones, el nombre del compuesto identificado, la relación m/z de los iones en cada pico bidimensional identificado, el ITS y la respuesta de intensidad dada por el detector. Por lo tanto, al poner esta plantilla en un nuevo cromatograma bidimensional, se establecen las correspondencias entre los mismos compuestos químicos, de modo que puede usarse en múltiples cromatogramas.

I. Análisis dirigido (*target*)

El primer paso es hacer un análisis dirigido (*target*) de todos los compuestos presentes en el cromatograma bidimensional, que se identificaron con la librería de espectros NIST MS Search algorithm, mediante la relación m/z y el IST de cada pico cromatográfico. De este modo, fue posible identificar un total de **119 compuestos** en las muestras de aceite, y que se indican en la tabla 2 del artículo científico publicado, fruto de esta investigación.

La figura 4.7 muestra un zoom del cromatograma bidimensional de la muestra de aceite de oliva de la parcela de Benamaurel, la réplica 1 en el estado de maduración 4 (enero 2015), que corresponde a un AOL, según la clasificación sensorial dada por el panel de cata (tabla 4.1).

En esta imagen se ve como eluyen los diferentes compuestos. Así, la región más alta de VOCs eluidos en la segunda columna, son los hidrocarburos ramificados insaturados (línea 1 de la figura 4.7). Además, se observa una separación por esta misma columna de los compuestos saturados e insaturados (líneas 3 y 4 de la figura 4.7).

El área de cada uno de los picos cromatográficos bidimensionales identificados, quedan seleccionados a través de ese *template* y se puede extraer con el *software* una tabla en Excel que indica el área de cada uno de ellos. El área de todos los picos cromatográficos identificados se normaliza con respecto al patrón interno α -tujona, que se selecciona previamente en el cromatograma bidimensional, y se relaciona con cada una de las "manchas" marcadas en el mapa de calor. La figura 4.8 muestra el cromatograma bidimensional de la réplica 1 de aceite de oliva obtenido en la etapa 4 de la parcela de Benamaurel. Las líneas de color rojo indican la unión entre el patrón interno y cada uno de los picos cromatográficos marcados.

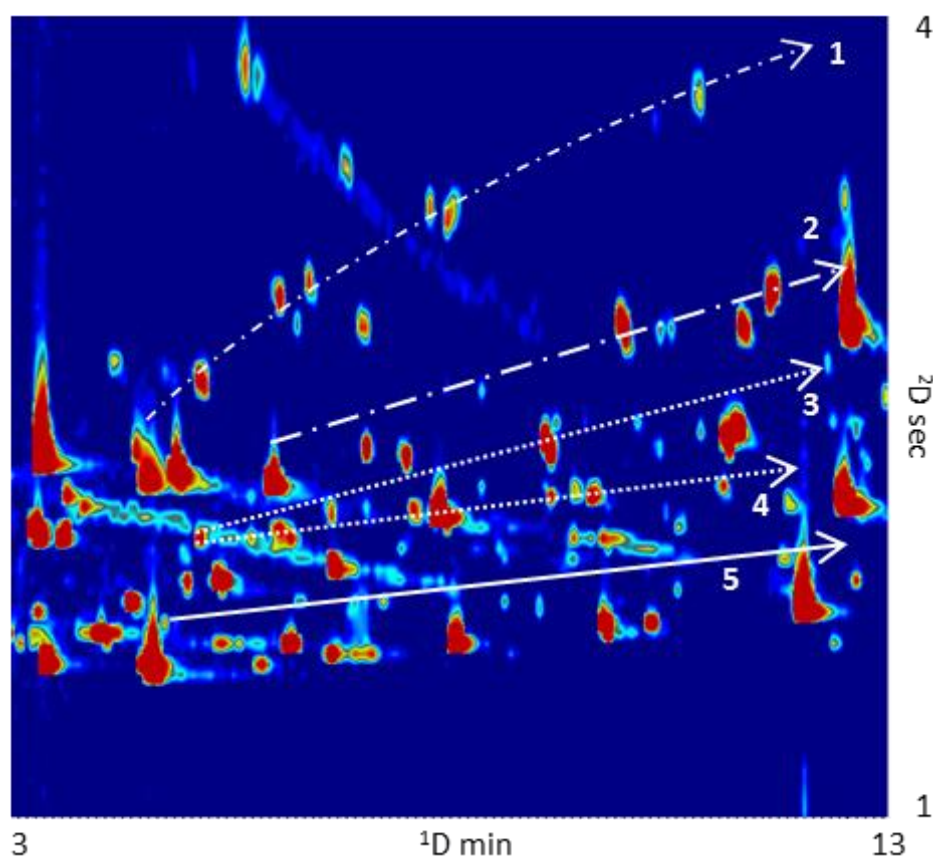


Figura 4.7. Distribución espacial de diferentes familias de compuestos orgánicos volátiles en el cromatograma, donde: (1) hidrocarburos ramificados insaturados; (2) ésteres; (3) aldehídos saturados; (4) aldehídos insaturados; (5) alcoholes.

Los compuestos identificados en el análisis dirigido son seleccionados con el objetivo de discriminar las muestras de aceite de oliva, en función del grado de madurez de la aceituna, pudiendo diferenciar con ellos los diferentes estados de maduración comprendidos entre noviembre 2014 y enero 2015, así como localizar indicadores específicos de la madurez en relación al atributo sensorial con el que se relacione.

Por lo tanto, se hace un análisis de componentes principales de la matriz de datos compuesta por las 48 muestras analizadas y el valor de área junto con el nombre de los 119 VOCs identificados para cada una de las muestras. Para la elaboración de PCA se usa la plataforma estadística XLSTAT.

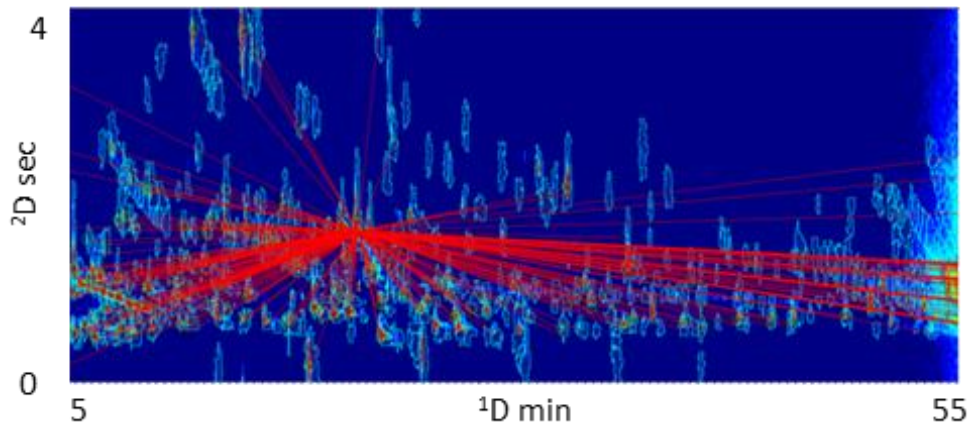


Figura 4.8. Relación entre el patrón α -tujona y el resto de compuestos orgánicos volátiles presentes en la muestra de aceite de oliva.

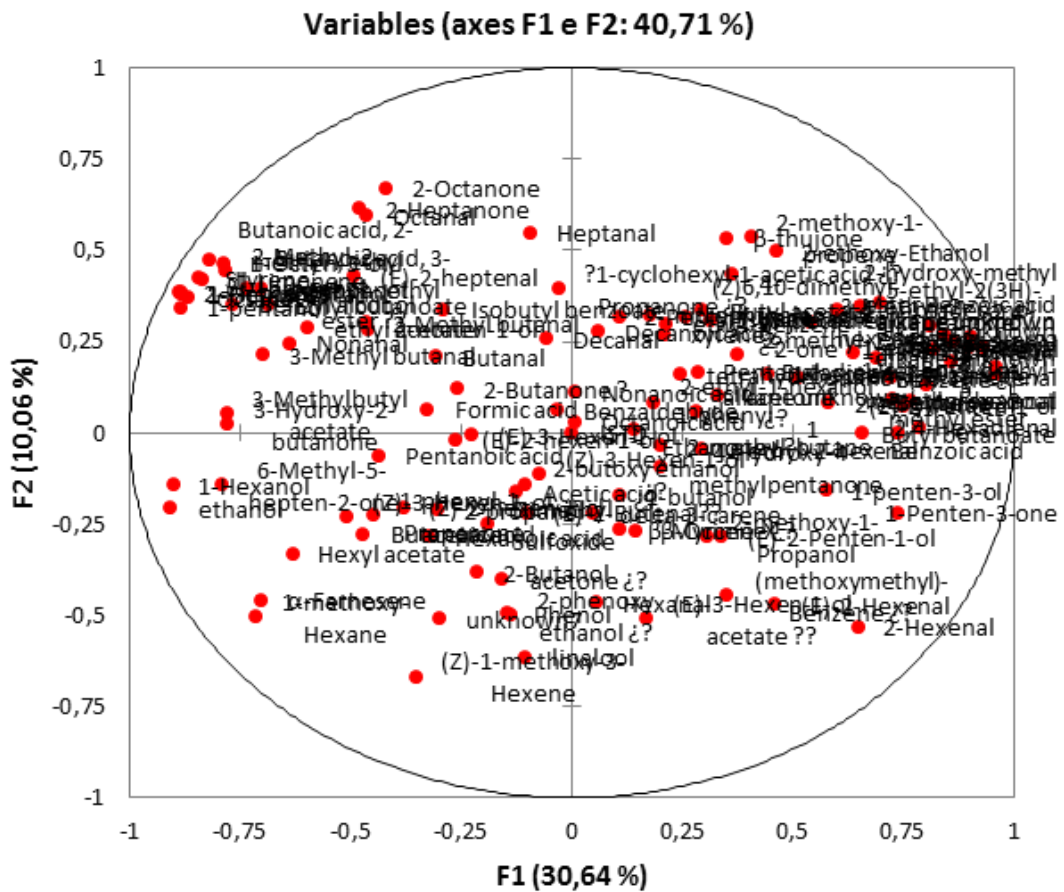


Figura 4.9. Scores del PCA de todos los compuestos identificados en todas las muestras de aceite de oliva analizadas.

Se obtiene el gráfico de *scores*, como el que se muestra en la figura 4.9, donde se observan todos los compuestos identificados distribuidos en todo el gráfico de las dos componentes principales seleccionadas.

A continuación y para poder seleccionar los compuestos responsables de cada uno de los estados de maduración o de su relación en función de la clasificación del AOVE, AOV o AOL, se llevan a cabo diferentes análisis de componentes principales.

En primer lugar, se elaboran los PCA para ver si existe agrupación por cada uno de los estados de maduración, en cada parcela, de las muestras de aceite de oliva. Así, por ejemplo, para el caso de la parcela de Caniles, son analizadas 8 muestras, con sus réplicas (un total de 16 análisis) y se hace el PCA con la matriz de dimensión 16×119, al igual que antes. La figura 4.10 muestra la gráfica de *scores* obtenidas, donde se observa la distribución de las muestras en función del estado de maduración en el que se recolectaron las aceitunas.

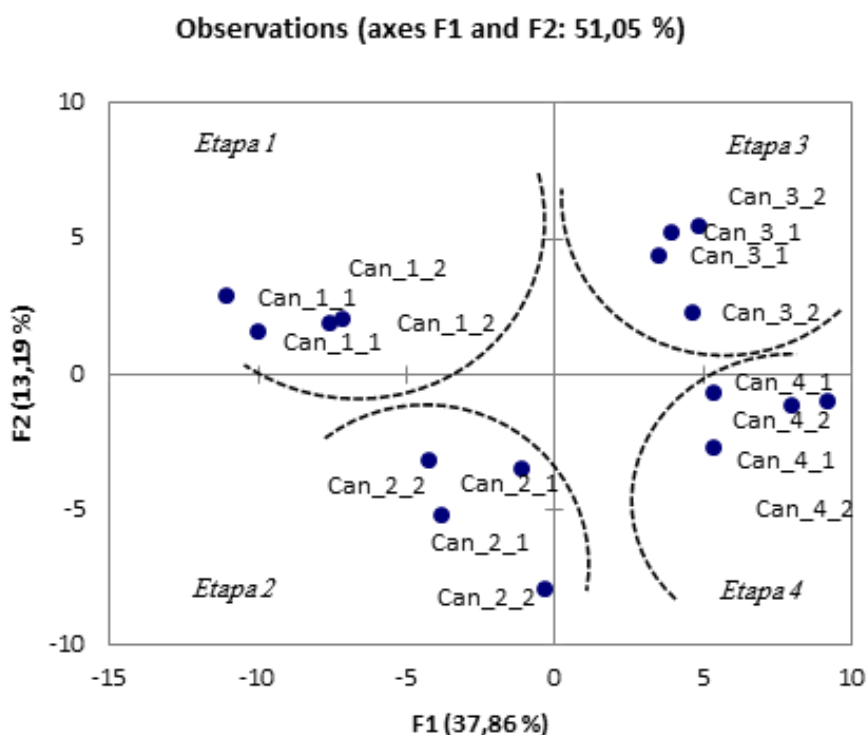


Figura 4.10. *Scores* del PCA de los aceites de oliva virgen extra de la parcela de Caniles. Cada etapa corresponde con la fecha en la que se recolectó la aceituna: 10-12 noviembre 2014 (etapa 1); 24-28 noviembre 2014 (etapa 2); 16-17 diciembre 2014 (etapa 3); 12-15 enero 2015 (etapa 4). En los códigos de la muestra aparece el nombre de la parcela (Can), la etapa de maduración (1-4) y la réplica (1,2).

En los PCA elaborados independientemente con las muestras de aceite de oliva de las tres parcelas se observa la agrupación de las muestras por estado de maduración. Así, se pueden seleccionar los compuestos orgánicos volátiles que aparecen en el gráfico de *scores*, en los extremos de cada agrupación por etapa de maduración.

Los compuestos que contribuyen a discriminar la etapa 1 (10-12 noviembre 2014) de la cosecha, son los compuestos: (Z)-2-hexenal, que son responsables de notas aromáticas de frutado; (Z)-3-hexenal, con olor a fruto verde; y (E,E)-2,4-hexadienal, que contribuyen con notas de frescor. Los siguientes compuestos contribuyen a la discriminación entre las etapas 1 y 2 (10-12 y 24-28 de noviembre 2014, respectivamente): (5Z)-3-etil-1,5-octadieno, (5E)-3-etil-1,5-octadieno, (E,E)-3,7-decadieno, y (E)-4,8-dimetil-1,3,7-nonatrieno. Todos estos compuestos decrecen con el estado de maduración en el que se recolecta la aceituna.

Teniendo en cuenta la clasificación de cada una de las muestras de aceite de oliva en una de sus tres categorías (ver tabla 4.1). En la parcela de Caniles se obtuvieron muestras de AOVE en todas sus etapas de maduración. En la parcela de Benamaurel son las muestras de aceite de oliva procedentes de las aceitunas de la cuarta etapa de maduración, las que se clasifican como AOV, y el resto son AOVE. Y en la parcela de Baza, se obtienen las muestras de aceite de oliva de las dos primeras etapas como AOVE, las de la siguiente etapa AOV, y la última AOL.

Por lo que, se hace un análisis de componentes principales con las muestras procedentes de las parcelas Baza y Benamaurel para ver si las muestras de aceite de oliva se agrupan en función a su categoría. En este caso, la matriz de datos está compuesta con las 8 muestras de aceite de oliva de la parcela de Baza, con sus replicados; así como las 8 muestras de aceite de oliva de la parcela de Benamaurel, con sus replicados, en los cuatro estados de maduración de la aceituna y con las áreas obtenidas de los 119 VOCs identificados. La matriz de datos es de 32×119. La figura 4.11 muestra la distribución obtenida.

En el gráfico de *scores* de los dos componentes principales, se observa la distribución de las muestras en relación con la clasificación del aceite de oliva.

A partir de esta distribución, se pueden obtener los VOCs relacionados con la clasificación del aceite de oliva, que se puede corroborar con los atributos positivos o defectos en bibliografía. Así, los volátiles de aldehídos saturados (heptanal, octanal y nonanal) e insaturados ((E)-2-heptenal) se relacionan con los defectos sensoriales del aceite de oliva. Además, la separación de los aceites lampantes, también se da por algunos alcoholes (propan-1-ol, 1-octen-3-ol, heptan-1-ol y octan-1-ol), cetonas (heptan-2-ona, octan-2-ona) y ésteres (acetato de etilo, etil-2-metil butanoato y etil-3-metil butanoato).

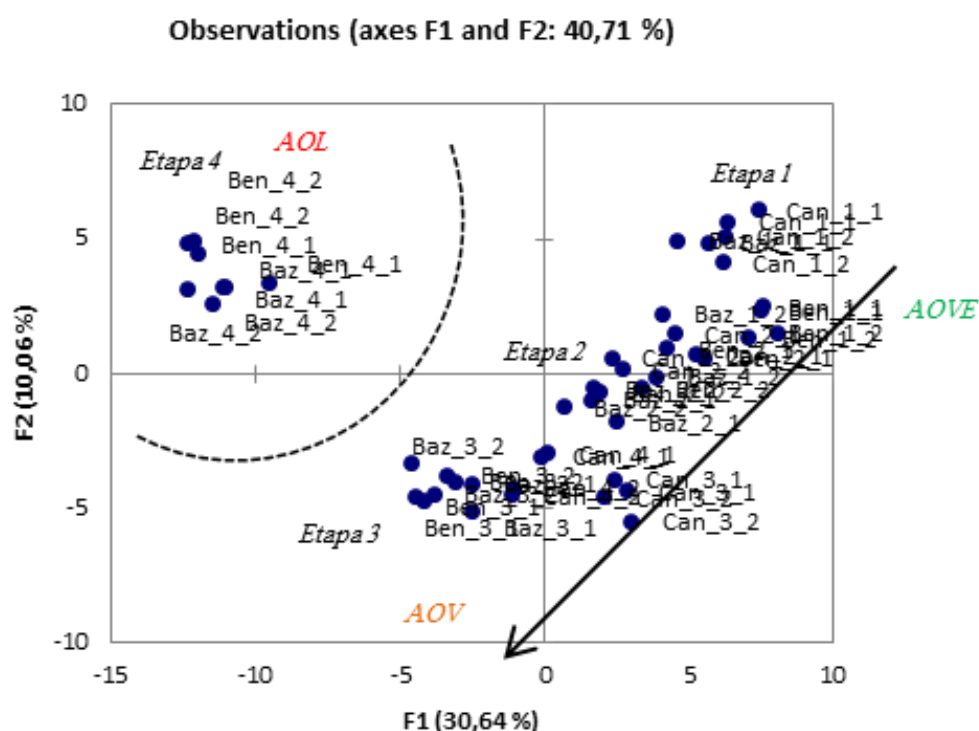


Figura 4.11. Scores del PCA de los aceites de oliva de la parcela de Baza y Benamaurel, en sus cuatro etapas de maduración de la aceituna: 10-12 noviembre 2014 (etapa 1); 24-28 noviembre 2014 (etapa 2); 16-17 diciembre 2014 (etapa 3); 12-15 enero 2015 (etapa 4). En los códigos de la muestra aparece el nombre de la parcela (Baz o Ben), la etapa de maduración (1-4) y la réplica (1,2). Así como: aceite de oliva virgen extra (AOVE), aceite de oliva virgen (AOV), aceite de oliva lampante (AOL).

II. Análisis no dirigido (*untargeted*)

El análisis no dirigido se lleva a cabo después de un análisis dirigido con el objetivo de poder encontrar más compuestos orgánicos volátiles, no identificados, en la muestra de aceite de oliva y que tengan relación con la

clasificación del aceite en función de su categoría o de la etapa de maduración en la que se recolecta la aceituna. Este análisis no dirigido se basa en la selección de las regiones de los picos en el cromatograma bidimensional, que se lleva a cabo automáticamente con el *software* GC Image Investigator R2.5. El análisis no dirigido incluye todas las regiones de los picos cromatográficos por encima de las 5000 cuentas de intensidad. En él se cargan todos los cromatogramas de cada una de las muestras de aceite de oliva, un total de 24 (más sus réplicas, 48 cromatogramas). En este caso, las regiones de los picos seleccionadas fueron 600, es el número total de picos cromatográficos bidimensionales, o manchas que el *software* seleccionó. En la figura 4.12 se muestra un zoom del mapa de calor obtenido para la muestra de Benamaurel, réplica 1, con el estado de maduración de la aceituna de enero de 2015. En la imagen se observa las regiones de los picos cromatográficos bidimensionales, con los VOCs identificados (que aparecen rodeados con un círculo).

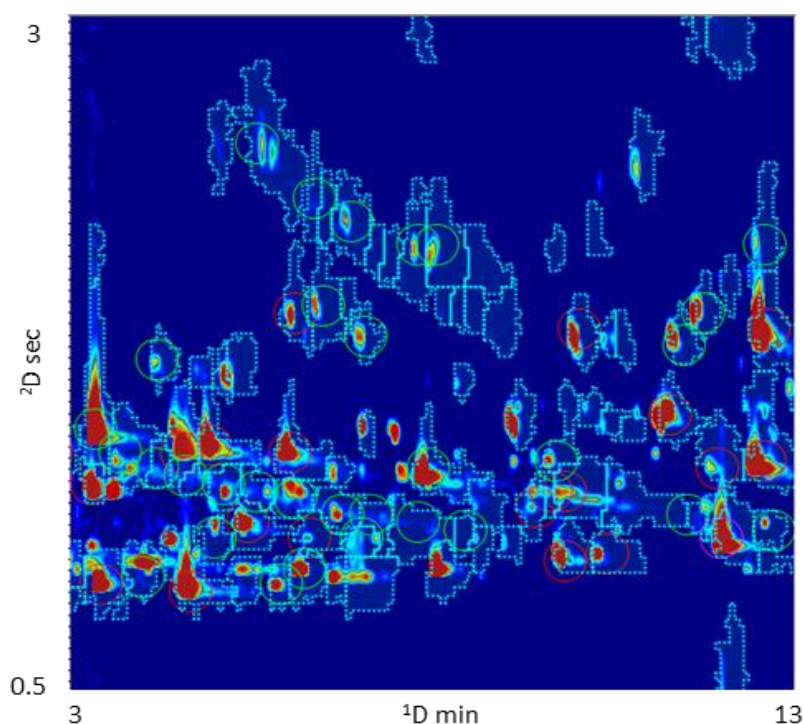


Figura 4.12. Cromatograma bidimensional de la muestra de aceite de oliva de la parcela de Benamaurel, en el estado de maduración de enero 2015. Aparece con recuadros las regiones de los picos cromatográficos seleccionados por el análisis no dirigido, y con círculos los picos cromatográficos que fueron identificados en el análisis dirigido.

Este tipo de análisis sirve para hacer una verificación cruzada, puesto que el análisis no dirigido señala los compuestos que han sido identificados en el análisis dirigido. Por lo tanto, este tipo de análisis sirve para corroborar los VOCs identificados, así como obtener información sobre compuestos no identificados que puedan tener relevancia en la investigación. El procedimiento que se lleva a cabo es el siguiente:

1. Detectar y registrar todos los picos obtenidos en el cromatograma bidimensional.
2. Localizar los picos cromatográficos desconocidos, así como verificar los picos identificados en el análisis dirigido. La verificación de estos picos la lleva a cabo el propio *software* indicando con círculos (como se muestra en la figura 4.12) los VOCs identificados en el mapa de calor del cromatograma bidimensional.
3. Alinear y combinar todos los cromatogramas para generar un cromatograma compuesto [19].
4. Define una región con cada uno de los picos cromatográficos detectados. Esto se lleva a cabo con el objetivo de extraer a través del software los valores de área de los picos cromatográficos seleccionados como desconocidos.
5. Crear y combinar el análisis dirigido y no dirigido en una única plantilla (*template*). Al generar un nuevo *template* con las muestras desconocidas, se puede aplicar a nuevas muestras que se analicen a posteriori, sin necesidad de volver a hacer el análisis no dirigido.

Con este análisis *untargete* se alinean todos los cromatogramas de todas las muestras y el *software* selecciona los compuestos desconocidos que coincidan en todas las muestras. En este caso, fueron detectados un total de **180 picos cromatográficos**, entre los identificados en el caso anterior (119) y los desconocidos tras hacer el análisis no dirigido.

Teniendo en cuenta estos resultados, se llevaron a cabo los mismos análisis de componentes principales que se realizaron en el análisis *targeted* anterior, pero

[19] C. Cordero, J. Kiefl, P. Schieberle, S.E. Reichenbach, C. Bicchi (2015). "Comprehensive two-dimensional gas chromatography and food sensory properties: potential and challenges", *Analytical and Bioanalytical Chemistry*, 407, 169-191.

teniendo en cuenta los 180 picos cromatográficos. Aparece la misma distribución que en el caso del análisis dirigido, lo que implica que los 61 compuestos desconocidos no son relevantes en esta investigación.

Los resultados dieron lugar a la publicación de un artículo científico, y se exponen en el siguiente apartado. En él se describe la selección de los compuestos orgánicos volátiles para la obtención de los índices de maduración de la aceituna, así como la estrategia *targeted* y *untargeted* usada.

4.5. ARTÍCULO

Combined untargeted and targeted fingerprinting with comprehensive two-dimensional chromatography for volatiles and ripening indicators in olive oil

(Artículo publicado en Julio 2016 en la revista: *Analytica Chimica Acta*)



Combined untargeted and targeted fingerprinting with comprehensive two-dimensional chromatography for volatiles and ripening indicators in olive oil



Federico Magagna ^{a,1}, Lucia Valverde-Som ^{b,1}, Cristina Ruíz-Samblás ^b,
Luis Cuadros-Rodríguez ^b, Stephen E. Reichenbach ^c, Carlo Bicchi ^a, Chiara Cordero ^{a,*}

^a Dipartimento di Scienza e Tecnologia del Farmaco, Università di Torino, Turin, Italy

^b Department Analytical Chemistry, Faculty of Science, University of Granada, C/Fuentenueva s/n, 18071, Granada, Spain

^c University of Nebraska – Lincoln, Lincoln, NE 68588-0115, USA

Abstract

Comprehensive two-dimensional gas chromatography (*GC×GC*) is the most effective multidimensional separation technique for in-depth investigations of complex samples of volatiles (VOC) in food. However, each analytical run produces dense, multi-dimensional data, so elaboration and interpretation of chemical information is challenging.

This study exploits recent advances of *GC×GC-MS* chromatographic fingerprinting to study VOCs distributions from Extra Virgin Olive Oil (EVOO) samples of a single botanical origin (Picual), cultivated in well-defined plots in Granada (Spain), and harvested at different maturation stages. A new integrated work-flow, fully supported by dedicated and automated software tools, combines *untargeted* and *targeted (UT)* approaches based on peak-region features to achieve the most inclusive fingerprinting.

Combined results from *untargeted* and *targeted* methods are consistent, reliable, and informative on discriminant features (analytes) correlated with optimal ripening of olive fruits and sensory quality of EVOOs. The great flexibility of the *UT fingerprinting* here adopted enables retrospective analysis with great confidence and provides data to validate the transferability of ripening indicators ((*Z*)-3-hexenal, (*Z*)-2-hexenal, (*E*)-2-pentenal, nonanal, 6-methyl-5-hepten-2-one, octane) to external samples sets. Direct image comparison, based on *visual* features, also is investigated for quick and effective pair-wise investigations. Its implementation with reliable metadata generated by *UT fingerprinting*

confirms the maturity of 2D data elaboration tools and makes advanced image processing a real perspective.

Key-words:

Comprehensive two-dimensional gas chromatography-mass spectrometry; untargeted and targeted fingerprinting; extra virgin olive oil; olives ripening; retrospective investigations.

1. Introduction

Comprehensive two-dimensional gas chromatography (*GC×GC*) is the most effective multidimensional separation technique for in-depth investigations of complex samples of volatiles in food [1]. The combination, in a single analytical platform, of two separation dimensions with mass spectrometric detection and, when possible, automated sample preparation, delivers highly efficient sample profiling (detailed analysis of single molecular entities) and fingerprinting (rapid, high-throughput screening of samples for distinctive analytical signatures) [2].

Each analytical run produces dense, multi-dimensional data, so elaboration and interpretation of chemical information is a challenging task. In addition, food samples generally have a high-degree of chemical multidimensionality [3] thus creating highly complex analytical challenges. In this context, data elaboration strategies should implement smart and productive processes, preferably with a high degree of automation, to make cross-samples analysis efficient and informative.

Within the existing methodologies for *GC×GC* data elaboration [4,5], the approach based on *peak-region* features has been very effective because of its comprehensive and uniform treatment of information from each sample constituent, both knowns and unknowns. Each single chemical entity is characterized by its chromatographic and spectrometric parameters (retention time in both dimensions, detector response, and mass spectral information) and by its absolute and relative position within the pattern of all detectable constituents. As a consequence, the 2D peak-retention pattern of a sample is a diagnostic fingerprint, informative of its composition; and pattern recognition approaches can be successfully applied to improve effectiveness and productivity in multi-sample data elaboration.

Although these concepts are not new for the *GC×GC* community [6], the full automation of these procedures and their implementation in commercial software packages has been achieved only recently. This has limited both routine adoption of the technique for food analysis and investigative strategies for profiling [2,7].

Analysis of olive oil volatiles is a challenging and important problem and *GC×GC* can yield deeper knowledge of the composition of this fraction offering new perspectives for quality and authenticity assessment [8].

In spite of the great potential of *GC×GC*, few studies are available in this field. Vaz Freire et al. [9] first proposed an image-features approach, or more generally a pattern recognition methodology, to investigate the characteristic distribution of volatiles from oils. They adopted open-

source image analysis software (Image J, National Institutes of Health) to extract information from small 2D regions located over the separation space and, by Principal Component Analysis (PCA), selected those regions with the highest discrimination potential. Then, they used targeted profiling to locate known analytes within informative 2D regions.

In 2010, Cajka and co-workers [10] exploited the targeted profiling potential of GC×GC-ToF-MS and identified 44 analytes able to discriminate samples of different geographical origin and production year. More recently, Purcaro et al. [8] combined targeted and untargeted analysis with the goal of a chemical blueprint of olive oil aroma defects. This inter-laboratory study confirmed the reliability of GC×GC for detailed profiling of olive oil volatile fractions and introduced an iterative strategy [11,12] to locate sensory-relevant analytes efficiently.

This study exploits the most recent advances of GC×GC-MS chromatographic fingerprinting to study VOC distributions from Extra Virgin Olive Oil (EVOO) samples of a single botanical origin (Picual), cultivated in well-defined plots in a single region (Granada, Spain), and harvested at different maturation stages. The principal interest in this application is the quality characteristics related to optimal ripening of olive fruits [13,14,15,16,17,18,19,20,21] and, as a consequence, olive oil classification and perceivable sensory quality [22,23]. In particular, this study proposes an integrated work-flow, fully supported by dedicated software tools, that performs cross-samples comparisons by contemporarily considering characteristic distributions (i.e., sample fingerprints) of both known and unknown compounds. This work-flow

integrates both *untargeted* and *targeted (UT) fingerprinting* to realize the most comprehensive results, and so is termed *UT fingerprinting*. Challenges of retrospective analysis and immediacy of image fingerprinting also are discussed because of the advantages they offer in specific investigations.

2. Materials and methods

2.1. Reference compounds and solvents

Pure reference standards of α -thujone, used as Internal Standard (ISTD), at a concentration of 100 mg/L in dibutyl phthalate, and *n*-alkanes (*n*-C₉ to *n*-C₂₅), used for linear retention index (I^T_S) determination, at a concentration of 100 mg/L in cyclohexane, were supplied by Sigma-Aldrich (Milan, Italy).

Solvents for *n*-alkanes dilution (toluene and cyclohexane HPLC-grade) and dibutyl phthalate also were from Sigma-Aldrich.

2.2. Olive oil samples

Olive oil samples of *Picual* variety, harvested in 2014, were supplied by "GDR Altiplano de Granada" (Spain) and were obtained from olives harvested in three different plots in Granada: "812 *Caniles*" (organic production and drip irrigation); "233-234 *Bazd*" (conventional production

and drip irrigation); and "701 *Benamaurel*" (conventional production and drip irrigation).

Each sample was available in duplicate and obtained by mixing olives from at least five different trees in the same plot to have homogeneous and representative samples. Olives were harvested at four different ripening stages: November 10-12, 2014; November 24-28, 2014; December 16-17, 2014; and January 12-15, 2015.

Samples were analyzed by an accredited laboratory to define quality parameters: acidity (%), peroxide index (mEq O₂/kg), and UV absorption. Samples also were submitted to sensory evaluation by a recognized/official panel [24]. Sample descriptions and acronyms are reported in Table 1, together with quality assessments.

2.3. Head-space Solid Phase Micro Extraction sampling devices and conditions

Volatiles were sampled from the headspace (HS) by HS Solid Phase Micro Extraction (HS-SPME). The sampling protocol was optimized in a previous study [8] and employs a divinylbenzene/carboxen/polydimethylsiloxane (DVB/CAR/PDMS) 50/30 μm , 2 cm length stableflex fiber from Supelco (Bellefonte, PA, USA).

The ISTD (α -tujone) was pre-loaded onto the fiber before sampling through the standard-in-fiber procedure. An ISTD solution, 2.0 μL , was placed into a 20 mL glass vial and submitted to HS-SPME at 50°C for 15 minutes (min). The fiber then was exposed to the head-space of olive oil

samples (1.500 g exactly weighted) in 20 mL glass vials, at 50°C for 40 minutes. Last, the sampled analytes were recovered by introducing the fiber into the S/SL injection port of the GCxGC system at 260°C and thermally desorbed for 5 minutes. Each sample was analyzed in duplicate.

Table 1: List of analyzed samples together with plot denomination, field replicate, harvest stage, acronym, quality parameters according to COMMISSION REGULATION (EEC) No 2568/91 of 11 July 1991, sensory evaluation results, and commercial classification.

Plot	Field replicate	Harvest stage	Sample Acronym	Acidity (%)	Peroxide	K ₂₃₂	K ₂₇₀	ΔK	Md [§]	Mf [§]	Classification
					index (mEq O ₂ /kg)						
Caniles	1	1	Can_1_1	0.2	4	2.02	0.2	0	0	7.9	EVOO
Caniles	1	2	Can_2_1	0.1	4	1.61	0.2	0	0	3.5	EVOO
Caniles	1	3	Can_3_1	0.3	4	1.41	0.2	0	0	3.6	EVOO
Caniles	1	4	Can_4_1	0.2	4	1.40	0.1	0	0	2.8	EVOO
Caniles	2	1	Can_1_2	0.2	4	2.23	0.2	0	0	6	EVOO
Caniles	2	2	Can_2_2	0.2	4	1.71	0.2	0	0	3	EVOO
Caniles	2	3	Can_3_2	0.3	4	1.38	0.2	0	0	3.4	EVOO
Caniles	2	4	Can_4_2	0.3	4	1.41	0.2	0	0	2.8	EVOO
Baza	1	1	Baz_1_1	0.2	5	1.84	0.2	0	0	5	EVOO
Baza	1	2	Baz_2_1	0.2	3	1.60	0.2	0	0	4.1	EVOO
Baza	1	3	Baz_3_1	0.2	5	1.17	0.2	0	> 0.00	1.3	VOO
Baza	1	4	Baz_4_1	0.4	11	1.11	0.1	0	> 0.00	0	LOO
Baza	2	1	Baz_1_2	0.2	4	1.92	0.2	0	0	5.2	EVOO
Baza	2	2	Baz_2_2	0.1	3	1.65	0.2	0	0	3.8	EVOO
Baza	2	3	Baz_3_2	0.2	6	1.28	0.1	0	> 0.00	1.7	VOO
Baza	2	4	Baz_4_2	0.4	13	1.12	0.1	0	> 0.00	0	LOO
Benamaurel	1	1	Ben_1_1	0.2	5	1.61	0.2	0	0	4.4	EVOO
Benamaurel	1	2	Ben_2_1	0.2	4	1.53	0.2	0	0	4.3	EVOO
Benamaurel	1	3	Ben_3_1	0.2	8	1.19	0.1	0	0	3.1	EVOO
Benamaurel	1	4	Ben_4_1	0.4	19	1.05	0.1	0	> 0.00	0	LOO
Benamaurel	2	1	Ben_1_2	0.1	4	1.64	0.2	0	0	4.2	EVOO
Benamaurel	2	2	Ben_2_2	0.2	3	1.48	0.2	0	0	4.3	EVOO
Benamaurel	2	3	Ben_3_2	0.2	6	1.51	0.1	0	0	2.9	EVOO
Benamaurel	2	4	Ben_4_2	0.2	14	1.05	0.1	0	> 0.00	0	LOO
Retrospective analysis - sampling from 2013 production											
Spain - blend	-		R-EVO 1	-	-	-	-	-	0	4.1	EVOO
PDO Monti Iblei - Sicily Italy			R-EVO 2	-	-	-	-	-	0	5.1	EVOO
PDO Dauno Gargano Puglia Italy			R-EVO 3	-	-	-	-	-	0	4.9	EVOO
Tuscany Italy - blend			R-EVO 4	-	-	-	-	-	0	4.1	EVOO
Tuscany Italy - blend			R-EVO 5	-	-	-	-	-	0	3.0	EVOO

§: Median of defects (Md) - Median of fruity note (Mf)

2.4. Comprehensive two-dimensional gas chromatographic system (GC×GC-MS) set-up and analysis conditions

GC×GC analyses were performed on an Agilent 6890 unit coupled to an Agilent 5975C MS detector (Agilent, Little Falls, DE, USA) operating in EI mode at 70eV. The GC transfer line was set at 270°C and the MS scan range was 40-240 m/z with a scanning rate of 12,500 amu/s to obtain a spectra generation frequency of 30 Hz.

The system was equipped with a two-stage KT 2004 loop-type thermal modulator (Zoex Corporation, Houston, TX) cooled with liquid nitrogen and with the hot jet pulse time set at 250 ms with a modulation time of 4 s for all experiments. Fused silica capillary loop dimensions were 1.0 m long and 0.1 mm inner diameter. The column set was configured as follows: ¹D SolGel-Wax column (100% polyethylene glycol)(30 m × 0.25 mm d_c , 0.25 μm d_f) from SGE Analytical Science (Ringwood, Australia) coupled with a ²D OV1701 column (86% polydimethylsiloxane, 7% phenyl, 7% cyanopropyl) (1 m × 0.1 mm d_c , 0.10 μm d_f) from Mega (Legnano, Milan, Italy).

Fiber thermal desorption into the GC injector port was under the following conditions: split/splitless injector in split mode, split ratio 1:20, injector temperature 250°C. Carrier gas was helium at a constant flow of 1.8 mL/min. The temperature program was: from 40°C (1 min) to 200°C at 3°C/min and to 250°C at 10°C/min (5 min).

The *n*-alkanes liquid sample solution for I_S^T determination was analyzed under the following conditions: split/splitless injector in split mode, split ratio 1:50, injector temperature 280°C, injection volume 1 μ L.

2.5. Raw data acquisition and GC \times GC data handling

Data were acquired by Agilent MSD ChemStation ver E.02.01.00 and processed using GC Image GC \times GC Software version 2.5 (GC Image, LLC Lincoln NE, USA). Statistical analysis was performed by XLstat (Addinsoft, New York, NY USA) and the PLS Toolbox (Eigenvector Research Inc., West Eaglerock Drive, Wenatchee, WA, USA) for Matlab[®] software (The Mathworks Inc., Natick, MA, USA).

2.6 Profiling and advanced fingerprinting work-flow

The bi-dimensional chromatographic data elaboration proposed here was organized in a sequential work-flow illustrated in Figure 1.

Untargeted and targeted analyses were performed by applying the *template matching* fingerprinting strategy, introduced by Reichenbach et al. in 2009 [6]. It uses the patterns of 2D peaks' metadata (retention times, MS fragmentation patterns, and detector responses) to establish reliable correspondences between the same chemical entities across multiple chromatograms. The output of template matching fingerprinting is a data matrix of aligned 2D peaks and/or peak-regions, together with their related metadata (¹D and ²D retention times, compound names for

target analytes, fragmentation pattern, single ions or total ions response), that can be used for comparative purposes.

Targeted analysis (Step 1 of Figure 1) focused on about 120 selected compounds, each reliably identified by matching their EI-MS fragmentation pattern (NIST MS Search algorithm, ver 2.0, National Institute of Standards and Technology, Gaithersburg, MD, USA, with Direct Matching threshold 900 and Reverse Matching threshold 950) with those collected in commercial (NIST2014 and Wiley 7n) and in-house databases. As a further parameter to support reliable identification, Linear Retention Indices (*ITS*) were considered and experimental values compared with tabulated ones.

Untargeted analysis (Step 2 of Figure 1) was based on a *peak-regions* features approach [5] and was performed automatically by GC Image Investigator™ R2.5 (GC-Image LLC, Lincoln NE, USA). The untargeted analysis included all *peak-regions* above the arbitrarily fixed peak response threshold of 5,000 counts together with target peaks from Step 1. This approach [25,26,27,28], briefly described in Section 3.2, re-aligned the 48 chromatograms using a set of *registration peaks*. The resulting data matrix was a 48 × 600 (samples × reliable *peak-regions*). Response data from aligned 2D *peak-regions* were used for PCA and results cross-compared to those obtained from target peaks distributions. (See Section 3.2 for the discussion of results.)

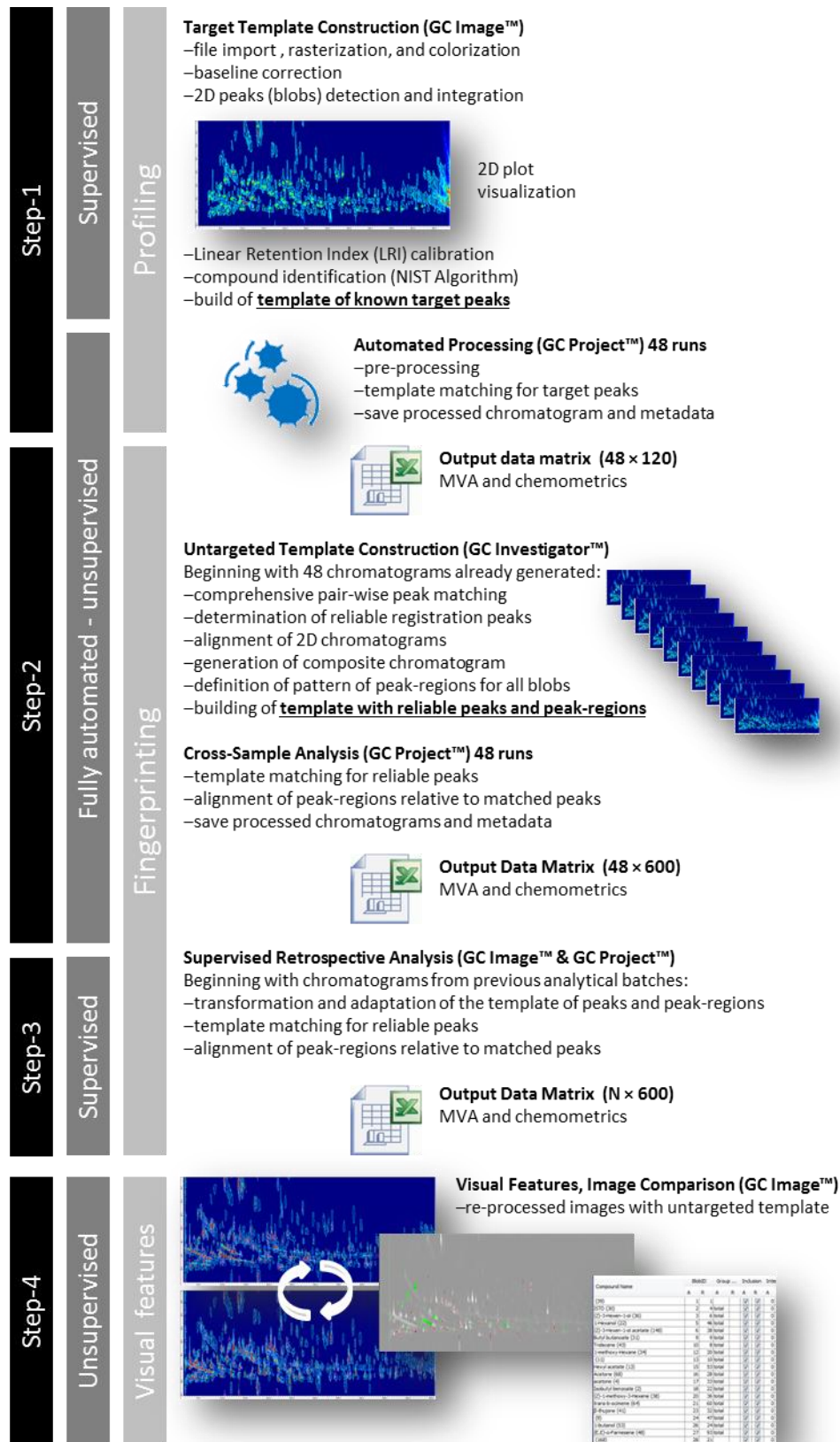


Figure 1: Two-dimensional chromatographic data elaboration work-flow.

Visual features fingerprinting, performed as pair-wise image comparison, was the last step of the study (Step 4 of Figure 1) and was rendered with "colorized fuzzy ratio" mode [cite Hollingsworth et al., JoCA 1105:51, 2006]. The algorithm computes the difference at each data point between pairs of TICs; a data point is the output of the detector at a point in time. These differences are mapped into Hue-Intensity-Saturation (HIS) color space to create an image for visualizing the relative differences between image pairs in the retention-times plane [29]. A detailed description is provided in *Section 3.4*.

3. Results and discussion

The goal of this study was to evaluate the potential of combining Untargeted and Targeted 2D data elaboration approaches based on untargeted *peak-region* features, target peaks, and *visual* features to approach to the most inclusive fingerprinting within EVOO volatiles: the UT fingerprinting strategy. Based on a sampling design focused on a single botanical variety and well-defined geographical locations, VOCs fingerprints were interpreted as a function of ripening stage and oil quality.

This strategy was inspired by a previous study focused on olive oil aroma defects [8], in which results clearly indicated that the informative potential of GC×GC-MS to delineate specific fingerprints for sensory quality classification of oils. These results showed that a "*strictly structured experimental design (considering more variables, such as*

cultivar, geographical origin, etc.) would be mandatory to “*robustly and reliably characterize specific markers and related characteristics concentration windows*” to support, or even replace, sensory evaluation [8]. In addition, it was clear that larger numbers of “external variables” affecting VOCs pattern reduce the effectiveness of untargeted approaches.

In the present study, with fewer sample-set variables, and a data elaboration process that combines untargeted and targeted approaches (i.e., *peak-region features, peaks, and visual features* methods), we achieve highly effective fingerprinting. The proposed work-flow is comprehensive yet efficient and fully supported by new commercial software. In addition, we validate both the data elaboration strategy and the informative role of some targets by a retrospective investigation on VOCs patterns from EVOO, VOO, and *lampante* oils (LOO) analyzed in previous studies.

Following this scheme, we first present and discuss results from targeted analysis, focusing on peaks for known informative chemicals and selected VOCs strictly related to the olives' geographical location, ripening stage, and product (oils) quality (presence/absence of sensory defects). Next, untargeted analysis based on peak-regions features, implemented in the second step, is discussed from the perspective of: (a) confirming sample classification results; (b) indicating new potential targets; (c) defining chemical indexes of ripening and quality through the ratio between informative analytes; and (d) validating the role of informative ratios through retrospective elaboration of samples analyzed in previous studies

by adopting the *UT* template created on the current sample set. The last part of the study aims at determining if classification based on peak-regions features could be replaced by direct image comparison without losing information about the chemical composition of this fraction. The following paragraphs illustrate the research steps and critically discuss results.

3.1 Targeted analysis and samples discrimination

The sample set is reported in Table 1 with their quality parameters, sensory evaluation results, and commercial classification. Quality metrics (acidity %, peroxide index, UV absorbance, and organoleptic assessment) indicated that 6 of the 24 samples were not compliant with *Extra-Virgin* classification [30]. These samples, classified as *Virgin* (VOO) and *Lampante* (LOO), were from the late ripening stages of the *Baza* and *Benamaurel* plots. This quality classification was confirmed by replicate sampling (i.e., *Baz_3_1/_2* and *Baz_4_1/_2*; *Ben_4_1/_2*) and was related to sensory defects revealed by the panel (Median of defects - Md >0.00). In addition, low-quality samples were connoted by a higher peroxide index and acidity %.

From the available literature [31,32,33,34,35,36,37,38,39], analytes detected in the *GC×GC* data were identified by their EI-MS fragmentation pattern and Linear Retention Indices (I_S^T) see section 2.6 for details). Following the work-flow in Figure 1, *template-matching*

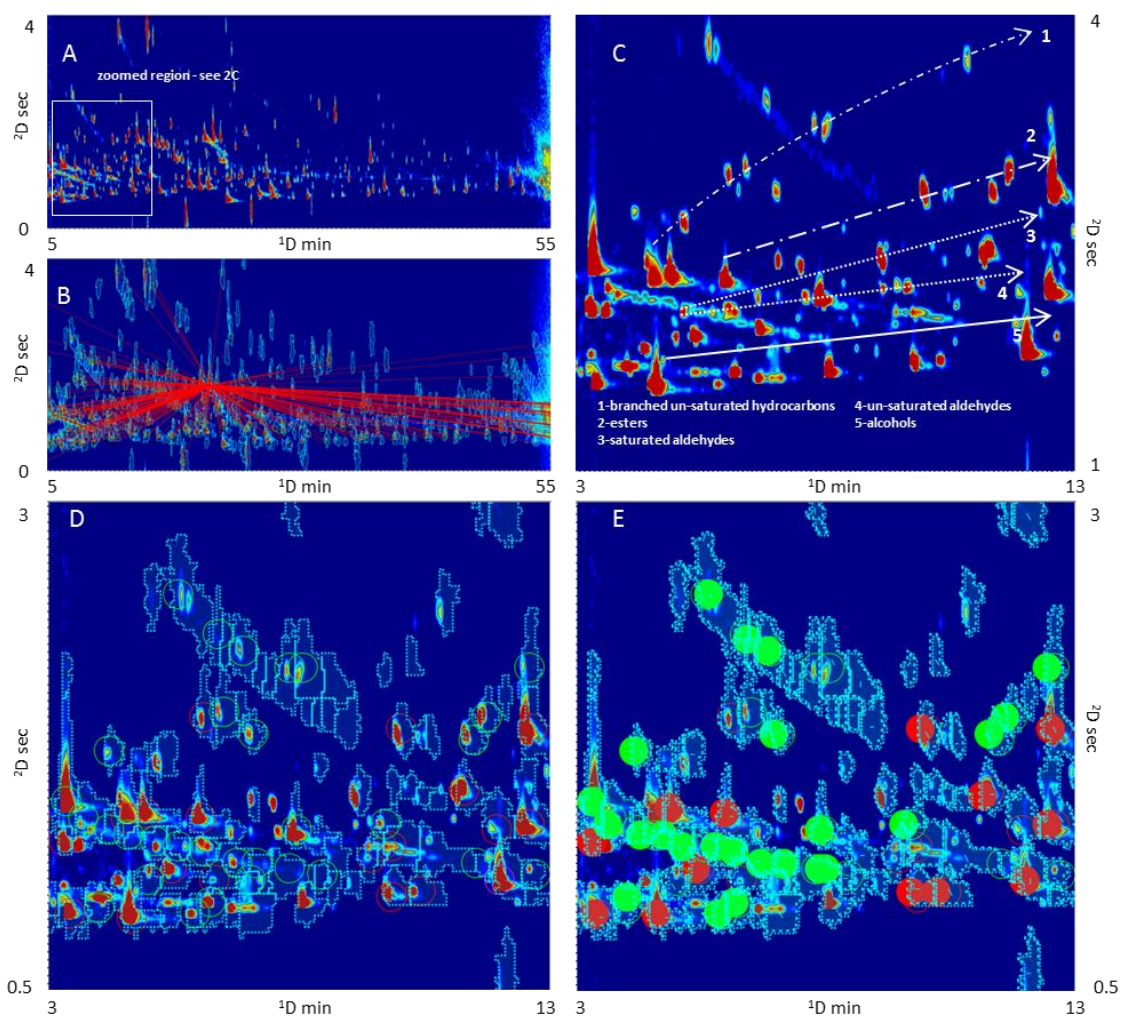
fingerprinting (see Paragraph 2.6) with 119 target peaks was used to map these informative chemicals across samples.

Figures 2A-B shows the pseudocolored GC×GC chromatogram of an EVOO sample from the *Benamaurel* plot harvested at stage 4 (in January 2015). Figure 2B locates the 119 known target peaks (empty light green circles) linked to the ISTD (α -tujone black circle) by red lines.

The quali-quantitative distribution of VOCs changed with the harvest stages. The number of detectable peaks above a Volume threshold of 4,000 (arbitrarily fixed on the Total Ion Current signal) was about 270-280 at the first stage and reached about 360 at the final stage (data not shown).

The effectiveness of GC×GC, in both peak-capacity and overall chromatographic resolution, plays a critical role in isolating the information for compounds with similar retention times in the ¹D dimension. A zoomed region, highlighted in Fig 2C, emphasizes the retention area of highly volatile compounds in which some branched hydrocarbons (eluting later in the ²D) are separated along the ²D from saturated and unsaturated and carbonyl compounds (e.g., pentanal, hexanal, (*E*)-2-butenal), and 1-penten-3-one, an odor-active volatile deriving from linolenic acid degradation), and alcohols (e.g., 1-propanol, 2-butanol, and 2-methyl-2-propanol).

The relative distribution (Normalized 2D Peak Volumes) of the 119 peaks is illustrated as heat-map in Supplementary Figure 1 (SF1).



Figures 2A-E: (2A) Pseudocolored $GC \times GC$ chromatogram of *Ben_4_1* harvested at stage 4 (in January 2015). (2B) Position of the 119 known target peaks (empty light green circles) linked to the ISTD (α -tujone black circle) by red lines. (2C) Retention area of highly volatile compounds referred to the white rectangle of Fig. 2A. (2D) Peak-regions delineated by light blue graphics together with targeted peaks (empty light blue circles). (2E) Results of comprehensive template matching for peak-regions, target peaks (green circles) and registration peaks (red circles). For details see text.

Columns are ordered left-to-right by 1D retention indices (polar phase column, 100% polyethylene glycol). The logarithmic colour map is based on $2D$ -Peak Volumes divided by ISTD response and is normalized by dividing single values by row standard deviations.

Table 2 reports the 119 target compounds together with their ^1D and ^2D retention times, I_S^T sensory descriptors, and correlation with oil defects as reported in reference literature [31,32,33,34,35,36,37,38,39]. 2D Peak Normalized Volumes (average values of two analytical replicates) are provided as Supplementary information (**Supplementary Table 1- ST1**).

The target analytes distribution (Normalized 2D volumes) was adopted as an informative fingerprint for possible discrimination of samples within different harvest stages and, in particular, to locate and validate specific indicators of ripeness, and, when feasible, odor-active compounds related to sensory quality.

Table 2: List of the 119 target analytes together with ^1D and ^2D retention times, I_S^T and sensory descriptors as reported in reference literature [31,32,33,34,35,36,37,38,39]. The 2D Peak Volume data is provided as Supplementary information in Supplementary Table 1 - ST1.

ID	Compound Name	1D (min)	2D (sec)	Exp ITS	Odour descriptor
1	1,4-Pentadiene	2.69	1.08	870	-
2	Octane	3.29	0.97	893	Solvent, unpleasant
3	Acetone	3.40	0.53	897	Pungent
4	Ethyl acetate	4.02	0.59	920	Fruity, aromatic, ethereal,
5	3-Methyl butanal	4.42	0.69	935	Malty
6	2-Propanol	4.61	1.11	942	-
7	Ethanol	4.68	0.49	945	Ethanol-like
8	1-Methoxy-Hexane	4.95	1.14	955	-
9	2-Ethylfuran	4.98	0.70	956	-
10	3,4-Diethyl-1,5-hexadiene (RS+SR) [§]	5.19	1.44	964	-
11	3,4-Diethyl-1,5-hexadiene (meso) [§]	5.35	1.41	970	-
12	Pentanal	5.47	0.79	974	Almond-like, pungent, malt
13	2-Butanone	5.49	0.59	975	Ethereal, fruity
14	(Z)-1-Methoxy-3-Hexene	6.08	1.11	997	-
15	(5Z)-3-Ethyl-1,5-octadiene [§]	6.14	1.73	1000	-
16	1-Penten-3-one	6.37	0.73	1008	Green
17	(5E)-3-Ethyl-1,5-octadiene [§]	6.55	1.77	1015	-

§: Ref.[40]

Table 2 (continued).

ID	Compound Name	1D (min)	2D (sec)	Exp ITS	Odour descriptor
18	1-Propanol	6.69	0.55	1020	Pungent, pineapple
19	Ethyl butanoate	6.69	1.00	1020	Fruity
20	Tetrahydro-Furan	6.98	0.55	1031	-
21	(<i>E</i>)-2-Butenal	6.99	0.66	1031	-
22	Ethyl 2-methylbutyrate	7.14	1.22	1037	Fruity
23	Ethyl 3-methylbutyrate	7.55	1.21	1053	Fruity
24	Butyl acetate	7.74	1.01	1059	Pear
25	(<i>E,Z</i>)-3,7-Decadiene	7.91	2.01	1066	-
26	Hexanal	8.03	1.01	1070	Tallowy, leaf-like
27	2-Methyl-2-propanol	8.17	0.59	1076	-
28	(<i>E,E</i>)-3,7-Decadiene	8.21	2.04	1077	-
29	3-Methylbutyl acetate	9.28	1.24	1117	-
30	(<i>E</i>)-2-Pentenal	9.54	0.87	1127	Pungent, apple-like
31	Xylene	9.82	1.08	1137	Plastic
32	(<i>Z</i>)-3-Hexenal	9.87	0.93	1139	Leaf-like
33	1-Butanol	9.98	0.63	1144	Winey
34	(<i>E</i>)-3-Hexenal	10.01	0.93	1145	-
35	Butyl 2-methylpropanoate	10.19	1.63	1151	-
36	ethyl-Benzene	10.33	1.33	1157	-
37	1-Penten-3-ol	10.53	0.64	1164	Pungent
38	2-Heptanone	11.41	1.26	1197	Sweet, fruity
39	Heptanal	11.50	1.27	1201	Oily, fatty, woody
40	2-Ethyl-hexanal	11.60	1.59	1204	-
41	Limonene	11.92	1.73	1216	Citrus, mint
42	(<i>Z</i>)-2-Hexenal	12.17	1.04	1225	Fruity
43	3-Methyl-1-butanol	12.28	0.69	1230	-
44	(<i>E</i>)-2-Hexenal	12.75	1.07	1211	Bitter almond, green
45	Butyl butanoate	12.82	1.65	1213	-
46	δ -3-carene	13.46	1.59	1231	-
47	Tridecene	13.87	3.77	1243	-
48	1-Pentanol	13.98	0.69	1246	Sweet, pungent
49	Styrene	14.09	0.93	1249	Balsamic, gasoline
50	Trans-b-ocimene	14.09	1.63	1249	-
51	Hexyl acetate	14.99	1.50	1274	Fruity
52	3-Hydroxy-2-butanone	15.19	0.66	1280	Buttery
53	2-Octanone	15.47	1.46	1288	Mould, green
54	Octanal	15.64	1.48	1293	Fatty, sharp
55	2-Ethyl-2-hexenal	15.99	1.47	1303	-
56	(<i>E</i>)-4,8-Dimethyl-1,3,7-nonatriene	16.40	1.86	1301	-
57	(<i>E</i>)-2-Penten-1-ol	16.47	0.64	1303	Mushroom
58	1-Penten-3-ol	16.49	0.66	1303	-
59	(<i>Z</i>)-3-Hexen-1-ol acetate	16.73	1.29	1309	Sweet
60	(<i>E</i>)-3-Hexen-1-ol acetate	16.80	1.26	1311	-

§: Ref.[40]

Table 2 (continued).

ID	Compound Name	1D (min)	2D (sec)	Exp ITS	Odour descriptor
61	(Z)-2-Penten-1-ol	16.81	0.66	1311	Green, plastic, rubber
62	(E)-2-heptenal	17.04	1.22	1317	Fatty, almond-like
63	6-Methyl-5-hepten-2-one	17.56	1.22	1330	Pungent, green
64	1-Hexanol	18.38	0.75	1351	Fruity, banana, soft
65	4-Hydroxy-4-methylpentanone	18.50	0.78	1354	-
66	(E)-3-Hexen-1-ol	18.62	0.72	1357	-
67	cis- β -Ocimene	19.04	1.59	1367	Sweat, herb
68	(Z)-3-Hexen-1-ol	19.44	0.73	1377	Leaf-like
69	Benzyl methyl ether	19.62	1.00	1382	-
70	Nonanal	19.95	1.64	1390	Fatty, waxy, pungent
71	(E,Z)-2,4-Hexadienal	20.11	0.89	1394	Green
72	1-Hepten-3-ol	20.12	1.36	1394	-
73	(E,E)-2,4-Hexadienal	20.37	0.89	1400	Green
74	(E)-2-Hexen-1-ol	20.37	0.73	1400	Green grass, leaves
75	(Z)-2-Hexen-1-ol	20.51	0.71	1404	Green grass, leaves
76	α -Thujone (ISTD1)	20.86	1.69	1406	-
77	1-ethenyl-3-ethyl-Benzene	21.57	1.22	1422	-
78	β -Thujone (ISTD2)	21.65	1.61	1424	-
79	1-ethenyl-4-ethyl-Benzene	21.98	1.22	1432	-
80	1-Octen-3-ol	22.27	0.89	1439	Mould, earthy
81	Acetic acid	22.34	0.48	1440	Sour, vinegary
82	1-Heptanol	22.47	0.85	1443	Herb
83	2-Ethyl-1-hexanol	23.87	0.94	1476	-
84	Decanal	24.26	1.73	1485	Penetrating, sweet, waxy
85	5-ethyl-2(3H)-Furanone	24.67	0.87	1495	-
86	Benzaldehyde	24.91	0.79	1505	Almond, burnt sugar
87	1,2-Dimethoxy propanol	25.23	0.85	1513	-
88	2,3-Butanediol	25.69	0.61	1525	-
90	1-Octanol	26.63	0.91	1548	Moss, nut, mushroom
91	Dimethyl Sulfoxide	26.96	0.70	1556	-
92	Formic acid	27.22	0.59	1562	-
93	Propanoic acid	27.55	0.54	1570	Fruity, pungent
94	5-Methyl-4-hexen-3-one	28.01	0.79	1582	-
95	3-Methyl-2-hexen-4-one	28.18	0.76	1586	-
96	Undecanal	28.39	1.83	1591	Sweet, fatty, waxy-floral-citrus
97	Methyl benzoate	28.77	0.91	1604	-
98	2-Ethoxy-ethanol	28.82	0.76	1605	-
99	2-(2-Ethoxyethoxy)ethanol	28.88	0.75	1607	-
100	Butanoic acid	29.38	0.52	1620	Sweaty, rancid
101	Dodecanal	32.52	1.88	1701	Fatty, citrus-like
102	Pentanoic acid	33.58	0.56	1730	Sweaty
103	5-ethyl-2(5H)-Furanone	33.79	0.82	1736	-
104	(E,E)- α -Farnesene	33.87	2.03	1739	-

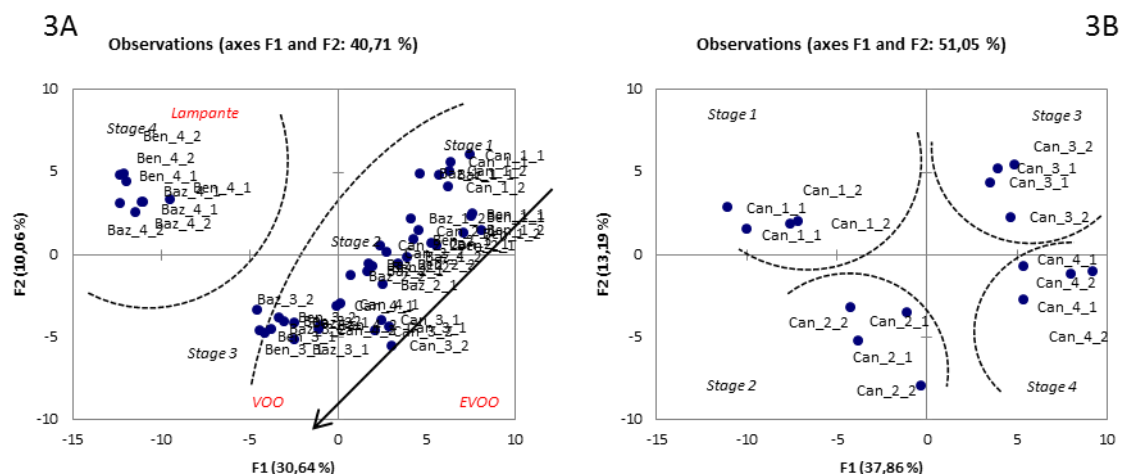
§: Ref.[40]

Table 2 (continued).

ID	Compound Name	1D (min)	2D (sec)	Exp ITS	Odour descriptor
105	Methyl salicylate	34.43	0.92	1754	-
106	2-Butoxy ethanol	35.35	0.87	1779	-
107	Hexanoic acid	37.41	0.59	1837	Goat-like, sweaty
108	(Z)-6,10-dimethyl-5,9-Undecadien-2-one	37.55	1.60	1842	-
109	(E)-Geranylacetone	37.61	1.52	1843	-
110	Isobutyl benzoate	37.64	1.23	1844	-
111	Benzyl alcohol	37.92	0.66	1852	Sweet, fruity
112	2-Phenyl ethanol	39.11	0.72	1885	Honey-like
113	Heptanoic acid	41.00	0.60	1941	-
114	Phenol	42.24	0.58	1977	-
115	Octanoic acid	44.52	0.64	2047	Sweaty
116	2-Phenoxy ethanol	46.53	0.69	2110	-
117	Nonanoic acid	47.86	0.66	2152	Sweaty, waxy
118	Decanoic acid	51.09	0.70	2259	Soap-like, fatty
119	γ -lactone	52.60	0.78	2311	-

§: Ref.[40]

Principal Component Analysis (PCA) maps the natural, unsupervised conformation of samples' groups and sub-groups [40]. Figure 3A shows the scores plot on the first two principal components (F1-F2 plane), based on the 48 × 119 matrix (samples × targets). The variance from the first principal component (F1) was 30.64% while for the second principal component (F2) was 10.06%. Autoscaling and mean centering were applied as pre-processing methods, because baseline correction already was applied for 2D data elaboration by GC Image. The corresponding loading plots are available as Supplementary information (Supplementary Figure SF2A).



Figures 3A-B: PCA results. (3A) Scores plot on the first two principal components (F1-F2 plane), based on targets distribution across all samples (48 × 119 matrix - samples × targets). (3B) Scores plot the first two principal components (F1-F2 plane) based on targets distribution across *Caniles* EVOO subset.

The PCA shows a clear discrimination between EVOO and VOO (clustered together in the right side) and *LOO* samples. Additionally, a further sub-classification according to harvesting stage is evident along F2 and within EVOO samples (see arrow).

The samples' structure/classification over the PCA loading plot (see Supplementary Information SF2A) indicates those analytes that are effectively responsible for the discrimination of *lampante* oils (stage 4 of *Benamaurel* and *Baza*). These analytes include saturated (e.g., heptanal, octanal, and nonanal) and unsaturated (e.g., (*E*)-2-heptenal) aldehydes, well-known from the literature to be correlated with specific sensory defects of olive oils. Moreover, the separation of *lampante* oils is also driven by other compounds, including some alcohols (e.g., propan-1-ol, 1-octen-3-ol, heptan-1-ol, and octan-1-ol), ketones (e.g., heptan-2-one and

octan-2-one) and esters (e.g., ethyl acetate, ethyl-2-methyl butanoate, and ethyl-3-methyl butanoate).

Additional insight on the targets' distribution as a function of harvest time was obtained by independently processing single subsets of samples by geographical location. In this way, all variables related to the pedoclimatic conditions and field treatments (organic or conventional) were excluded, and indications on variables (markers) correlated with ripening stage are more clearly evidenced. Figure 3B shows the score plot for the *Caniles* EVOO subset and the corresponding loadings plot is provided as Supplementary information (Supplementary Figure 2B -SF2B).

Compounds that contribute most to discriminating harvest stage 1 are: (*Z*)-2-hexenal, connoted by a *fruity* note; (*Z*)-3-hexenal, with *green* odor; and (*E,E*)-2,4-hexadienal, contributing a *fresh* note to the overall perception. A group of unsaturated hydrocarbons, tentatively identified from the literature [41], was found to be distinctive in the discrimination of the earlier harvest stages (1-2): 3,4-diethyl-1,5-hexadiene (RS+SR), 3,4-diethyl-1,5-hexadiene (meso), (5*Z*)-3-Ethyl-1,5-octadiene, (5*E*)-3-Ethyl-1,5-octadiene, (*E,Z*)-3,7-decadiene, (*E,E*)-3,7-decadiene, and (*E*)-4,8-Dimethyl-1,3,7-nonatriene. As expected, all these markers decrease in later ripening stages.

Interestingly, but not surprisingly, the evolution of (*Z*)-2-hexenal and (*Z*)-3-hexenal, which provide a *fruity* note (Mf), through time is in accordance with the sensory evaluation of the panel (as reported in Table 1). The relative abundance of these analytes shows a marked decrease from

samples harvested in November (2014) to late January (2015). This observation is confirmed by data from *Baza* oils where (*Z*)-3-hexenal falls below method Limit of Detection (LOD) at stages 3 and 4 consistent with the perception of defects ($M_d > 0.00$) leading to their classification as *lampante* oils, while (*Z*)-2-hexenal in *Benamaurel* samples was not detected even at first harvesting time.

On the other hand, some other target analytes, for example octane (*sweetly, alcane*), nonanal (*fatty, waxy*), and 6-methyl-5-hepten-2-one (*pungent, green*), showed an opposite trend by increasing their relative abundance from stage 1 to the 4. Their presence was not revealed by the panel, possibly because of their relatively high odor thresholds (octane 940 $\mu\text{g}/\text{Kg}$, nonanal 150 $\mu\text{g}/\text{Kg}$, and 6-methyl-5-hepten-2-one 1000 $\mu\text{g}/\text{Kg}$ [22]).

These results are consistent with those reported by Aparicio and Morales [19], Raffo et al. [42] and other researchers [22,43] who hypothesized an increase of autoxidation products (e.g., octane and 6-methyl-5-hepten-2-one) accompanied by a decrease of lipoxygenase pathway products (e.g., (*Z*)-2-hexenal, (*Z*)-3-hexenal, and (*E,E*)-2,4-hexadienal) with later harvest times.

PCA carried out on *Baza* and *Benamaurel* oils (scores and loadings provided as Supplementary information, Supplementary Figures 3A and 3B - SF3A and SF3B) confirms, with some exceptions (e.g., in *Benamaurel* samples), the distribution of the samples and the trend of these specific chemicals over time from stage 1 to 4.

3.2. Untargeted analysis

Untargeted analysis was performed to extend the comparative process to the entire pattern of detected VOCs. The unsupervised fingerprinting was based on the peak-region feature approach and implemented by Image Investigator in the GC Image software package. This data elaboration step was made more informative by considering the 2D peaks included in the targeted template built within the Step 1 of the work-flow (illustrated in Figure 1), thus preserving all information about known analytes within the fingerprinting.

The fully automated procedure of *peak-regions* fingerprinting delineates a small 2D retention-times window (or *region*) per 2D *peak* over the chromatographic space. Regions are shown in Figures 2B and 2D, delineated with light blue graphics. In this context, the process approaches "one-feature-to-one-analyte" selectivity, typical of peak features methods, with all the advantages of regional features matching [5,27,28]. These advantages includes unambiguous cross-detection/matching of trace peaks that may be detected in some samples but not in others and co-eluting analytes that may be resolved in some chromatograms but not in others.

The unsupervised procedure is:

1. Detect and record 2D peaks in individual chromatograms.

2. Locate *registration* peaks, i.e., peaks that reliably match across all chromatograms (connoted by red circles in Figures 2D and 2E). This is verified for a sub-group of targeted peaks.
3. Align and combine all chromatograms to create a composite chromatogram [5].
4. Define a pattern of *region* features around every 2D peak detected in the composite chromatogram.
5. Create a combined targeted and untargeted template from:
 - a. the registration peaks from Step 2,
 - b. the peak-regions from Step 4, and
 - c. the targeted peaks.

The programmed output of the Image Investigator is the template that includes only (a) and (b). An innovation of this work is the addition of (c) targeted peaks.

Once the resulting template, as shown in Figure 2B superimposed on the image of the *Baz_4_1* sample - analytical replicate 1, is matched to a target chromatogram, the analysis includes peak-regions (light blue graphics), targeted peaks (green circles), and registration peaks (red circles). Feature regions are aligned relative to corresponding peaks, and the characteristics of those features including all metadata (retention times in both chromatographic dimensions, detector response, relative/absolute intensity, peaks' EI-MS fragmentation pattern, response factors, etc.) are computed to create a feature vector for the target

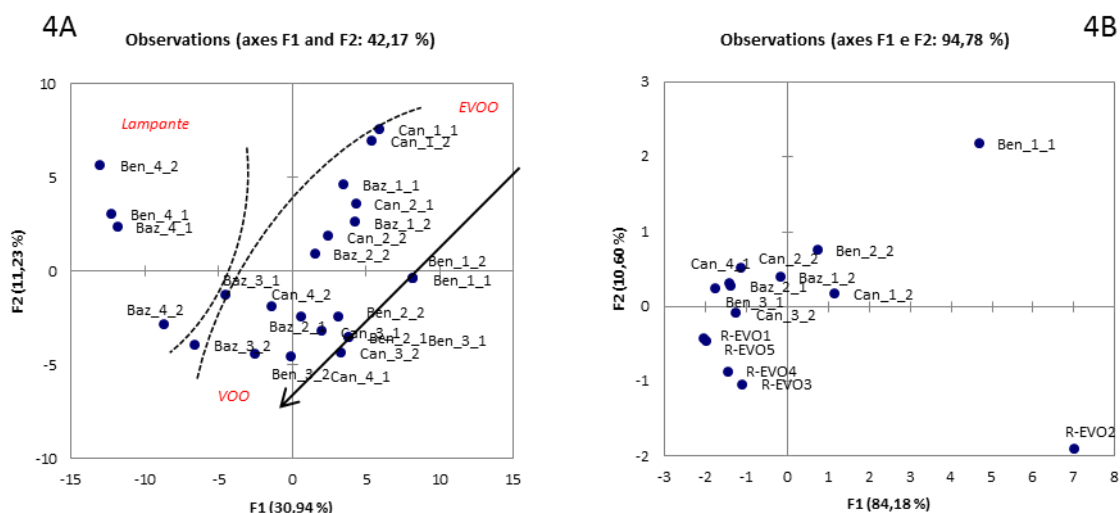
chromatogram to be adopted for cross-sample analysis. The final output is a data matrix where peak-regions and template peaks are cross-aligned within all samples' chromatograms and the response data are available for further chemometrics.

Results based on 180 reliable peak-regions (i.e., those that matched in all-but-one chromatogram of the set) are shown in Figure 2B, and visualized by PCA of Figure 4A. They confirmed what already was evidenced by the known targets distribution: a clear discrimination of *lampante* oils from VOO and EVOO while maintaining the sub-classification based on harvesting period. These results account for a total variability of 42%, in line with previous elaborations.

Targeted peak-regions cross-validate the classification based on PCA: (*Z*)-2-hexenal, (*Z*)-3-hexenal, (*E,E*)-2,4-hexadienal, 1,4-pentadiene, (5*Z*)-3-ethyl-1,5-octadiene, and (*E*)-4,8-dimethyl-1,3,7-nonatriene contribute to the discrimination of stages 1 and 2 against the others, as obtained in the previous elaboration. Untargeted analysis does not discover additional informative roles of un-identified features and confirms the coverage of the targeted peaks.

One interesting and positive aspect of these results is the strong accordance between targeted and untargeted fingerprinting in terms of sample discrimination effectiveness. This result was not observed when, for example, sampling conditions included too many variables known to impact the VOCs fingerprint (e.g., cultivars, geographical origin, harvesting period/year, technological process, bad practices etc.) [8]. In

those less-controlled cases, the sensitivity and effectiveness of untargeted methodologies were lower and targeted analysis gave better results. In such cases with more experimental variables, much larger numbers of samples may be required for effective discrimination.



Figures 4A-B: PCA results. (4A) Scores plot on the first two principal components (F1-F2 plane), based on reliable *peak-regions* distribution across all samples (48 × 180 matrix - samples × reliable *peak-regions*). (3B) Scores plot the first two principal components (F1-F2 plane) based on informative ratios between ripening markers. For details see text.

Another interesting outcome, in line with previous studies on flavor blueprint [8], is the accordance between sensory quality scores and samples sub-classes. Because sensory profiles by descriptive analysis were not available, a direct correlation between odor-active compounds distribution and sensory quality was not possible. However, positive attributes (Mf in Table 1) had high scores for samples harvested at stages 1 and 2 that rapidly decreased at stages 3 to 4. Along the same

Principal Component (e.g., F2) samples discrimination is in accordance with both variables (i.e., quality score and ripening stage).

Cross-validation of fingerprinting results reinforces and confirms the role played by some ripening markers responsible for positive attributes (green, fruity and fresh) [22,34]. These compounds appear at stage 1, last up to the stage 2, and then start to decrease. From these results, and in agreement with quality parameters (Table 1), the optimal harvest period to obtain a product with high sensory quality from *Picual* variety appears to have been November within stages 1 and 2.

Several informative analytes positively and/or negatively correlated with ripening and oil quality, were therefore selected and their ratio profiled as a function of harvest stages. In addition, a retrospective analysis on EVOO samples' pattern acquired during a previous study [8] was performed to verify the reliability and consistency of these indicators.

3.3 Retrospective analysis and definition of reliable chemical indexes of ripening

Relative ratios (based on 2D Peak Volumes) from the informative chemicals highlighted by the *UT fingerprinting* were calculated and trends observed along harvest stages.

These ratios are functions of sampling parameters (phase ratio, β ; sampling temperature; and time), but derive from analyses conducted under highly standardized and head-space linearity conditions. These VOCs fingerprints are therefore informative and replicable, and these

ratios could be transferred to other studies/ batches and considered as chemical indices of ripening.

Analytes chosen to discriminate samples at stages 1 and 2 were: (*Z*)-2-Hexenal, (*Z*)-3-Hexenal, (*E*)-2-Pentenal, and (*E,E*)-2,4-Hexadienal; those chosen that contributed to discriminate the late harvest stage were: octane, 6-methyl-5-hepten-2-one, and nonanal. Their ratios for the *Baza* samples set are illustrated by the box-plot graphics in Figure 5 and correspond to: (*Z*)-3-Hexenal/6-Methyl-5-hepten-2-one, (*Z*)-3-Hexenal/Nonanal, (*Z*)-2-Hexenal/6-Methyl-5-hepten-2-one, (*E,E*)-2,4-Hexadienal/6-Methyl-5-hepten-2-one, (*Z*)-3-Hexenal/Octane, (*E*)-2-Pentenal/Nonanal, (*E*)-2-Pentenal/6-Methyl-5-hepten-2-one, and (*E*)-2-Pentenal/Octane. Trends were estimated by fittings with exponential, polynomial, or linear functions to delineate their evolution along harvest stage, resulting functions are reported in Table 3. The accuracy of fittings is as assessed by the determination coefficient (R^2).

As a general consideration, most of the ratios followed an exponential or second order polynomial trend with the exception of (*E*)-2-pentenal/octane index whose evolution was relatively linear. In addition, non-linear trends are connoted by higher informative potential because of their sudden changes between optimal and non-optimal ripening stages. Notably, their numeric values decreased one order-of-magnitude between harvest stages where oil quality changed from EVO to VO or *lampante*.

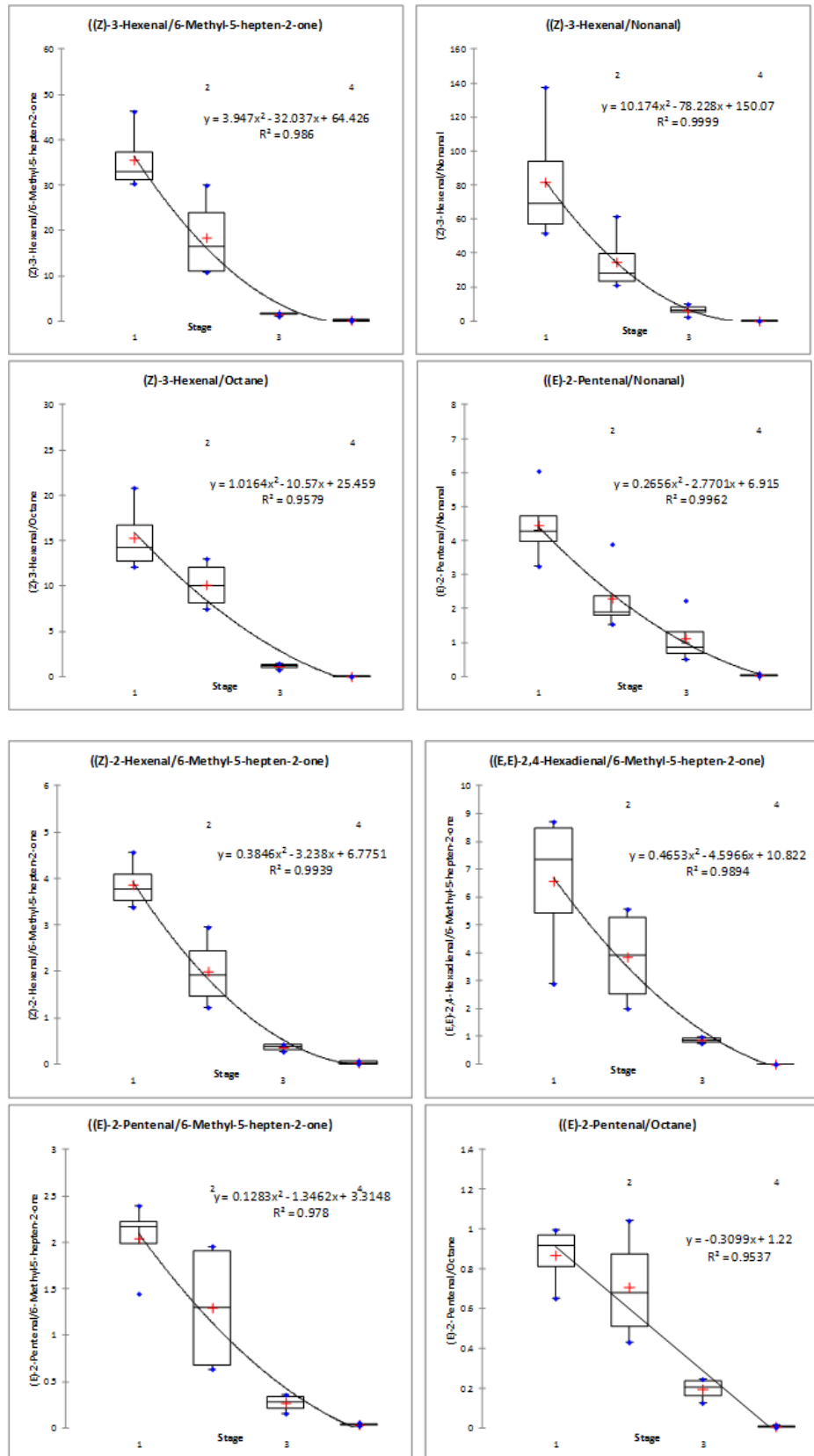


Figure 5: Box-plot graphics showing the evolution of different informative ratios between ripening markers along harvest stages for Baza plot samples.

Table 3: Ripening informative markers evolution trends along harvest stages. The quality of fittings is referred as Coefficient of Determination (R^2).

Informative ratios	Benamaurel	R^2	Baza	R^2	Caniles	R^2
(Z)-3-Hexenal/6-Methyl-5-hepten-2-one	$y = 31.09x^2 - 227.3x + 413.1$	0.999	$y = 3.947x^2 - 32.03x + 64.42$	0.986	$y = 103.8e-0.59x$	0.939
(Z)-3-Hexenal/Nonanal	$y = -4.729x^2 - 21.60x + 158.3$	0.956	$y = 10.17x^2 - 78.22x + 150.0$	0.999	$y = 167.7e-0.53x$	0.956
(Z)-2-Hexenal/6-Methyl-5-hepten-2-one	$y = 2.606x^2 - 18.63x + 32.95$	0.998	$y = 0.384x^2 - 3.238x + 6.775$	0.993	$y = 11.66e-0.59x$	0.959
(E.E)-2.4-Hexadienal/6-Methyl-5-hepten-2-one	$y = 236.1e-1.71x$	0.969	$y = 0.465x^2 - 4.596x + 10.82$	0.989	$y = 2.283x^2 - 15.30x + 26.81$	0.998
(Z)-3-Hexenal/Octane	$y = 5.22x^2 - 46.37x + 100.9$	0.987	$y = 1.016x^2 - 10.57x + 25.45$	0.957	$y = -4.758x + 25.81$	0.862
(E)-2-Pentenal/Nonanal	$y = 27.60e-1.28x$	0.794	$y = 0.265x^2 - 2.770x + 6.915$	0.996	$y = x^2 - 6.839x + 13.15$	0.995
(E)-2-Pentenal/6-Methyl-5-hepten-2-one	$y = 42.31e-1.58x$	0.962	$y = 0.128x^2 - 1.346x + 3.314$	0.978	$y = 5.705e-0.51x$	0.910
(E)-2-Pentenal/Octane	$y = -0.601x + 2.250$	0.951	$y = -0.309x + 1.22$	0.953	$y = -0.303x + 1.630$	0.971

The usefulness of such ratios also might be evaluated from a wider perspective where, for example, VOCs fingerprints are adopted for quality classification of EVO oils. Within selected volatiles produced during the climacteric stage of ripening [22], (Z)-3-hexenal is a product of the lipoxygenase pathway and, in EVO and VO oils, it contributes to the fresh aroma perception thanks to its relatively low odor threshold [20,21,34,37]. This compound is also a cultivar-specific marker for the *Picual* variety [34], as is 6-methyl-5-hepten-2-one that, as a counterpart, is connoted by a negative odor perception and an incremental trend along

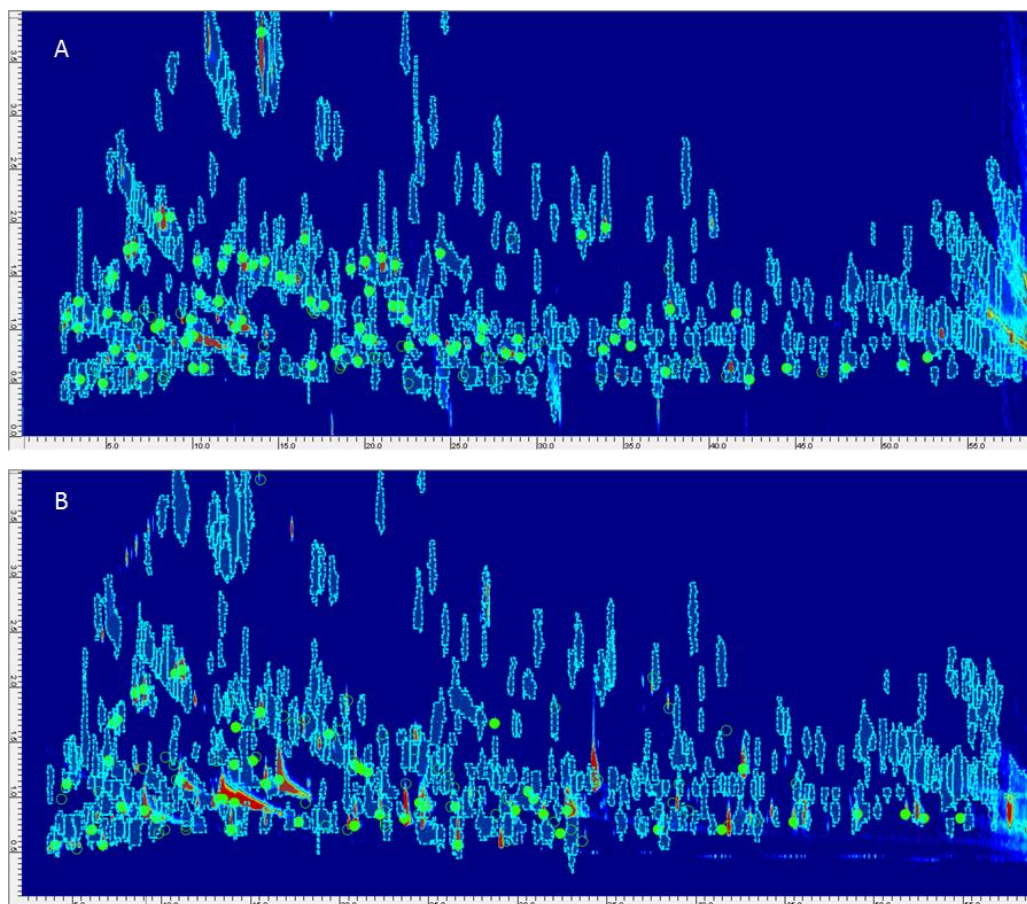
ripening stages. Nonanal provides information about oxidation state as well as octane [36,37,44,45].

To evaluate the consistency and the transferability of this approach for informative chemical indexes, a retrospective analysis was attempted by re-processing chromatograms from a previous study [8]. Samples consisted of EVOO from different botanical/geographical origins and technological processes and from olives harvested in 2013. They were analyzed previously, in the authors' laboratory, with the same nominal HS-SPME sampling protocol and GC×GC-MS conditions. Details are reported in Table 1.

The peak-regions template created in this study (and shown in Figure 6A with the *Can_1_2* sample) was matched to these older, GC×GC chromatograms (as shown in Figure 6B with the EVOO oil from Sicily PDO Monti Iblei) after a supervised transformation of the template to compensate for non-linear retention times differences in both dimensions [46].

These chromatographic inconsistencies are not infrequent because, in a time frame of two years, column sets were replaced and/or columns have altered retention behaviour (in particular the 1D PEG polar phase) producing minimal, but not negligible, pattern alterations. However, thanks to the specificity of the matching methods, 2D peaks that positively match are just those with EI-MS fragmentation pattern similarity above 700 (direct match) or 900 (reverse match). Cross-aligned results are in

consequence reliable and consistent, making possible retrospective investigations.



Figures 6A-B: (6A) 2D chromatogram of *Can_1_2* sample together with the reliable peak-regions template. (6B) 2D chromatogram of *R-EVOO 2* sample (PDO Monti Iblei - Sicily Italy) together with the reliable peak-regions template transformed and adapted with a supervised approach.

Ratios between informative markers for the EVOO samples from *Baza*, *Caniles*, and *Benamaurel*, plus five samples from the previous study (R-EVOO from 1 to 5), were analyzed by PCA and the results are shown in Figure 4B. The discrimination power of the first two PCs reaches 94%, confirming the informative power of the combination of variables. Picual

samples (*Baz*, *Can* and *Ben*) are clustered together, with the exception of the *Benamaurel* EVOO at the earliest harvest stage, whereas along F2, there is evident discrimination for ripening stage. On the other hand, R-EVOO samples clustered together close to stages 2 and 3, with the only exception of R-EVOO 2 PDO (Monti Iblei Sicily, Italy), which showed a very high value for (Z)-3-Hexenal/6-Methyl-5-hepten-2-one because of the high abundance of (Z)-3-Hexenal (which accounted for 12% of Total Volume). The corresponding loading plot for informative ratios is provided as Supplementary information Supplementary Figure 4B - SF4B.

The proposed ratios are consistent within Picual variety, but to be considered as general indices for ripening classification their reliability should be verified and validated by analyzing samples from different harvest years and location, and their transferability to other botanical origins and geographical locations should be investigated *ex-novo* by screening samples after a rigorous sampling design.

3.4 Fingerprinting by image features approach

The last part of this study focuses on a fingerprinting approach based on *visual* features and it is suitable for rapid and effective pair-wise pattern comparisons. The approach is one of the earliest introduced in GC×GC data elaboration [5], and is still adopted when distinctive patterns have to be compared on an untargeted basis to immediately reveal compositional differences.

Previous studies demonstrated the potentials of this simple and intuitive approach by exploring the volatile fraction of roasted coffee and juniper [47], volatiles emitted from *Chrysolina herbacea* bugs fed by *Mentha* spp. leaves [48], and primary metabolites distribution in mice urine after dietary manipulation [26]. The same approach was used iteratively, by cross matching sample pairs, to reveal a chemical blueprint of odor active compounds responsible of sensory defects [8].

In this application, where VOCs variations are mainly related to harvest/ripening stage, the visual approach would be effective to immediately highlight 2D peaks and/or analytes that have significantly different relative distributions between sample pairs. In addition, by comparing samples within the same production plot, the effect of fruit maturation is magnified while keeping constant the effect of local pedoclimatic changes.

This fully automated approach, namely Image Comparison (GC Image v2.5b), if implemented with peak-regions fingerprinting template, provides immediate information about targeted or untargeted peak-regions variations between pair-wise compared samples.

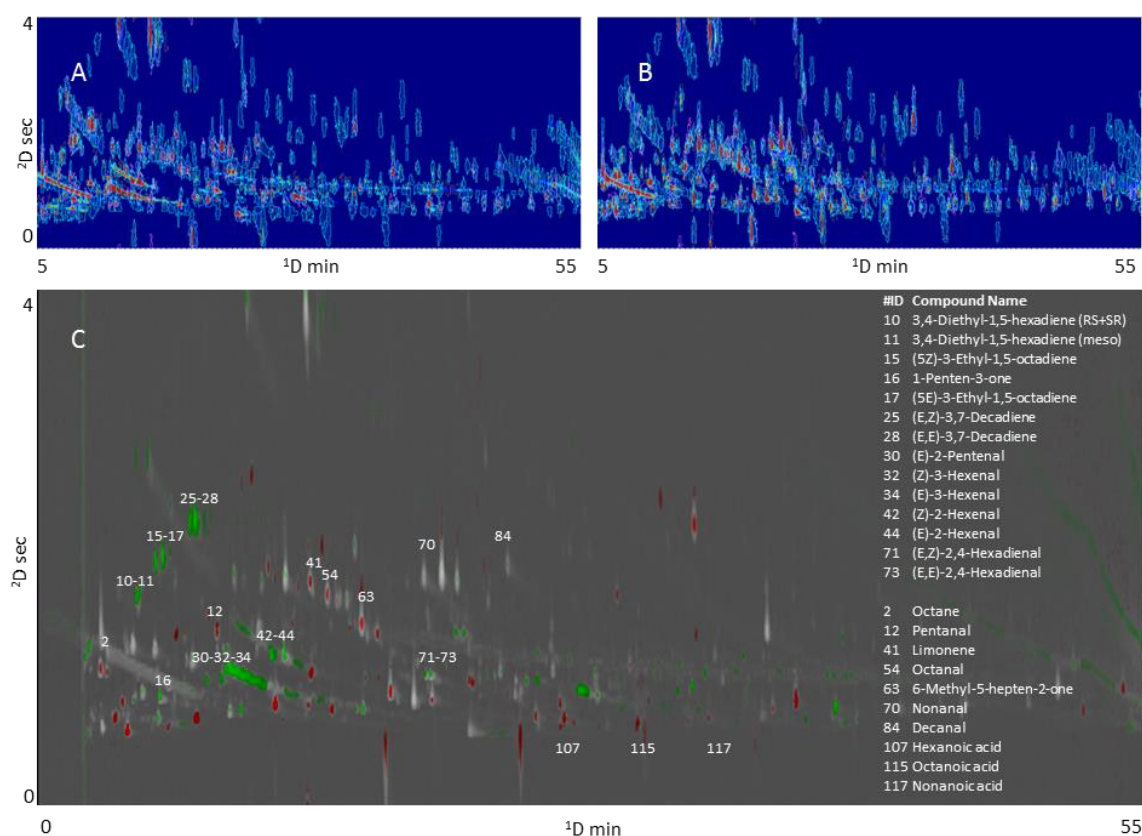
The example here illustrated refers to a *Benamaurel* oil sample obtained at stage 1 (averaged normalized image from *Ben_1_1* and *Ben_1_2*) arbitrarily considered as the *analyzed* image versus the stage 4 samples (averaged normalized image from *Ben_4_1* and *Ben_4_2*) arbitrarily considered as the *reference* image.

Figure 7 shows the image comparison results between the average image of harvest stage 1 (Fig. 7A) and stage 4 (Fig. 7B) obtained by averaging the 2D chromatograms from two replicate locations and two analytical runs. The resulting image (Fig. 7C) is rendered as "colorized fuzzy ratio" that uses the Hue-Intensity-Saturation (HIS) color space to color each pixel in the retention-times plane. The algorithm computes the difference at each data point between aligned pair-wise images. If a pixel is colored green, then the difference is positive, indicating a larger detector response in the *analyzed* image (*Ben_1_1* and *Ben_1_2*). If a pixel is colored red, then the difference is negative, indicating a larger detector response in the *reference* image (*Ben_4_1* and *Ben_4_2*). Brightness depends on the magnitude of the difference, and so white saturation indicates pixels at which peaks have detector responses that are nearly equal in the analyzed and reference images.

Because the 2D chromatograms submitted to the image comparison were already pre-processed by *peak-region* fingerprinting, results are implemented with the information about 2D peaks' identity (if known) or unique identification numbering (#) for unknowns.

Results of *visual* features fingerprinting are intuitive and promptly give information on discriminant peaks. Green colored regions in the upper part of the 2D plot at lower ¹D retention correspond to unsaturated alkanes [41] (#ID 10, 11, 15, 17, 21), unsaturated aldehydes (#30 (E)-2-Pentenal, #32 (Z)-3-Hexenal, #34 (E)-3-Hexenal, #42 (Z)-2-Hexenal, #44 (E)-2-Hexenal, #71 (E,Z)-2,4-Hexadienal, and #73 (E,E)-2,4-Hexadienal), whereas red areas corresponds to limonene (#41), short chain fatty acids

(#107 hexanoic, #115 octanoic and #117 nonanoic acid), linear saturated aldehydes (#12 pentanal, #54 octanal and #70 decanal), and some ketones, such as 6-methyl-5-hepten-2-one (#63).



Figures 7A-C: (7A) Averaged 2D-chromatogram of *Benamaurel* oil samples (field replicates and analytical replicates) obtained at stage 1 and (7B) at stage 4. (7C) Image comparison results between average image of harvest stage 1 (7A) and stage 4 (7B). The resulting image is rendered as "colorized fuzzy ratio". Analytes that varied between stages are listed together with their unique ID numbering (ref. Table 2).

4. Conclusions

This study evidences and emphasizes the potentials of fingerprinting based on *GC*×*GC*-MS separations and highlights the synergism between untargeted and targeted methodologies to investigate complex fractions

of volatiles in depth. Their combination enables to achieve the most inclusive/comprehensive fingerprinting (*UT fingerprinting*) and if compared to previous studies, the degree of automation implemented in the data elaboration work-flow is promising. Experimental results on EVOO volatiles definitely confirm the maturity of available software tools to exploit dense and multi-level data set effectively.

The consistency and reliability of cross-sample analysis results in revealing informative/discriminant features is confirmed by matching results from different approaches, and is of interest in this challenging application field where accurate fingerprinting can be very useful: (a) to support studies aimed at improving product quality; (b) to define a distinctive chemical fingerprint to discriminate samples of a certain botanical/geographical origin; and (c) to re-investigate, on a retrospective projection, samples in light of new informative features.

Acknowledgments

The authors are grateful to "GDR Altiplano de Granada" (Spain) for olives samples.

The research was carried out thanks to the support "International mobility program for young researchers (PhD)" by University of Granada and CEI BioTic Granada.

Note: S. E. Reichenbach has a financial interest in *GC Image, LLC*.

References

- [1] P.Q. Tranchida, P. Donato, F. Cacciola, M. Beccaria, P. Dugo, L. Mondello, Potential of comprehensive chromatography in food analysis. *TrAC Trends Anal Chem* 52 (2013) 186-205
- [2] C. Cordero, J. Kiefl, P. Schieberle, S.E. Reichenbach, C. Bicchi, Comprehensive two-dimensional gas chromatography and food sensory properties: potential and challenges. *Anal. Bioanal. Chem.* 407 (2015) 169-191.
- [3] C. Giddings, Sample dimensionality: a predictor of order-disorder in component peak distribution in multidimensional separation, *J. Chromatogr. A* 703 (1995) 3-15
- [4] Z. Zeng, J. Li, H.M. Hugel, G. Xu, P.J. Marriott, Interpretation of comprehensive two-dimensional gas chromatography data using advanced chemometrics. *TrAC Trends Anal Chem* 53 (2014) 150-66
- [5] S.E. Reichenbach, X. Tian, C. Cordero, Q. Tao, Features for non-targeted cross-sample analysis with comprehensive two-dimensional chromatography, *J. Chromatogr. A* 1226 (2012) 140-148
- [6] S.E. Reichenbach, P.W. Carr, D.R. Stoll, Q. Tao, Smart templates for peak pattern matching with comprehensive two-dimensional liquid chromatography. *J. Chromatogr. A* 1216 (2009). 3458-3466.
- [7] P.Q. Tranchida, G. Purcaro, M. Maimone, L. Mondello, Impact of comprehensive two-dimensional gas chromatography with mass spectrometry on food analysis. *J Sep Sci* 39 (2016) 149-61
- [8] G. Purcaro, C. Cordero, E. Liberto, C. Bicchi, L. S. Conte, Toward a definition of blueprint of virgin olive oil by comprehensive two-dimensional gas chromatography. *J. Chromatogr. A*, 1334 (2015) 101-111.
- [9] L.T. Vaz-Freire, M.D.R.G. da Silva, A.M.C. Freitas, Comprehensive two-dimensional gas chromatography for fingerprint pattern recognition in olive oils produced by

- two different techniques in portuguese olive varieties galega vulgar, cobrançosa e carrasquenha. *Anal Chim Acta* 633 (2009) 263-70.
- [10] T. Cajka, K. Riddellova, E. Klimankova, M. Cerna, F. Pudil, J. Hajslova, Traceability of olive oil based on volatiles pattern and multivariate analysis. *Food Chem* 121 (2010) 282-9
- [11] C. Cordero, E. Liberto, C. Bicchi, P. Rubiolo, S.E. Reichenbach, X. Tian, Q. Tao, Targeted and non-targeted approaches for complex natural sample profiling by GC×GC -qMS. *J Chromatogr Sci* 48 (2010) 251-61.
- [12] C. Cordero, E. Liberto, C. Bicchi, P. Rubiolo, P. Schieberle, S.E. Reichenbach, Q. Tao, Profiling food volatiles by comprehensive two-dimensional gas chromatography coupled with mass spectrometry: Advanced fingerprinting approaches for comparative analysis of the volatile fraction of roasted hazelnuts (*Corylus avellana* L.) from different origins. *J Chromatogr A* 1217 (2010) 5848-58.
- [13] M.N. Franco, J. Sánchez, C. De Miguel, M. Martínez, D. Martín-Vertedor. Influence of the fruit's ripeness on virgin olive oil quality. *J. Oleo Sci.*, 64 (3) (2015), 263-273.
- [14] L.C. Matos, S.C. Cunha, J.S. Amaral, J.A. Pereira, P.B. Andrade, R.M. Seabra, B.P.P. Oliveira, Chemometric characterization of three varietal olive oils (Cvs. Cobrançosa, Madural and Verdeal Transmontana) extracted from olives with different maturation indices. *Food Chem.*, 102 (2007). 406-414.
- [15] L. Martínez Nieto, G. Hodaifa, J.L. Lozano Peña, Changes in phenolic compounds and rancimat stability of olive oils from varieties of olives at stages of ripeness. *J Sci Food Agric.* 90, (2010) 2393-2398.
- [16] T.Gallina-Toschi, L. Cerretani, A. Bendini, M. Bonoli-Carbognin, G. Lercker, Oxidative stability and phenolic content of virgin olive oil: an analytical approach by traditional and high resolution techniques. *J. Sep. Sci.* 28 (2005). 859-870.

- [17] L. Nasini, P. Proietti, Olive harvesting. In: Peri C. (ed). The extra-virgin olive oil handbook. John Wiley Blackwell, Chichester (UK) (2014). pp. 89-105.
- [18] M.N. Franco, J. Sánchez, C. De Miguel, M. Martínez, D. Martín-Vertedor, Influence of the fruit's ripeness on virgin olive oil quality. *J. Oleo Sci.*, 64 (3) (2015). 263-273.
- [19] R. Aparicio, M.T. Morales, Characterization of olive ripeness by green aroma compounds of virgin olive oil. *J Agric Food Chem*, 46 (1998) 1116-1122.
- [20] Morales M.T., Aparicio R., Calvente J.J.. Influence of olive ripeness on the concentration of green aroma compounds in virgin olive oil. *Flav. Frag. J.*, 11 (1996) 171-178.
- [21] R. Aparicio, G. Luna, Characterisation of monovarietal virgin olive oils. *Eur. J. Lipid Sci. Technol.* 104, (2002). 614-627.
- [22] C.M. Kalua, M.S. Allen, Jr. Bedgood, A.G. Bishop, P.D. Prenzler, K. Robards, Olive oil volatile compounds, flavour development and quality: A critical review. *Food. Chem.* 100 (2007) 273-286.
- [23] A.M. Inarejos-García, M. Santacatterina, M.D. Salvador, G. Fregapane, S. Gómez-Alonso, PDO virgin olive oil quality-Minor components and organoleptic evaluation. *Food Res. Int.* 43 (2010). 2138-2146.
- [24] Regulation EU No 1348/2013 on the characteristics of olive oil and olive-residue oil and on the relevant methods of analysis
- [25] S.E. Reichenbach, D.W. Rempe, Q. Tao, D. Bressanello, E. Liberto, C. Bicchi, S. Balducci, C. Cordero, Alignment for comprehensive two-dimensional gas chromatography with dual secondary columns and detectors. *Anal Chem* ;87(19) (2015) 10056-63
- [26] D. Bressanello, E. Liberto, M. Collino, S.E. Reichenbach, E. Benetti, F. Chiazza, C. Bicchi, C. Cordero, Urinary metabolic fingerprinting of mice with diet-induced metabolic derangements by parallel dual secondary column-dual detection two-

- dimensional comprehensive gas chromatography. *J Chromatogr A* ;1361 (2014) 265-76.
- [27] S.E. Reichenbach, X. Tian, A.A. Boateng, C.A. Mullen, C. Cordero, Q. Tao, Reliable peak selection for multisample analysis with comprehensive two-dimensional chromatography. *Anal Chem* 85(10) (2013)4974-81.
- [28] S.E. Reichenbach, X. Tian, Q. Tao, E.B. Ledford, Jr., Z. Wu, O. Fiehn. Informatics for cross-sample analysis with comprehensive two-dimensional gas chromatography and high-resolution mass spectrometry (*GCxGC-HRMS*). *Talanta* 83(4) (2011) 1279-88.
- [29] B. Hollingsworth, S. Reichenbach, Q. Tao, and A. Visvanathan. "Comparative Visualization for Comprehensive Two-Dimensional Gas Chromatography." *Journal of Chromatography A*, 1105(1-2):51-58, 2006
- [30] COMMISSION REGULATION (EEC) No 2568/91 of 11 July 1991 on the characteristics of olive oil and olive-residue oil and on the relevant methods of analysis and successive emendments.
- [31] K. Kotsiou, M. Tasioula-Margari, Changes occurring in the volatile composition of Greek virgin olive oils during storage: oil variety influences stability. *Eur. J. Lipid Sci. Technol.* 117 (2015). 514-522.
- [32] I.Romero, D. L. García-González, R. Aparicio-Ruiz, M. T. Morales, Validation of SPME-GCMS method for the analysis of virgin olive oils volatiles responsible for sensory defects. *Talanta*, 134 (2015) 394-401.
- [33] L. Cerretani, M. Desamparados Salvador, A. Bendini, G. Fregapane, Relationship between sensory evaluation performed by Italian and Spanish official panels and volatile and phenolic profiles of virgin olive oils. *Chem. Percept.* 1 (2008) 258-267.
- [34] G. Luna, M. T.Morales, R. Aparicio, Characterisation of 39 varietal virgin olive oils by their volatile composition. *Food Chem.*, 98 (2006) 243-252.

- [35] R. Iraqi, C. Vermeulen, A. Benzekri, A. Bouseta, S. Collin, Screening for key odorants in Moroccan green olives by gas chromatography-olfatometry/aroma extract dilution analysis. *J. Agric. Food Chem.*, 53 (2005) 1179-1184.
- [36] M. T.Morales, G. Luna, R. Aparicio, Comparative study of virgin olive oil sensory defects. *Food Chem.*, 91 (2005) 293-301.
- [37] Reiners J., Grosch W. Odorants of virgin olive oils with different flavor profiles. *J. Agric. Food Chem.* 46 (1998) 2754-2763.
- [38] M. Rychlik, P. Schieberle, W. Grosch (1998). Compilation of odor thresholds, odor qualities and retention indices of key food odorants. Deutsche Forschungsanstalt für Lebensmittelchemie and Institut für Lebensmittelchemie der Technischen Universität München. Garching, Germany.
- [39] <http://www.flavornet.org/flavornet.html> [Check: 02 May 2016].
- [40] R. Leardi (2003). Chemometrics in data analysis. In: Lees M. (ed). Food authenticity and traceability. Ed. Woodhead, Cambridge.
- [41] F. Angerosa, L. Camera, N. D'Alessandro, G. Mellerio Characterization of seven new hydrocarbon compounds present in the aroma of virgin olive oils, *J. Agric. Food Chem.* 46 (2) (1998) 648-653
- [42] A. Raffo, R. Bucci, A. D'Aloise, G. Pastore Combined effects of reduced malaxation oxygen levels and storage time on extra-virgin olive oil volatiles investigated by a novel chemometric approach. *Food Chem.* 182 (2015) 257-267.
- [43] M.D.R.Gomes da Silva, A.M. Costa Freitas, M.J.B. Cabrita, Garcia Raquel. Olive oil composition: volatile compounds, in: Boskou D. Olive oil-constituents, quality, health properties and bioconversions, Rijeka, Croatia, 2011, pp. 17-46.
- [44] V.Messina, A. Sancho, N. Walsøe de Reca Monitoring odour of heated extra-virgin olive oils from Arbequina and Manzanilla cultivars using an electronic nose. *Eur. J. Lipid Sci. Technol.* 117 (2015) 1-6.

- [45] A. Kanavouras, P. Hernandez-Münoz, F. Coutelieris, S. Selke Oxidation-derived flavor compounds as quality indicators for packaged olive oil. *J. Am. Oil Chem. Soc.*, 81, (2004) 251-257.
- [46] GC Image, LLC. *GC Image GCxGC Edition Users' Guide*, R2.6, 2016
- [47] C. Cordero, E. Liberto, C. Bicchi, P. Rubiolo, S.E. Reichenbach, X. Tian, Q. Tao. Targeted and non-targeted approaches for complex natural sample profiling by GCxGC-qMS. *J. Chromatogr. Sci.* 48 (2010) 251.
- [48] C. Cordero, S.A. Zebelo, G. Gnani, A. Griglione, C. Bicchi, M.E. Maffei, P. Rubiolo, HS-SPME-GCxGC-qMS volatile metabolite profiling of *Chrysolina herbacea* frass and *Mentha* spp. leaves. *Analytical and Bioanalytical Chemistry* 402 (2012) 1941-1952.

4.6. CONCLUSIONES

En esta investigación se han desarrollado dos tipos de metodologías de tratamiento de datos: *targeted* y *untargeted*, a partir de la "huella dactilar" cromatográfica obtenidas por cromatografía bidimensional, de los compuestos orgánicos volátiles presentes en muestras de aceite de oliva, obtenidas en diferentes etapas de maduración de la aceituna.

El estudio da evidencias de la eficacia de trabajar con análisis dirigido y no dirigido de las "huellas dactilares" cromatográficas con el equipo GC×GC. Puesto que es posible extraer un mayor número de información de los picos cromatográficos presentes en las muestras, sean identificados o no.

Ha sido posible identificar un total de 119 compuestos presentes en las muestras y que se han relacionado con diferentes atributos sensoriales, teniendo en cuenta la bibliografía.

Se han seleccionado los volátiles (Z)-2-hexenal, (Z)-3-hexenal, (E)-2-pentenal y (E,E)-2,4-hexadienal como discriminantes de los primeros estados de maduración de la aceituna, correspondientes al mes de noviembre 2014, en esta investigación. Así como los compuestos octano, 6-metil-5-hepten-2-ona y nonanal, responsables de las últimas etapas de maduración de la aceituna (diciembre 2014 y enero 2015). Estos compuestos se han relacionado entre sí para la obtención de un índice de maduración de la aceituna.

Como complemento a esta investigación, sería necesario analizar un mayor número de muestras que permitan cerciorar la utilidad de estos marcadores, así como la inclusión de aceites de oliva de otra variedad, para eliminar la posible relación de los compuestos seleccionados con el tipo de variedad de aceituna o el origen geográfico de las mismas.

4.7. COMUNICACIÓN A CONGRESO

- ❖ L. Valverde-Som, C. Ruiz-Samblás, L. Cuadros-Rodríguez, F. Magagna, C. Bicchi, C. Cordero. **Monitoring the evolution of the profile of volatile organic compounds (VOCs) from olive oil at different ripening stage by comprehensive two-dimensional gas chromatography.** XX Reunión de la Sociedad Española de Química Analítica (SEQA). **Póster.**

CONCLUSIONES FINALES



CONCLUSIONES FINALES

En este capítulo final se exponen las conclusiones derivadas de las aplicaciones que han sido presentadas en esta Tesis Doctoral.

- I. Se ha desarrollado y optimizado un método analítico para la identificación y cuantificación de los principales compuestos (precursores y derivados) asociados a la ruta metabólica del colesterol, mediante GC-(IT)MS, en modo "monitorización con un único ión" (*single ion monitoring*, SIM).
 - I.a. Se ha implementado el uso de una pre-columna para mejorar la relación señal/ruido de los cromatogramas obtenidos. Esta configuración cromatográfica no ha sido aplicada anteriormente en este tipo de muestras.
 - I.b. Se han obtenido cada uno de los espectros de masas, con el analizador de trampa de iones, para un número elevado de patrones correspondientes a los principales precursores y derivados del colesterol.
 - I.c. Se han analizado muestras de plasma de ratón y se ha comprobado que son buenas candidatas para la detección e identificación de los principales compuestos de la ruta metabólica del colesterol.
 - I.d. Se han verificado posibles biomarcadores de la enfermedad del Alzheimer para la detección mediante análisis de sangre.

- II. Se ha elaborado un "tutorial" sobre la aplicabilidad de la metodología de la "huella dactilar" instrumental en la autenticación de alimentos, además se han explicado algunas de las metodologías de tratamiento de datos quimiométricos.
 - II.a. Se han establecido los fundamentos de la metodología de "huellas dactilares".

-
- II.b. Se ha introducido y definido el término "identitación", y se ha diferenciado del término más habitual de "identificación".
- II.c. Se ha efectuado una profunda revisión bibliográfica sobre el estado del arte en lo relativo a la aplicación de la citada metodología.
- III. Se ha desarrollado y optimizado un método analítico para la detección de compuestos orgánicos volátiles presentes en muestras de aceite de oliva mediante GC-FID. Paralelamente, los cromatogramas obtenidos se han convertido en vectores de datos y se han aplicado métodos de análisis de la similitud para llevar el control de la estabilidad organoléptica de los aceites.
- III.a. Se ha llevado a cabo un procedimiento para la selección de marcadores mediante el análisis de componentes principales en diferentes grupos de muestras.
- III.b. Se han aplicado diferentes índices de similitud que permiten comparar cromatogramas.
- III.c. Se ha llevado a cabo la monitorización de la estabilidad organoléptica utilizando como parámetros de control los índices de similitud descritos.
- III.d. Se han propuesto dos nuevos índices de similitud: índice de proximidad (NEAR) e índice de manejabilidad (HAND), descritos por primera vez, en esta Tesis Doctoral, como normalización de los valores de distancia Euclídea y de Mahalanobis, respectivamente.
- III.e. Se ha puesto de manifiesto el uso de la media cuadrática calculada a partir de la T^2 de Hotelling y Q^2 residual para la monitorización de la estabilidad en gráficos de control multivariantes.
- III.f. Se aplica una aproximación del cálculo de la distancia de Mahalanobis para el caso en el que el número de muestras sea menor que el número de variables.
- IV. Se ha desarrollado un método analítico para la detección, identificación y cuantificación de compuestos orgánicos volátiles presentes en muestras de

aceite de oliva mediante GC×GC-(Q)MS y la relación de dichos VOCs con atributos sensoriales.

- IV.a. Se ha caracterizado el perfil de compuestos orgánicos volátiles de aceites de oliva en función de cuatro estadios de maduración de la aceituna, comprendidos entre los meses noviembre-enero.
- IV.b. Se han seleccionado posibles marcadores de la madurez de la aceituna.
- IV.c. Se han estudiado diferentes formas de trabajar con la "huella dactilar" cromatográfica bidimensional (2D) mediante el análisis dirigido y no dirigido de los compuestos presentes en ella, desarrollando así el método a llevar a cabo para el tratamiento de datos.

Principales aportaciones

Las principales contribuciones de esta Tesis Doctoral, atendiendo al interés para la comunidad científica, se pueden resumir en los siguientes puntos:

- ✓ Un método analítico de cromatografía de gases acoplada a espectrometría de masas con analizador de trampa de iones para la detección, identificación y cuantificación de precursores y derivados del colesterol en muestras de plasma de ratón, con el uso de una precolumna y aplicando modo SIM para la identificación de los metabolitos.
- ✓ Un método analítico de cromatografía de gases acoplada a un detector de ionización de llama para la detección de compuestos orgánicos volátiles en muestras de aceite de oliva.
- ✓ Un método de tratamiento de datos que permite comparar cromatogramas a través de índices de similitud.
- ✓ Varios enfoques, en función del índice de similitud seleccionado, que permiten construir gráficos de control multivariados, y que pueden ser

aplicados para monitorizar la estabilidad organoléptica de patrones de aceite de oliva.

- ✓ Un método analítico de cromatografía de gases bidimensional acoplada a un detector de espectrometría de masas con analizador de cuadrupolo simple, para la detección, identificación y cuantificación de compuestos orgánicos volátiles en aceite de oliva. Además de la selección de diferentes indicadores de la madurez de la aceituna en relación a la obtención de aceites de oliva.

Se puede concluir que ha sido posible desarrollar diferentes métodos analíticos en función de la aplicación requerida en el ámbito de ciencias de los alimentos y de la salud. Se han puesto en conocimiento diferentes formas de trabajar con la técnica de cromatografía de gases, así como la espectrometría de masas y el detector de ionización de llama.

Perspectivas futuras

El trabajo presentado hace referencia a estudios en los que se han obtenido buenos resultados cumpliendo con los objetivos propuestos a priori, aun así, se pueden considerar nuevas aportaciones que permitan realizar una serie de investigaciones futuras.

Teniendo en cuenta la investigación desarrollada en el capítulo 1 sobre la determinación de precursores y derivados del colesterol en muestras de sangre, se podría aumentar el número de muestras a analizar en plasma de ratón para confirmar así los marcadores propuestos para la detección de la enfermedad del Alzheimer.

Además, sería necesario poder extrapolar esta experiencia a humanos y contar con muestras de pacientes enfermos y sanos. Así, a partir de los datos obtenidos con el método en modo SIM, cabría la posibilidad de elaborar un modelo quimiométrico con el cromatograma TIC procedente de la unión de cada uno de los segmentos generados en el método analítico y que permita la clasificación de

enfermos de Alzheimer y pacientes sanos, sin la necesidad de la cuantificación de los metabolitos.

Para el caso de la investigación correspondiente al capítulo 3 sobre la monitorización de la estabilidad de patrones de aceite de oliva para análisis sensorial, sería interesante identificar los marcadores seleccionados mediante un detector de espectrometría de masas. Además, se abre un amplio abanico de posibilidades generadas en otra dirección a la que inicialmente se propuso esta investigación científica. Por ejemplo, cabría la posibilidad de someter diferentes aceites de oliva a condiciones de estrés, de luz y temperatura, en botellas transparentes, para establecer una relación entre un aceite de oliva de calidad virgen extra y su deterioro. Estos problemas siempre han preocupado al sector del aceite de oliva y a veces ocurren en determinados puntos de ventas o por el transporte de los mismos. Mediante el uso de la cromatografía y el tratamiento de datos quimiométrico, así como los diferentes índices de similitud propuestos, se podría obtener un indicador de este tipo de alteración en el aceite.

Por último, y teniendo en cuenta el estudio llevado a cabo durante la estancia doctoral (capítulo 4), se plantea aplicar los índices del estado de maduración propuestos en otro tipo de aceites de oliva. Así como llevar a cabo un análisis con mayor número de muestras, que permita obtener la relación de los compuestos orgánicos volátiles con el estado de maduración de la aceituna, de aceitunas procedentes de diferentes variedades, diferentes cosechas, diferentes orígenes geográficos y que permitan confirmar la fiabilidad de los posibles marcadores seleccionados en el estudio.

Valoración personal

Esta tesis pretende ser una contribución de diferentes aplicaciones, en dos ámbitos científicos, que ponen de manifiesto el uso de "huellas dactilares" instrumentales como alternativas a la cromatografía clásica de detección y cuantificación de analitos. Dentro de estas aplicaciones se han desarrollado diferentes métodos analíticos y se ha trabajado en diferentes matrices, que aportan nuevos métodos a la comunidad científica, así como un amplio campo por explorar aún, siguiendo las perspectivas futuras que se podrían desarrollar a partir de los estudios aquí propuestos.

Con lo que respecta a las competencias básicas adquiridas por la doctoranda y que se mencionaron en el capítulo de presentación de esta Tesis Doctoral, regidas por el Real Decreto 99/2011, se puede concluir que se han alcanzado cada una de ellas con éxito. Así:

- ✓ Con el desarrollo de la Tesis Doctoral se ha capacitado a la doctoranda para el trabajo autónomo en el ámbito de la I+D+i proponiéndole problemas y circunstancias que han acontecido a lo largo de su investigación. Dándole la posibilidad de resolver los problemas sola, así como en equipo mediante debates y planteando nuevos puntos de pista.
- ✓ Se ha demostrado la capacidad y competencia para desarrollar diferentes tipos de métodos analíticos en cromatografía de gases, en función del tipo de problema de partida planteado.
- ✓ Ha sido posible poner en práctica y adaptar la información encontrada en otros trabajos y por otros investigadores, permitiendo así, diseñar y crear nuevos métodos de trabajo.
- ✓ Además, con la divulgación científica expuesta a lo largo de la Tesis, mediante diferentes comunicaciones a congresos, tanto en formato póster como oral, y con la elaboración y publicación de diferentes artículos científicos, se ha puesto de manifiesto la capacidad de la doctoranda para exponer e informar a la comunidad científica, así como a cualquier tipo de público, de las aportaciones realizadas.

Así mismo, la realización de los estudios dentro de un grupo de investigación favorece tanto el trabajo en equipo, como de forma autónoma. La estancia realizada ha contribuido al aprendizaje en un contexto internacional y multidisciplinar y que le ha permitido conocer otros sistemas de trabajo.

Por último, la defensa pública ante un tribunal especializado en el campo de estudio, favorecerá a desarrollar la crítica y defensa intelectual de soluciones por parte de la doctoranda.

DESPEDIDA



DESPEDIDA

Ahora sólo queda despedirme y seguir hacia una nueva etapa de mi vida...

Quiero decir que todo en esta vida es mucho más fácil de lo que nos pensamos, mucho más fácil de lo que nosotros lo hacemos...

Hay que VIVIR y olvidarse de los miedos, porque son estos miedos los que nos paralizan, los que nos hacen débiles, los que nos hacen no querernos como somos ni querer a los demás...

Tienes que ser fuerte, pero no desde el miedo de la debilidad, si no desde esa fuerza que crece de dentro, que crece desde el amor y que te hace ser quien eres realmente.

Si sientes algo, déjalo sentir, no te pongas escudos, si tienes miedo PÁRATE y piensa realmente de donde viene ese miedo, qué es lo que está ocurriendo, realmente es tan grave? realmente te vas a morir? realmente no puedes vivir con eso? realmente te duele? realmente es necesario? Todo puede hablarse, puedes hablar con la persona que está en frente y explicarle lo que sientes, explicarle lo que necesitas, mostrarle quien eres realmente.

Esta vida es solo una, y hay que vivirla lo mejor posible, sin miedos, sin vergüenzas, queriendo a las personas, demostrándoles el amor de dentro, el amor real, no por el simple hecho de que te quieran, de que te presten atención, si no, porque de verdad lo sientes!!

Eres una persona muy grande, eres una persona muy especial y puedes con todo, solo tienes que ATREVERTE. Atreverte a sentir, atreverte a disfrutar, ATREVERTE a SER FELIZ. Coge las riendas de tu vida, que nadie la coja por ti y no permitas que te hieran, di lo que te duele, expresa lo que te molesta, lo que quieres, lo que sientes. Las personas que te quieren te quieren tal como eres, bueno o malo, con defectos y con virtudes, ahora es el momento de que te quieras tú y te acerques con ese amor a los demás, un amor sano, un amor limpio, un amor puro.

VIVE,

DISFRUTA y

SE FELIZ.

Lucía Valverde Som.

