Pontificia Universidad JAVERIANA Cali

Facultad de Ingeniería y Ciencias

Ingeniería Biomédica

Informe Final de Trabajo de Grado

Estudio de la Correlación entre los Compuestos Orgánicos Volátiles en el Aliento Exhalado Condensado (AEC) Medidos y el Desarrollo de Diabetes Mellitus Tipo 2.

Daniela Gutiérrez
Jerónimo Delvasto

Director
Dr. Andrés Jaramillo

Asesora
Viviana Ortega Burbano

19 de julio de 2025

Santiago de Cali, 19 de julio de 2025

Señores
Pontificia Universidad Javeriana – Cali
Dr. Hernán Camilo Rocha Niño
Decano
Facultad de Ingeniería y Ciencias
Ciudad

Cordial Saludo.

Por medio de la presente nos permitimos presentarle el Trabajo de Grado titulado "Estudio de la Correlación entre los Compuestos Orgánicos Volátiles en el Aliento Exhalado Condensado (AEC) Medidos y el Desarrollo de Diabetes Mellitus Tipo 2.".

Esperamos que este trabajo reúna todos los requisitos académicos, cumpla el propósito para el cual fue creado y sirva de apoyo para futuros proyectos relacionados con la profesión.

Atentamente,

Daniela Gutiérrez

Jerónimo Delvasto

Santiago de Cali, 19 de julio de 2025

Señores
Pontificia Universidad Javeriana – Cali
Dr. Hernán Camilo Rocha Niño
Decano
Facultad de Ingeniería y Ciencias
Ciudad

Cordial Saludo.

Certifico que el presente Trabajo de Grado titulado "Estudio de la Correlación entre los Compuestos Orgánicos Volátiles en el Aliento Exhalado Condensado (AEC) Medidos y el Desarrollo de Diabetes Mellitus Tipo 2.", realizado por Daniela Gutiérrez y Jerónimo Delvasto, estudiantes de Ingeniería Biomédica, se encuentra terminado y puede ser presentado para su sustentación.

Atentamente,

Dr. Andrés Jaramillo
Director Trabajo de Grado

Agradecimientos

Expresamos nuestro más sincero agradecimiento a todas las personas e instituciones que hicieron posible la realización de este proyecto de investigación.

En primer lugar, agradecemos a nuestro director el dr.Andrés Jaramillo y nuestra asesora Viviana Ortega por su orientación académica, por las observaciones críticas y por el acompañamiento constante durante todas las etapas del proceso investigativo. Su guía fue fundamental para consolidar la calidad metodológica y científica de este trabajo.

Agradecemos profundamente a los habitantes del municipio de Padilla, Cauca, quienes con generosidad y disposición participaron en la toma de muestras y permitieron el avance de esta investigación. Su colaboración fue esencial para el cumplimiento de los objetivos planteados.

Reconocemos el valioso apoyo brindado por el instituto iÓmicas y el equipo del programa Good Food Fellows, así como por las instituciones académicas y técnicas que facilitaron el acceso a recursos materiales, analíticos y humanos indispensables para el desarrollo del estudio.

Extendemos también nuestro agradecimiento a nuestras familias, amistades y personas cercanas, quienes con paciencia y comprensión, nos acompañaron durante este proceso.

Finalmente, dedicamos este trabajo a todas las personas afectadas por la diabetes mellitus tipo 2, con el deseo de que esta investigación contribuya al desarrollo de estrategias preventivas accesibles, no invasivas y eficaces en el diagnóstico temprano de esta enfermedad.

Glosario

Acrónimos y Abreviaturas

DM2	Diabetes Mellitus tipo 2
COVs	Compuestos Orgánicos Volátiles
AEC	Aliento Exhalado Condensado
GC-MS	Cromatografía de Gases acoplada a Espectrometría de Masas
SPME	Microextracción en Fase Sólida
HbA1c	Hemoglobina Glucosilada A1c
PCA	Análisis de Componentes Principales
OPLS-DA	Análisis Discriminante por Proyecciones a Estructuras Latentes Ortogonales
ROC	Curva Característica Operativa del Receptor
ADA	Asociación Americana de Diabetes
EASD	Asociación Europea para el Estudio de la Diabetes
RT	Tiempo de retención

Resumen

El presente trabajo tuvo como objetivo estudiar los compuestos orgánicos volátiles (COVs) presentes en el aliento exhalado condensado (AEC) para establecer su relación con el riesgo de desarrollar diabetes mellitus tipo 2. Se empleó cromatografía de gases acoplada a espectrometría de masas (GC-MS) como técnica principal de análisis, lo que permitió detectar y cuantificar los metabolitos volátiles presentes en las muestras. La investigación se estructuró en etapas que incluyeron una revisión teórica, el diseño de un protocolo para la recolección y análisis de muestras, y el posterior tratamiento estadístico de los datos obtenidos.

Como resultado, se identificaron diez metabolitos significativamente presentes en individuos clasificados como de alto riesgo de desarrollar diabetes tipo 2. De estos, cinco metabolitos presentan rutas metabólicas de origen endógeno bien descritas, entre los cuales octanal y 6-methyl-5-hepten-2-one destacaron por su relevancia estadística. Estos compuestos son propuestos como posibles biomarcadores tempranos de la enfermedad. La evidencia obtenida respalda el uso del análisis de COVs en AEC como herramienta diagnóstica no invasiva para la detección precoz del riesgo metabólico, contribuyendo al desarrollo de estrategias de medicina predictiva.

Palabras Clave: Aliento Exhalado Condensado (AEC), Biomarcadores, Compuestos Orgánicos Volátiles (COVs), Cromatografía de Gases (GC), Diabetes Mellitus Tipo 2 (T2DM), Espectrometría de Masas (MS).

Abstract

The present study aimed to investigate the presence of volatile organic compounds (VOCs) in exhaled breath condensate (EBC) and their potential association with the risk of developing type 2 diabetes mellitus (T2DM). Gas chromatography–mass spectrometry (GC-MS) was employed as the primary analytical technique, enabling the detection and quantification of VOCs in the collected samples. The research was conducted in several stages, including a theoretical review, the development of a sample collection and analysis protocol, and the subsequent statistical analysis of the data obtained.

A total of ten metabolites were identified as significantly elevated in individuals categorized as high-risk for developing T2DM. Among these, five metabolites were found to have well-established endogenous metabolic pathways. Notably, octanal and 6-methyl-5-hepten-2-one emerged as the most statistically significant candidates. These compounds are proposed as potential early biomarkers for the disease. The findings support the use of VOC analysis in EBC as a non-invasive diagnostic approach for the early identification of metabolic risk, contributing to the advancement of predictive and personalized medicine strategies.

Keywords: Exhaled Breath Condensate (EBC), Biomarkers, Volatile Organic Compounds (VOCs), Gas Chromatography (GC), Type 2 Diabetes Mellitus (T2DM), Mass Spectrometry (MS).

Índice general

1. Introducción **1**

2. Planteamiento del Problema **3**

3. Justificación **5**

4. Objetivos **9**
4.1. Objetivo General . . . 9
4.2. Objetivos Específicos . . . 9

5. Marco de Referencia **11**
5.1. Contextualización del Macroproyecto . . . 11
5.2. Áreas Temáticas . . . 11
5.3. Marco Teórico . . . 11
5.3.1. Compuestos Orgánicos Volátiles . . . 11
5.3.2. Diabetes Mellitus Tipo 2 . . . 17
5.3.3. Relación entre COVs y enfermedades metabólicas . . . 19
5.3.4. Cromatografía de gases acoplada a espectrometría de masas . . . 22
5.3.5. Microextracción en fase sólida SPME . . . 26
5.4. Trabajos Relacionados . . . 28

6. Materiales y Métodos **33**
6.1. Desarrollo de protocolos . . . 33
6.2. Evaluación del Riesgo de Diabetes Tipo 2 . . . 34
6.3. Recolección de Muestras de Aliento Exhalado Condensado . . . 34
6.4. Preparación de Muestras . . . 35
6.5. Justificación metodológica . . . 35
6.5.1. GC-MS: metodología seleccionada . . . 35
6.5.2. Comparación con metodologías alternativas . . . 35
6.6. Optimización y Análisis por SPME-GC-MS . . . 37
6.6.1. Optimización del proceso . . . 37
6.6.2. Análisis de Muestras . . . 39
6.6.3. Reactivos y Estándares . . . 39
6.7. Procesamiento y Análisis Estadístico de Datos: . . . 40
6.7.1. Preprocesamiento de datos en MS-DIAL . . . 40
6.7.2. Análisis de muestras blanco en MS-DIAL . . . 41
6.7.3. Preparación y procesamiento de los datos . . . 42
6.8. Análisis estadístico: . . . 43

6.8.1. Correlación y análisis multivariado 44
6.8.2. Representación gráfica de abundancia relativa 44

7. Resultados y Discusión **45**

8. Conclusiones **57**

9. Trabajos futuros **59**

10. Anexos **61**

Anexos **61**

Bibliografía **71**

Índice de figuras

5.1. Compuestos orgánicos volátiles del microbioma humano. (Fuente: [1]) 13
5.2. Proceso para la detección de compuestos orgánicos volátiles en el aliento mediante un sistema de nanotecnología y de procesamiento de datos. (Fuente: [2]) 14
5.3. Protocolo de recolección de muestras con bolsas Tedlar y análisis posterior. (Fuente: [3]) 15
5.4. Principio de recolección de COVs por medio de una nariz electrónica. (Fuente: [4]) . 16
5.5. Proyección del incremento de la T2DM a nivel global entre 2021 y 2045 realizada por la Federación Internacional de Diabetes. (Fuente: [5]) 17
5.6. Relación entre obesidad y diabetes mellitus tipo 2. Interacción entre susceptibilidad genética y factores ambientales. (Fuente: [6]) 18
5.7. Condiciones fisiológicas de las células β pancreáticas; (A) Condiciones normales; (B) Frente a la diabetes mellitus tipo 2. (Fuente: [7]) 19
5.8. Mecanismos de relación entre la producción de COVs y estrés oxidativo. (Fuente: [7]) 20
5.9. Diagrama simplificado de un sistema de cromatografía de gases. (Fuente: [8] 23
5.10. Espectro de masas obtenido por espectrometría de masas. (Fuente: [9]) 24
5.11. Diagrama de un sistema de cromatografía de gases acoplado a espectrometría de masas (GC-MS). (Fuente: [9]) 25
5.12. Cromatogramas obtenidos mediante GC-MS mostrando los picos correspondientes a PFBIHA y acetona oxima en diferentes condiciones experimentales. (Fuente: [10]) . . 26
5.13. Esquema de diferentes técnicas de microextracción en fase sólida (SPME): (a) SPME en solución directa (DI-SPME), (b) SPME en espacio de cabeza (HS-SPME), y (c) SPME protegida por membrana. (Fuente: [11]) 27

6.1. Relación entre el tiempo de recolección y la masa de aliento exhalado condensado . . 34
6.2. Scatter plot sobre la eliminación de compuestos detectados en muestras blanco. . . . 42
6.3. Scatter plot de metabolitos retenidos y eliminados tras el filtrado por ubiquidad. . . 42

7.1. Gráfico de volcán de los compuestos diferenciadores entre RA y RB. 45
7.2. Mapa de calor de metabolitos significativos relevantes en RA. 47
7.3. Diagramas de caja de los compuestos significativos relevantes en RA. 48
7.4. Comparación de concentraciones semicuantitativas promedio ($\mu g/mL$) de los metabolitos significativos entre los grupos RA y RB. 55

Índice de tablas

5.1. Clasificación de los compuestos orgánicos volátiles según su punto de ebullición. (Fuente: [12] . . . 12
5.2. Relación entre compuestos orgánicos volátiles (COVs) y la diabetes tipo 2 (DT2). . . 21

6.1. Factores experimentales y niveles evaluados en el diseño del método SPME-GC-MS. . 37
6.2. Factores y rangos evaluados en la optimización del método SPME. . . . 38
6.3. Parámetros iniciales SPME-GC-MS. . . . 39
6.4. Parámetros utilizados en el preprocesamiento e identificación de compuestos en MS-DIAL. . . . 41

7.1. Compuestos significativos identificados en el grupo RA . . . 46
7.2. Compuestos significativos identificados en el grupo RA con Mass Hunter y la librería NIST . . . 47
7.3. Mapeo del rango de masas de los 4 estándares . . . 52
7.4. Asignación de estándares para los metabolitos según la diferencia de masas . . . 53
7.5. Concentración semicuantitativa ($\mu g/mL$) de cada metabolito en las muestras analizadas (primer serie de muestras) . . . 53
7.6. Concentración semicuantitativa ($\mu g/mL$) de cada metabolito en las muestras analizadas (segunda serie de muestras) . . . 54
7.7. Concentración semicuantitativa ($\mu g/mL$) de cada metabolito en las muestras analizadas (tercera serie de muestras) . . . 54
7.8. Concentración semicuantitativa promedio ($\mu g/mL$) de cada metabolito según tipo de muestra RA y RB . . . 54
7.9. Resumen estadístico de los 10 compuestos orgánicos volátiles (COVs) significativos más relevantes diferenciadores en RA. . . . 55

Capítulo 1

Introducción

La diabetes mellitus tipo 2 es una enfermedad crónica que se manifiesta principalmente a través de niveles elevados de glucosa en la sangre, lo cual ocurre debido a la incapacidad del cuerpo para producir suficiente insulina o a una respuesta inadecuada a esta hormona [13] . Este tipo de diabetes es el más común y, por lo general, está estrechamente asociado con factores de riesgo como el sobrepeso, la inactividad física y una dieta desequilibrada. Estos componentes favorecen la resistencia a la insulina y el surgimiento de la enfermedad, provocando que los síntomas iniciales sean variables y frecuentemente pasen inadvertidos [13][14]. Por lo tanto, numerosas personas pueden no estar al tanto de su estado hasta que la diabetes ha evolucionado hacia una condición más grave.

Es esencial identificar rápidamente la diabetes mellitus tipo 2, pues puede evitar complicaciones serias a largo plazo, como afecciones cardiovasculares, daños en los riñones y un incremento en el riesgo de infecciones recurrentes[15][16]. Estas dificultades de salud son especialmente frecuentes en las fases avanzadas de la diabetes y pueden afectar de manera significativa la calidad de vida del paciente. Estudios han demostrado que la intervención temprana y el manejo adecuado de la diabetes pueden reducir notablemente el riesgo de estas complicaciones, lo que subraya la necesidad de un diagnóstico oportuno y un tratamiento adecuado.

Para diagnosticar y anticipar el riesgo de desarrollar diabetes tipo 2, se utilizan métodos convencionales, como la prueba de glucosa en ayunas y la medición de hemoglobina glucosilada (HbA1c), que permiten un seguimiento adecuado de los niveles de glucosa[14][17]. Sin embargo, estas técnicas presentan ciertas limitaciones, lo que ha impulsado la búsqueda de alternativas avanzadas. En este contexto, la cromatografía de gases acoplada a espectrometría de masas (GC-MS) ha surgido como una herramienta valiosa. Esta técnica permite analizar y cuantificar compuestos en mezclas complejas, lo cual es útil para examinar compuestos orgánicos volátiles (COVs) en el aliento exhalado, los cuales podrían indicar la presencia de ciertas enfermedades, incluida la diabetes mellitus tipo 2 [18][19][20]. La GC-MS se destaca por su alta sensibilidad y especificidad, permitiendo la detección precisa de pequeñas cantidades de compuestos relevantes en el aliento.

El estudio de la relación entre los COVs en el aliento exhalado condensado (AEC) y la diabetes mellitus tipo 2 busca identificar patrones específicos que podrían estar presentes en personas en riesgo de desarrollar la enfermedad. Esta investigación sugiere que el análisis del aliento podría convertirse en una técnica de diagnóstico eficaz para la detección de la diabetes[18][19]. El reconocimiento de COVs propios de la diabetes tipo 2 podría ayudar a desarrollar dispositivos para el diagnóstico rápido y el monitoreo del avance de la enfermedad. Esta perspectiva no sólo proporcionaría una opción accesible y cómoda para los pacientes, sino que también posibilitaría una identificación constante y no invasiva en grupos de riesgo. La habilidad de supervisar la enfermedad de forma constante permitiría realizar modificaciones rápidas en los tratamientos, proporcionando a los doctores datos

exactos acerca del estado metabólico de cada paciente y, finalmente, mejorando la calidad de vida de aquellos que padecen esta condición[21][22][18].

En el presente estudio, mediante el análisis de muestras de AEC por GC-MS, se identificaron diez metabolitos significativamente presentes en individuos clasificados como de alto riesgo de desarrollar diabetes tipo 2: 2-pentanol, 3(s)-citronelol, ethylene glycol monophenyl ether, 2-octanone, 2-hexanone, butanedioic acid dimethyl ester, 3,5-dimethylpyridine, 2-phenylpropionaldehyde, octanal y 6-methyl-5-hepten-2-one. De estos, cinco compuestos cuentan con rutas metabólicas endógenas bien descritas en la literatura científica (2-pentanol, 2-octanone, 2-hexanone, octanal y 6-methyl-5-hepten-2-one), lo que permite proponerlos como los biomarcadores más prometedores dentro del conjunto evaluado. En particular, octanal y 6-methyl-5-hepten-2-one mostraron una mayor significancia estadística, sugiriendo su fuerte potencial como indicadores de riesgo en etapas tempranas de la enfermedad. Estos hallazgos fortalecen la hipótesis de que el análisis de compuestos volátiles en el aliento podría constituir una estrategia diagnóstica no invasiva y eficaz en el contexto del tamizaje metabólico.

Capítulo 2

Planteamiento del Problema

La diabetes mellitus tipo 2 (DM2) es una enfermedad crónica que altera la regulación de la glucosa en sangre, lo cual genera consecuencias graves para la salud metabólica general. Esta condición está estrechamente relacionada con factores de riesgo modificables, como el sobrepeso, el sedentarismo y dietas hipercalóricas [[13]]. En la actualidad, la DM2 constituye una de las principales causas de morbilidad a nivel mundial, especialmente en adultos de mediana edad y personas mayores [23].

Un aspecto crítico en la evolución de esta enfermedad es la prediabetes, un estado intermedio en el que los niveles de glucosa son más altos que lo normal, pero no lo suficiente como para cumplir con los criterios diagnósticos de diabetes [14]. Esta condición representa una ventana de oportunidad para intervenir antes de que la enfermedad se manifieste clínicamente. Sin embargo, el diagnóstico temprano sigue siendo un reto. Las pruebas tradicionales, como la glucosa plasmática en ayunas, si bien son herramientas estándar, tienen limitaciones: se trata de procedimientos invasivos, dependen de condiciones previas como el ayuno, y no siempre reflejan adecuadamente el riesgo metabólico a largo plazo. Además, la variabilidad interindividual puede afectar la sensibilidad diagnóstica de estas pruebas, lo cual puede retrasar el inicio de intervenciones oportunas [14][17].

La DM2 presenta una fisiopatología compleja, caracterizada por una combinación de resistencia a la insulina y una producción insuficiente de esta hormona. Esta disfunción metabólica favorece la acumulación de glucosa en sangre, desencadenando hiperglucemia sostenida. La progresión silenciosa de la enfermedad, con síntomas que pueden pasar inadvertidos durante años, hace que el diagnóstico a menudo ocurra en fases avanzadas. Esto incrementa significativamente el riesgo de complicaciones crónicas, como daño vascular, neuropatías, nefropatías y enfermedades cardiovasculares. Entre las más graves se encuentran las lesiones en extremidades inferiores, que pueden llevar a infecciones severas, gangrena e incluso amputaciones [24][25][25].

Frente a este panorama, surge la necesidad urgente de identificar métodos diagnósticos más sensibles, accesibles y no invasivos. Una alternativa emergente es el análisis de compuestos orgánicos volátiles (COVs) presentes en el aliento exhalado condensado (AEC), los cuales pueden actuar como biomarcadores metabólicos. Estos compuestos reflejan procesos bioquímicos internos y pueden variar en función del estado de salud del individuo. Diversas investigaciones han sugerido que ciertos perfiles de COVs están asociados con enfermedades metabólicas, incluyendo la diabetes [22][18].

No obstante, la aplicación clínica de este enfoque aún requiere ser validada. En particular, es necesario establecer qué COVs específicos se correlacionan de manera significativa con el riesgo de desarrollar DM2 y qué tan reproducibles y confiables son estos hallazgos mediante técnicas como la cromatografía de gases acoplada a espectrometría de masas (GC-MS). Por ello, el presente estudio busca aportar evidencia sobre la viabilidad del uso de COVs en AEC como herramienta diagnóstica

temprana, contribuyendo así al desarrollo de estrategias en medicina predictiva y prevención de enfermedades metabólicas.

Capítulo 3

Justificación

Caracterizada por un desorden en la regulación de glucosa en sangre, la diabetes mellitus tipo 2 (T2DM) es una de las enfermedades con mayor impacto en la salud pública por sus impactos significativos en la morbilidad y mortalidad a nivel global. En los últimos años ha tenido un crecimiento alarmante en donde diversas poblaciones sin distinción de edad, género o ubicación geográfica se han visto afectadas. Según estimaciones realizadas por la Federación Internacional de Diabetes, en el año 2021 eran aproximadamente 537 millones de personas afectadas por diabetes a nivel mundial, lo que representaba un 10.5 % de la población adulta, lo cual significó también un gasto sanitario de hasta 966 mil millones de dólares. De la misma forma, se realizó una proyección para el año 2045, en donde se estimó un incremento en el número de casos a 783 millones, lo que deja clara la necesidad inminente de nuevas y mejores estrategias de detección temprana y manejo preventivo [26].

En el caso de América Latina, los reportes de diabetes mellitus tipo 2 han aumentado preocupantemente, siendo una de las principales causas de morbilidad y mortalidad. Un artículo reciente de la revista Med, de la Editorial Neogranadina, señala que tanto la falta de estrategias eficaces como la escasez de recursos y la desigualdad socioeconómica han contribuido significativamente al aumento de casos en los últimos años, siendo la pandemia del COVID-19 uno de sus principales factores, pues buena parte de la población adoptó malos hábitos alimentarios y de sedentarismo. En el caso de Colombia, la Asociación Colombiana de Diabetes y diversas instituciones reportaron una prevalencia del 8.1 % que afecta a aproximadamente 2.671.400 personas entre los 20 y 79 años, de los cuales, un preocupante 40 % de los casos no están diagnosticados. De este modo, resulta relevante destacar la necesidad de seguir estrategias más eficaces para el diagnóstico temprano, sobre todo si consideramos que la diabetes tipo 2 es una enfermedad que se presenta en silencio y que, sin diagnóstico oportuno, puede desembocar en complicaciones severas que van desde retinopatías, neuropatías y hasta enfermedades cardiovasculares [27].

El diagnóstico de T2DM tradicionalmente se realiza mediante la prueba de hemoglobina glucosilada A1c, un análisis de sangre que se encarga de estimar el promedio de azúcar en sangre en los últimos 2 o 3 meses por medio de la adherencia de glucosa a la hemoglobina [28]. Existen otros métodos para el diagnóstico de diabetes mellitus tipo 2 como el examen aleatorio de glucosa en sangre o de glucemia en ayunas, la prueba de tolerancia a la glucosa oral y exámenes de detección, sin embargo, todos estos son exámenes invasivos y ofrecen un diagnóstico binario: presencia o ausencia de la enfermedad, dejando de lado la posibilidad de evaluar el nivel de riesgo de la persona [23].

Debido a esto, un desafío importante en el contexto de la T2DM es la gran proporción de personas que desconocen su condición, lo que retrasa el inicio del tratamiento y aumenta el riesgo de desarrollar complicaciones severas. Según estimaciones recientes de la Federación Internacional de Diabetes, casi la mitad de los adultos entre 20 y 79 años que padecen diabetes no están diag-

nosticados. Donde la mayor proporción se encuentra en regiones como África (53.6 %), el Pacífico Occidental (52.8 %) y el Sudeste Asiático (51.3 %), mientras que las tasas más bajas de diabetes no diagnosticada se observan en América del Norte y el Caribe, con un 24.2 %. En los países de ingresos bajos y medios, los más afectados, la enfermedad frecuentemente se detecta sólo tras la aparición de complicaciones microvasculares y macrovasculares, tales como retinopatía, neuropatía, nefropatía y enfermedades cardiovasculares [29].

De esta forma, resalta la necesidad de mejorar los métodos de diagnóstico de la diabetes, así como de implementar herramientas de detección más efectivas y menos invasivas, lo que hace que la investigación de biomarcadores y compuestos orgánicos volátiles (COVs) en el aliento exhalado sea un enfoque innovador con un gran potencial. La identificación temprana de riesgo alto en individuos podría facilitar los mecanismos de prevención y terapia oportuna, reduciendo así la morbilidad y mortalidad asociada con la diabetes.

Este proyecto se centra en la identificación de patrones de COVs específicos asociados con un alto riesgo de desarrollar diabetes mellitus tipo 2, abordando una necesidad importante en el ámbito de la medicina preventiva. Algunos estudios actuales se centran en la investigación de COVs para el diagnóstico de T2DM, sin embargo este proyecto adopta un enfoque innovador al buscar identificar moléculas en el aliento exhalado que correlacionen con un riesgo elevado de desarrollar la enfermedad [30]. De esta manera, este enfoque diferencial no solo contribuye a la investigación teórica sobre biomarcadores en enfermedades metabólicas, sino que también propone una herramienta valiosa para la intervención clínica temprana, adaptada a las características metabólicas individuales de los pacientes.

El proyecto emplea la cromatografía de gases acoplada a espectrometría de masas (GC-MS), la cual sirve para el reconocimiento de los compuestos en concentraciones muy bajas, esta técnica es clave en el ámbito del proyecto, en el sentido de que permitirá identificar tanto la presencia como la variación de los compuestos en el aliento condensado que actúan como marcadores de la aparición de la diabetes mellitus tipo 2 y establecen patrones específicos en el aliento exhalado.

El estudio de los COVs y cómo se relacionan con el riesgo de contraer diabetes tipo 2 no sólo es importante para el avance en el conocimiento de biomarcadores metabólicos, sino que también ofrece oportunidades para crear formas más sencillas de realizar pruebas de diagnóstico menos invasivas, incrementando así la posibilidad de poder predecir enfermedades antes de que éstas se hayan desarrollado por completo y que, por lo tanto, tienen repercusión para la salud pública y la calidad de vida.

Por último, cabe mencionar que este proyecto busca un enfoque interdisciplinario que articula conjuntamente la química, la ingeniería biomédica y la nutrición en la búsqueda integrada a los problemas que se asocian con la diabetes mellitus tipo 2; la química da las herramientas para la identificación y el análisis de los compuestos orgánicos volátiles como biomarcadores de riesgo, la ingeniería biomédica lleva estos hechos a traducir los hallazgos para futuras tecnologías diagnósticas avanzadas y accesibles, y la nutrición permite la contextualización de los patrones metabólicos identificados y asumirlos con los patrones de factores dietéticos y hábitos de vida. Esta dependencia entre las disciplinas no sólo da consistencia y relevancia al proyecto, sino que, además, asegura la generación de soluciones innovadoras y aplicables que suponen, por tanto, un importante avance en

la medicina preventiva y la salud pública.

Capítulo 4

Objetivos

4.1. Objetivo General

Estudiar la correlación entre los compuestos orgánicos volátiles (COVs) presentes en aliento exhalado condensado (AEC) y el desarrollo de la diabetes mellitus tipo 2.

4.2. Objetivos Específicos

1. Diseñar y validar un protocolo estandarizado para la recolección de AEC y para la medición de COVs usando cromatografía de gases acoplada a espectrometría de masas.
2. Establecer un conjunto de COVs en el AEC que se correlacionen significativamente con un mayor riesgo de desarrollar T2DM cuando la enfermedad todavía no es clínicamente detectable.
3. Posibilitar una base conceptual y experimental para considerar a los COVs como biomarcadores potenciales para la evaluación de diagnósticos iniciales y como prevención de T2DM en etapas preclínicas.

Capítulo 5

Marco de Referencia

5.1. Contextualización del Macroproyecto

El presente proyecto se ejecuta dentro del contexto del macropoyecto denominado "Gestión del riesgo de diabetes en la comunidad", el cual es financiado por Good Food Fellows y está liderado por un grupo interdisciplinario de investigadores de la Pontificia Universidad Javeriana. Este macropoyecto busca la promoción de la prevención de la diabetes mellitus tipo 2 (T2DM) a través de la aplicación de estrategias de participación comunitaria e intervenciones en relación con la alimentación.

En este contexto, se realiza la recolección de datos y de muestras en el municipio de Padilla, ubicado en el departamento del Cauca, una zona con una población mayoritariamente afrocolombiana, donde la diabetes mellitus tipo 2 representa un problema de salud pública de especial relevancia. Esto con el fin de evaluar los factores de riesgo en relación a la diabetes y realizar propuestas de aportes que combinen tecnología y prácticas saludables.

5.2. Áreas Temáticas

- Engineering in medicine and biology - Medical services - Medical conditions - Diabetes.
- Instrumentation and measurement - Instrumentation and measurement - Gas chromatography.
- Engineering in medicine and biology - Biology - Biostatistics.
- Mathematics - Statistics - Correlation.
- Materials, elements, and compounds - Compounds - Organic compounds - Volatile organic compounds.

5.3. Marco Teórico

5.3.1. Compuestos Orgánicos Volátiles

Los compuestos orgánicos volátiles son compuestos químicos orgánicos, es decir, cuya principal estructura atómica es el carbono, y que presentan una volatilidad considerable a temperatura ambiental (20°C) y presión atmosférica estándar (101,3 kPa). Los COVs son emitidos no solo desde fuentes antropogénicas tales como materiales de construcción, productos de limpieza y combustibles,

sino que provienen también de fuentes biológicas; incluso se producen emisiones endógenas por parte del cuerpo humano como resultado de los procesos metabólicos. La volatilización de los COVs, así como la interacción de estos con las membranas biológicas hacen que los COVs sean biomarcadores adecuados para el diagnóstico y el seguimiento de la enfermedad, así como también factores de interés medioambiental por sus efectos potencialmente perjudiciales a la salud y el medio ambiente [31].

Los COVs también se caracterizan por tener una gran diversidad estructural, es decir, además de estar compuestos principalmente por átomos de carbono e hidrógeno, también pueden tener otro tipo de átomos como oxígeno, nitrógeno, azufre y halógenos, generando una gran gama de grupos químicos, incluyendo alcoholes, alcanos, éteres, ésteres, cetonas, entre otros [32]. En la atmósfera, los COVs pueden provenir de formas biogénicas, por medio de emisiones de plantas, microorganismos, insectos y animales, o de formas antropogénicas, abarcando actividades industriales, combustión de fósiles y procesos de manufactura. Desde una perspectiva ambiental, los COVs son precursores en la formación de ozono troposférico, smog fotoquímico y aerosoles orgánicos secundarios, afectando la calidad del aire y la salud pública. Además, ciertos COVs poseen propiedades tóxicas y carcinogénicas, lo que resalta la importancia de su monitoreo y control en entornos industriales y urbanos [33][34][35].

Los COVs se clasifican comúnmente en 3 categorías según su punto de ebullición, aquellos con puntos de ebullición más bajos son conocidos como Compuestos Orgánicos Muy Volátiles, mientras que los más altos, a partir de 240-260 °C, son llamados Compuestos Orgánicos Semivolátiles. El cuadro 1 amplía la información acerca de la clasificación de los COVs según el punto de ebullición.

Cuadro 5.1: Clasificación de los compuestos orgánicos volátiles según su punto de ebullición. (Fuente: [12]

Descripción	**Abreviatura**	**Rango del punto de ebullición (°C)**	**Ejemplo de compuestos**
Compuestos orgánicos muy volátiles (gaseosos)	VVOC	<0 a 50-100	Propano, butano, cloruro de metilo
Compuestos orgánicos volátiles	VOC	50-100 a 240-260	Formaldehído, d-limoneno, tolueno, acetona, etanol, 2-propanol, hexanal
Compuestos orgánicos semivolátiles	SVOC	240-260 a 380-400	Pesticidas (DDT, clordano, plastificantes, retardantes de fuego)

El Volatiloma Humano Al conjunto de todos los COVs que produce un organismo, se le conoce como volatiloma. En el cuerpo humano, los COVs pueden provenir desde diversas matrices biológicas, ya sea por procesos endógenos o exógenos, dentro de estas se incluyen la sangre, el aliento, la orina,

la saliva, secreciones cutáneas, heces, entre otras; y logran reflejar la actividad fisiológica interna y la interacción con factores externos, lo que los hace indicadores potenciales de estados de salud o enfermedad [36]. Históricamente, la investigación sobre COVs tiene raíces que se remontan al siglo XVIII, cuando Antoine Lavoisier identificó dióxido de carbono y oxígeno en el aliento exhalado, sentando las bases para técnicas modernas como la capnografía. A lo largo del tiempo, avances clave han incluido la identificación de acetona en pacientes diabéticos (Wilhelm Petters, 1857) y el análisis cuantitativo de hidrocarburos volátiles como biomarcadores de estrés oxidativo [37].

Suele asociarse el origen biológico de los COVs con determinados procesos naturales de carácter metabólico como la oxidación de ácidos grasos, la fermentación bacteriana del intestino y la degradación de aminoácidos y lípidos. Por otro lado, algunos de estos compuestos como los aldehídos o las cetonas son el resultado de la peroxidación lipídica y la descarboxilación de oxoácidos. Las transformaciones pueden tener lugar en distintos tejidos como el hígado o el pulmón y apuntan al rol de los órganos en la modificación de los COVs y su traslado hacia fluidos corporales específicos.

Por otra parte, no se producen COVs únicamente por procesos metabólicos endógenos, sino que los microorganismos que forman parte del cuerpo humano, también liberan COVs que son expresiones únicas de su esencia funcional y que dejan huellas en distintas matrices biológicas. Los COVs microbianos son relevantes en la interacción entre el microbioma y el huésped, y la huella química generada puede dar cuenta de infecciones específicas o de alteraciones en la microbiota [37].

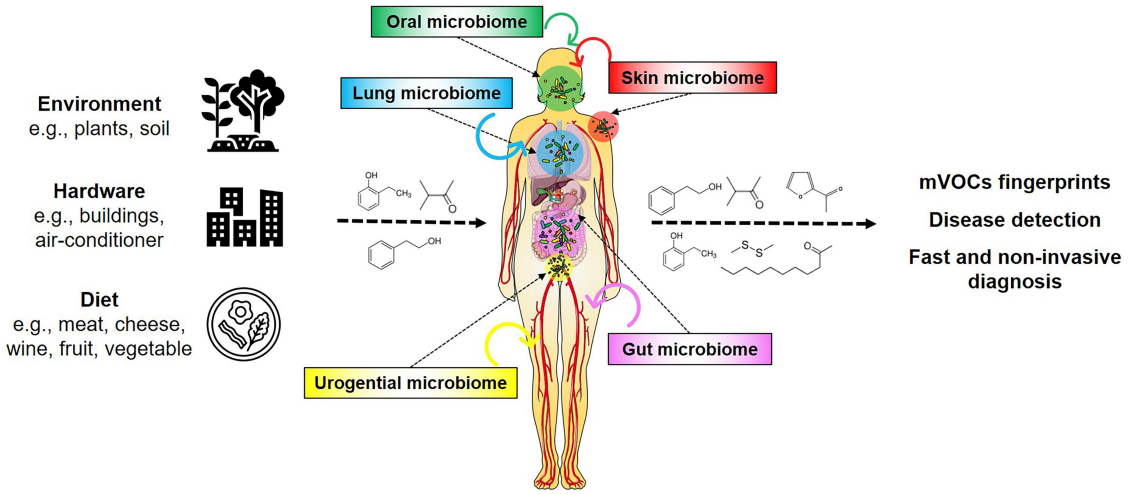

Figura 5.1: Compuestos orgánicos volátiles del microbioma humano. (Fuente: [1])

Los avances tecnológicos han hecho posible la caracterización del volatiloma mediante la cromatografía de gases acoplada a espectrometría de masas (GC-MS) o bien las tecnologías de tiempo real como la PTR-MS o la SIFT-MS. A su vez, se han identificado biomarcadores representados por los COVs en saliva y orina para la detección de enfermedades como el cáncer de próstata, vejiga o mama o incluso infecciones del tracto urinario [38][39].

Importancia de los COVs en la medicina En la medicina, los COVs han surgido como un recurso con gran potencial a tener en consideración para la identificación de enfermedades de forma no invasiva. Su análisis puede ser usado para establecer perfiles químicos del estado fisiológico o patológico de los pacientes, facilitando así el diagnóstico y la aplicación de tratamientos. Estudios recientes han evidenciado la aplicación de técnicas como GC-MS para la identificación con una alta sensibilidad de COVs asociados a enfermedades infecciosas y enfermedades metabólicas, realizando una diferenciación entre una persona sana y una enferma [40].

De las fuentes de COVs más estudiadas en el ámbito médico son la saliva, el sudor y el aliento exhalado, esto debido a la facilidad de su recolección y análisis en tiempo real, lo que ha hecho posible la identificación y asociación de COVs con enfermedades e infecciones por medio de tecnologías emergentes como narices electrónicas y sensores específicos. Biomarcadores como el metil nicotinato han sido descritos en pacientes con tuberculosis activa ofreciendo una alternativa no invasiva y rápida respecto a los métodos tradicionales [40]. Este hecho puede ser de particular interés en escenarios como las unidades de cuidados intensivos o en poblaciones pediátricas donde los métodos invasivos están asociados a riesgos o a prácticas poco factibles.

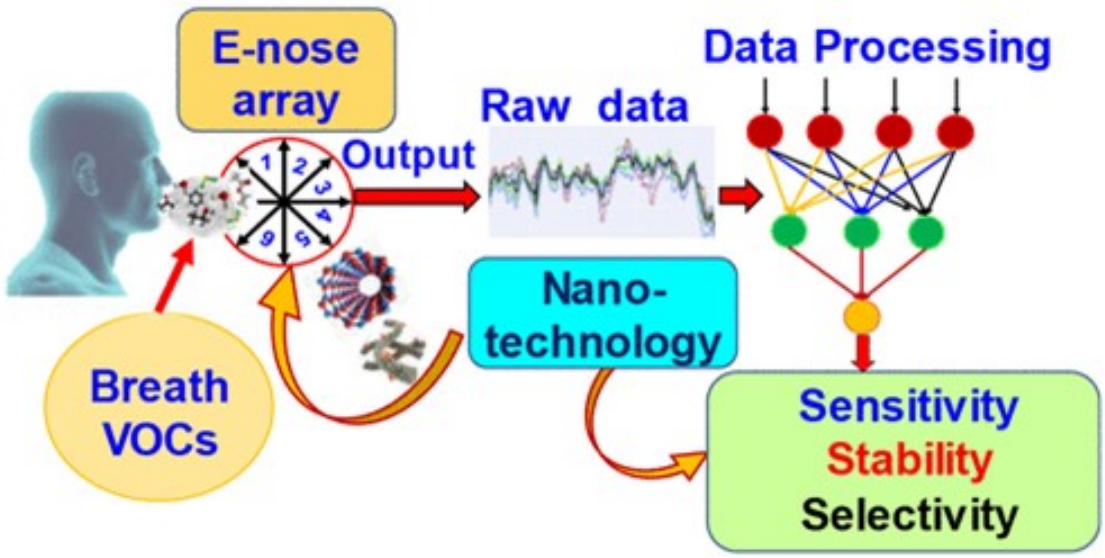

Figura 5.2: Proceso para la detección de compuestos orgánicos volátiles en el aliento mediante un sistema de nanotecnología y de procesamiento de datos. (Fuente: [2])

A pesar de estos avances, todavía existen desafíos significativos. Por ejemplo, la estandarización de los métodos de recolección y análisis, junto con la validación de los biomarcadores en poblaciones amplias y diversas, son pasos críticos para garantizar la implementación clínica de estas tecnologías. Sin embargo, los progresos en instrumentación analítica, como los dispositivos portátiles y espectró-

metros de movilidad iónica, prometen superar estas barreras, haciendo de los COVs una herramienta cada vez más accesible y útil en la práctica clínica [40][41].

Métodos más comunes de recolección de COVs

- Bolsas de polímero: son un medio muy utilizado en la recolección de COVs en los estudios de aliento debido a su portabilidad y fácil uso y almacenamiento en campo. Estas bolsas desechables permiten capturar el aliento exhalado, sin embargo, tienen limitaciones a largo plazoz. Diversos estudios describen un proceso en el que los participantes deben soplar de manera lenta a través de un tubo conectado a las bolsas, gracias a la duración del tiempo de mantenimiento de la respiración antes de la exhalación, se logra tener una mayor proporción de aliento alveolar en las muestras captadas [3][42].

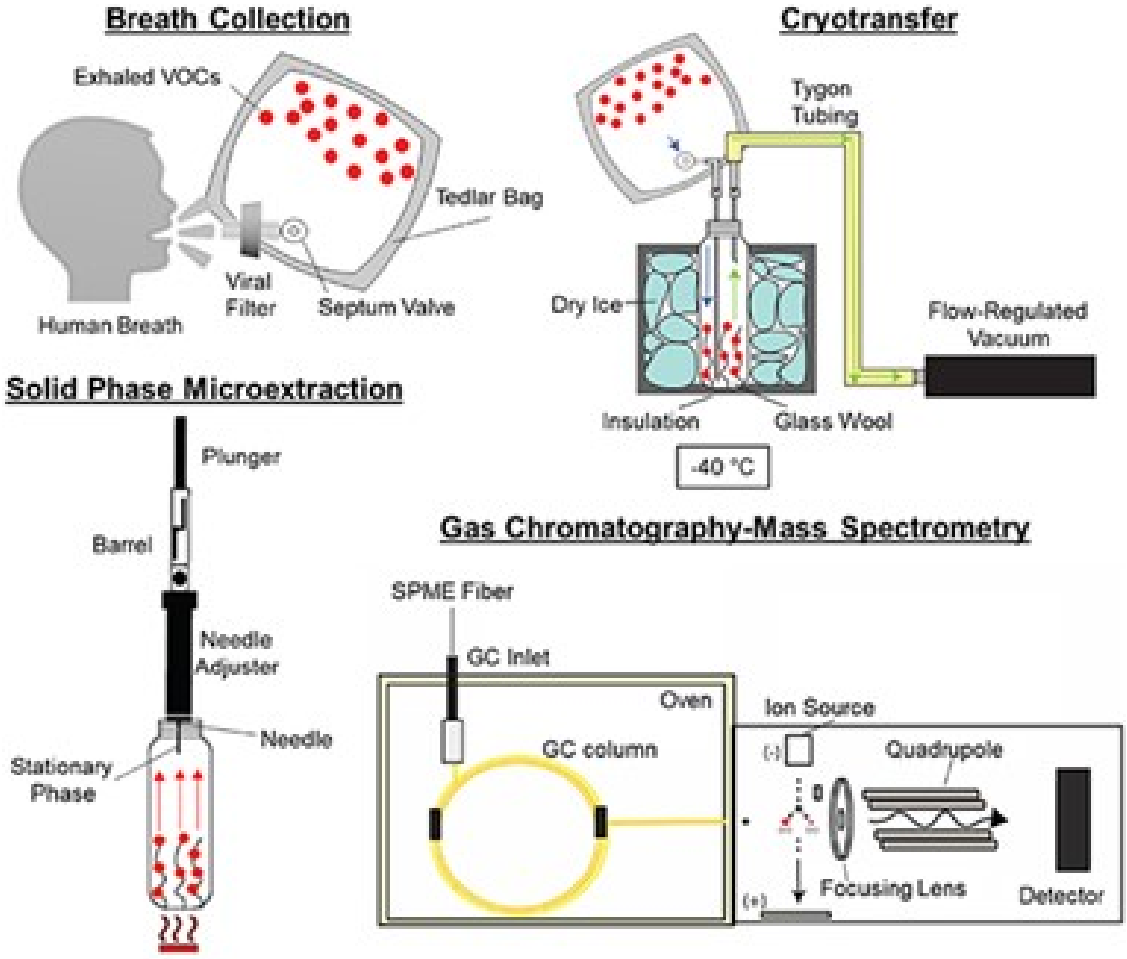

Figura 5.3: Protocolo de recolección de muestras con bolsas Tedlar y análisis posterior. (Fuente: [3])

- Condensación de aliento exhalado: el protocolo más usado para la condensación de aliento exhalado consiste en respirar por 10 minutos de forma normal a través de un tubo colocado en hielo o cualquier otro mecanismo que logre enfriarlo lo suficiente. Existen equipos comerciales que realizan este proceso de una forma más estandarizada como RTube, Turbodecss y EcoScreen [43].

- Nariz electrónica: uno de los métodos no invasivos de recolección de COVs más novedosos, basado en el uso de matrices de sensores químicos que responden de una forma distinta en base a los patrones de COVs, simulando el sistema olfativo humano. En el procedimiento habitual en la recolección de datos, los olores, ya sean de frutas, plantas, animales o aliento humano, entran a un circuito cerrado conectado al dispositivo, que registra los cambios en los parámetros eléctricos resultados por la interacción entre los compuestos y los sensores. Se utilizan algoritmos, como análisis de componentes principales (PCA) o análisis discriminante lineal (LDA), para el procesamiento y análisis de los datos obtenidos [44][45].

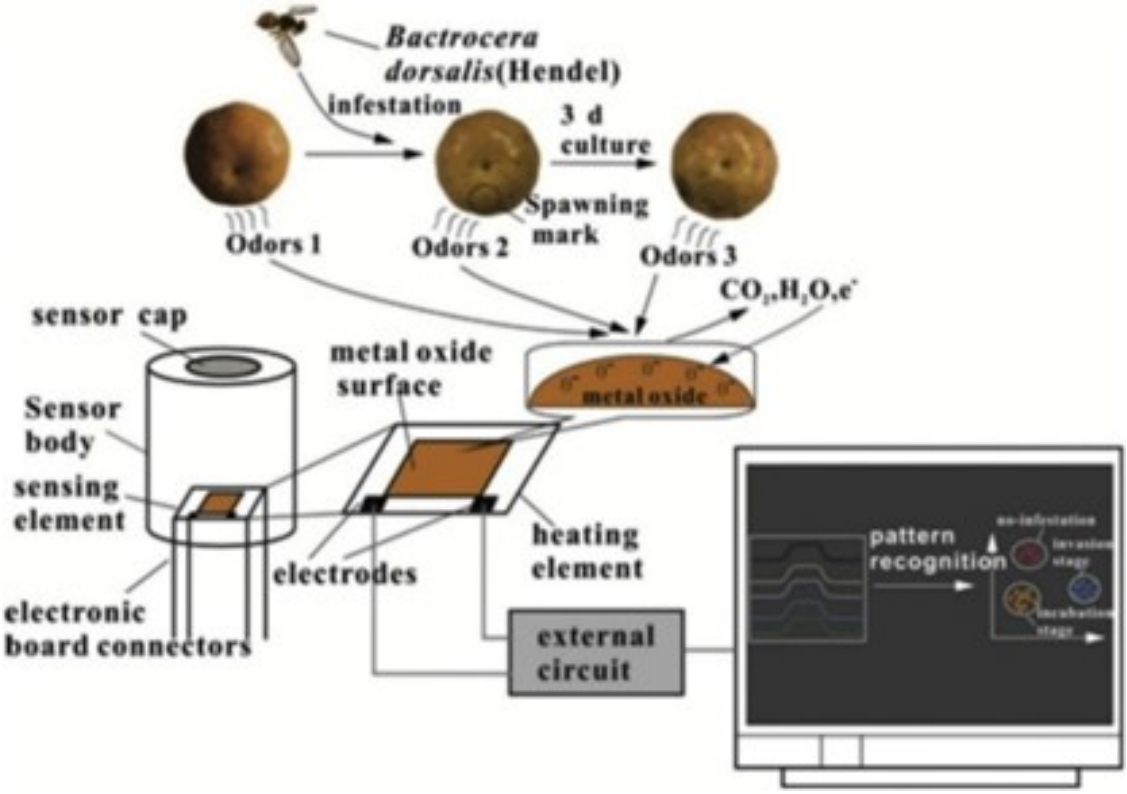

Figura 5.4: Principio de recolección de COVs por medio de una nariz electrónica. (Fuente: [4])

5.3.2. Diabetes Mellitus Tipo 2

Epidemiología y factores de riesgo La diabetes mellitus tipo 2 representa uno de los principales retos en cuanto a la salud pública a nivel mundial con una prevalencia y morbilidad que aumenta año tras año a causa de los hábitos de la población en la alimentación y el sedentarismo, factores determinantes en la patología.

Geográficamente y sociodemográficamente existen diferencias notables en los estudios epidemiológicos de diabetes mellitus tipo 2, en países desarrollados, como Estados Unidos y algunos países de Europa occidental, donde se lleva un estilo de vida occidental y se consumen más alimentos procesados, la prevalencia es mayor; no obstante, los países en vía de desarrollo, por la rápida urbanización y la transición nutricional, han experimentado un acelerado incremento. Por otro lado, existen regiones, como las islas del pacífico, en donde se ha encontrado una predisposición genética que a su vez interactúa con los cambios en la dieta [46][47].

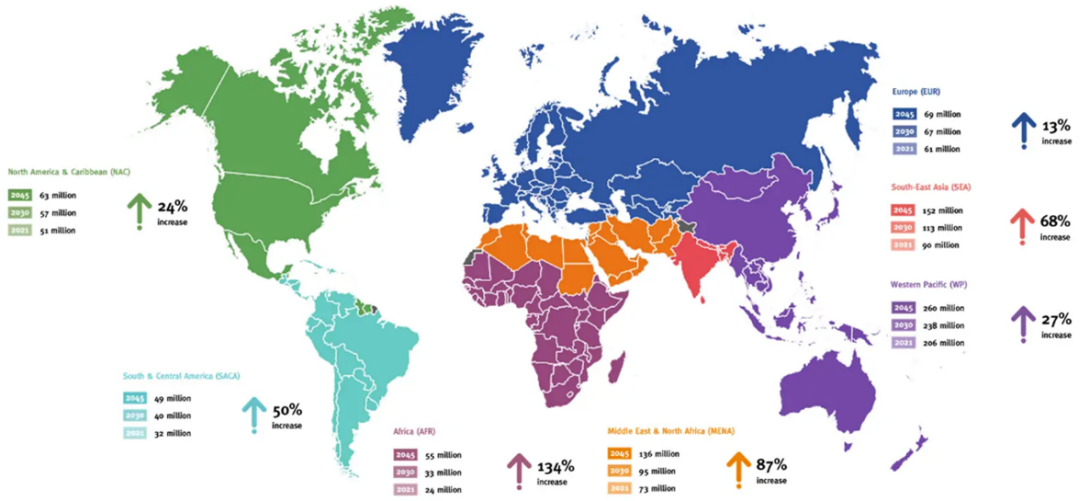

Figura 5.5: Proyección del incremento de la T2DM a nivel global entre 2021 y 2045 realizada por la Federación Internacional de Diabetes. (Fuente: [5])

La edad también es un factor de gran peso en la epidemiología de la T2DM. Las personas mayores son las que principalmente se ven afectadas por la enfermedad, sin embargo, el aumento de casos en jóvenes adultos y niños ha sido considerablemente grande, sobre todo en comunidades con alta prevalencia de obesidad infantil. A pesar de que no hay diferencias significativas en la prevalencia entre ambos sexos, en los hombres suele manifestarse de forma más temprana [46][48].

La T2DM se asocia a comorbilidades múltiples crónicas, y a desenlaces adversos, dentro de las complicaciones crónicas se encuentran enfermedades macrovasculares como la cardiopatía isquémica

y accidente cerebrovascular, o microvasculares como la retinopatía, nefropatía o neuropatía diabéticas. En un análisis reciente, se estableció que la enfermedad vascular periférica y la nefropatía diabética son las complicaciones de mayor prevalencia, mientras que la hipoglucemia severa y el estado hiperosmolar no cetósico son las más comunes entre las complicaciones agudas [47].

Fiopatología de la T2DM La T2DM se caracteriza por ser un trastorno metabólico progresivo que se da como el resultado de una resistencia a la insulina y la disfunción de las células β, encontradas en el páncreas. El fenómeno principal de esta enfermedad es la resistencia a la insulina, la cual es causada principalmente por la obesidad, pues la expansión ectópica del tejido adiposo genera un microambiente inflamatorio y disfuncional que altera las vías metabólicas normales, llevando al desarrollo de hiperglucemia crónica y, posteriormente, diabetes. Dicha inflamación sistémica, causada por la obesidad, juega un papel crítico en la resistencia a la insulina, esto se debe a que el estado inflamatorio está mediado por la secreción disfuncional de citoquinas proinflamatorias. Estas moléculas, al activar vías de señalización, interfieren con la señalización de la insulina, lo que promueve la lipotoxicidad, es decir, la acumulación de lípidos tóxicos en órganos periféricos como el hígado y los músculos esqueléticos [6].

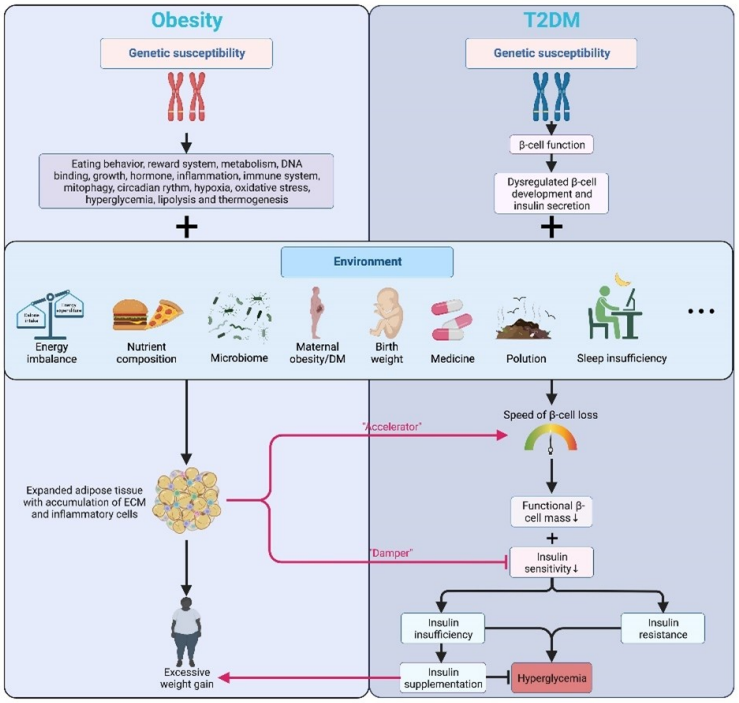

Figura 5.6: Relación entre obesidad y diabetes mellitus tipo 2. Interacción entre susceptibilidad genética y factores ambientales. (Fuente: [6])

En cuanto al impacto a las células β pancreáticas, la acumulación de grasa ectópica en la obesidad, en etapas iniciales, lleva a estas células a intentar compensar la resistencia a la insulina

mediante hiperplasia y un aumento de la secreción de insulina, lo que al final conlleva a una hiperinsulinemia. Además, la exposición crónica a hiperglucemia y niveles elevados de lípidos lleva al desarrollo de glucolipotoxicidad, causando estrés oxidativo, estrés retículo endoplásmico y disfunción autoinmune, culminando en la apoptosis progresiva de las células β y, por consiguiente, en la disminución de la capacidad secretora de insulina [6] [7].

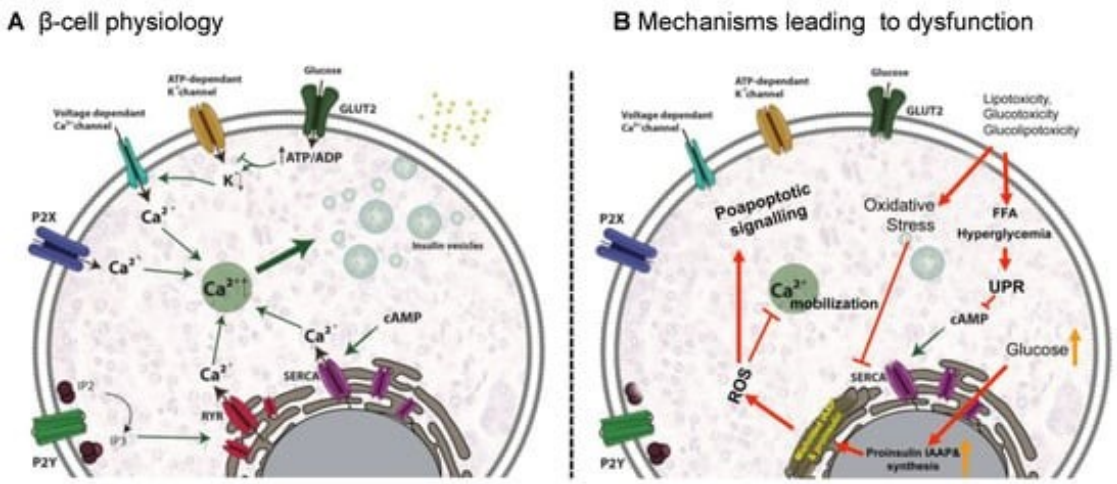

Figura 5.7: Condiciones fisiológicas de las células β pancreáticas; (A) Condiciones normales; (B) Frente a la diabetes mellitus tipo 2. (Fuente: [7])

5.3.3. Relación entre COVs y enfermedades metabólicas

Impacto del metabolismo alterado en la producción de COVs En respuesta a las alteraciones metabólicas, las concentraciones y perfiles de los compuestos orgánicos volátiles suelen variar. Estas disfunciones, que pueden ser el resultado de diversos factores como la obesidad, la diabetes mellitus tipo 2 y otras enfermedades metabólicas, se encuentran directamente relacionadas al estrés oxidativo y la inflamación crónica, procesos que modifican las rutas metabólicas esenciales.

La obesidad provoca un estado inflamatorio de bajo grado que induce la producción de especies reactivas de oxígeno, las cuales son moléculas potencialmente dañinas, pues pueden causar daño en lípidos, proteínas y ácidos nucleicos, y producen subproductos volátiles en el proceso de peroxidación de los lípidos, como aldehídos y ciertos hidrocarburos [49]. La acumulación de especies reactivas de oxígeno causa el arranque de otras vías metabólicas alteradas, como el ciclo de los polioles y las rutas de la glicación avanzada, lo que se traduce de nuevo en un incremento en la producción de COVs a partir de ácidos grasos poliinsaturados y de proteínas dañadas. Las moléculas volátiles producidas reflejan todo el daño metabólico, así como la actividad de las enzimas prooxidativas, y son detectadas también en la respiración y en otros fluidos biológicos [50].

En la diabetes, el efecto de estas alteraciones en la producción de COVs es aún más evidente e importante debido al control glucémico insuficiente y a la activación prolongada de las vías oxida-

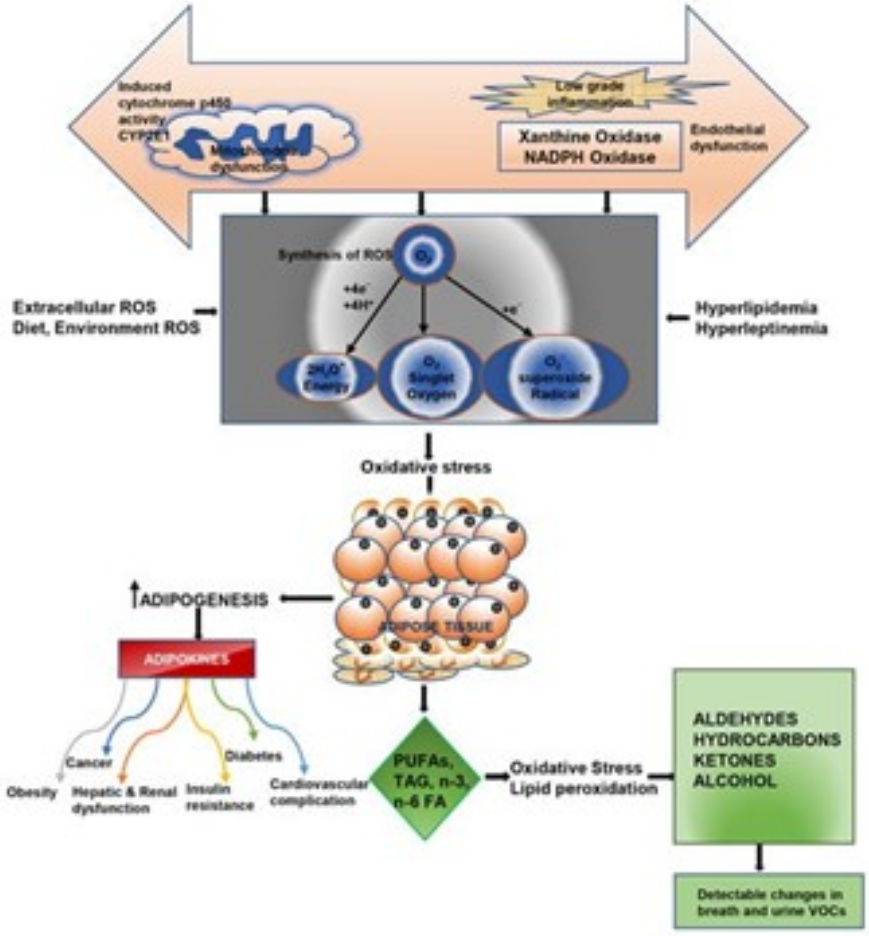

Figura 5.8: Mecanismos de relación entre la producción de COVs y estrés oxidativo. (Fuente: [7])

tivas. En la hiperglucemia crónica, el incremento de la producción de especies reactivas de oxígeno y de las especies reactivas de nitrógeno da lugar a compuestos como la acetona, el 2-propanol y los aldehídos volátiles [51]. Estos compuestos pueden ser tomados como marcadores del daño oxidativo, pero también forman parte de las alteraciones en la oxidación de los ácidos grasos y de los aminoácidos, las cuales se ven potenciadas por la disfunción mitocondrial [49]. La evidencia también muestra que las alteraciones metabólicas producidas por la obesidad y por la T2DM impactan en la actividad enzimática en el tejido adiposo y en otros órganos. La sobrerregulación de la familia CYP450, famosa por realizar reacciones oxidativas, puede incrementar la producción de alcoholes, cetonas y aldehídos, que son tomados del daño de las moléculas lipídicas y proteicas [49]. Este proceso se relaciona con el aumento en la peroxidación de lípidos y en los COVs que se perciben en la respiración de los pacientes obesos y diabéticos. Adicionalmente, el estrés hipóxico en tejidos metabólicamente activos contribuye a la producción de COVs debido a la activación de vías alternas de metabolismo oxidativo [49] [50]. De esta forma, en el contexto de las variaciones de los perfiles

metabólicos, también se producen cambios en la microbiota intestinal, la cual es un elemento importante en la producción de ciertos COVs. La disbiosis, alteración muy presente en la obesidad y la diabetes favorece el aumento de los volátiles generados a partir del metabolismo bacteriano como puede ser el amoníaco y los sulfuros de hidrógeno. Dichas alteraciones no son sólo un efecto de las alteraciones metabólicas, sino que también favorecen un incremento en el estrés sistémico y las rutas oxidativas [50].

COVs identificados como biomarcadores de la T2DM

Cuadro 5.2: Relación entre compuestos orgánicos volátiles (COVs) y la diabetes tipo 2 (DT2).

COV	**Origen Metabólico**	**Relación con DT2**	**Matriz de Detección**	**Referencia**
Acetona	Oxidación de ácidos grasos durante cetosis	Indicador de resistencia a la insulina y cetosis	Aliento	[52], [53], [49]
Isopreno	Metabolismo del colesterol	Relacionado con disfunción hepática	Aliento	[52]
Hexanal y Octanal	Peroxidación lipídica	Marcadores de daño oxidativo	Plasma, aliento	[52]
Ácidos grasos de cadena corta (propiónico y butírico)	Microbioma intestinal	Disbiosis intestinal asociada a DT2	Heces, sangre	[53]
Trimetilamina	Metabolismo hepático y renal	Indicador de complicaciones hepáticas y renales	Aliento, orina	[53]
2-Propanol y Metanol	Estrés oxidativo	Indicadores de disfunción metabólica	Aliento	[52]
Fenilacetaldehído y Benzaldehído	Oxidación de aminoácidos aromáticos	Alteraciones metabólicas en estados hiperglucémicos	Plasma, orina	[52]
Sulfuro de hidrógeno y Metilmercaptano	Metabolismo bacteriano	Asociados a complicaciones renales y hepáticas	Orina, aliento	[53]
Hexano	Peroxidación lipídica en lípidos poliinsaturados	Biomarcador de estrés oxidativo, asociado con obesidad y daño celular por ROS	Aliento, sangre	[49]

COV	Origen Metabólico	Relación con DT2	Matriz de Detección	Referencia
Pentano	Peroxidación lipídica	Indicador no invasivo de estrés oxidativo en obesidad	Aliento	[49]
Etano	Peroxidación lipídica	Indicador de estrés oxidativo; vinculado a obesidad y daño en membranas celulares	Aliento	[49]

5.3.4. Cromatografía de gases acoplada a espectrometría de masas

La cromatografía de gases acoplada con la espectrometría de masas (GC-MS), es una técnica analítica que se utiliza ampliamente en la comunidad científica y en la industria gracias a que permite separar, identificar y cuantificar compuestos orgánicos, ya sean volátiles o semivolátiles.

Cromatografía de Gases La cromatografía de gases (GC) es una técnica de separación que depende de la volatilidad y de la interacción que tienen con una fase estacionaria. La cromatografía de gases se introdujo por primera vez en el año 1952, dentro de los primeros avances en esta tecnología estuvo el desarrollo de la cromatografía gas-líquido que era un método para separar mezclas complejas de compuestos volátiles [54].

La GC emplea un gas portador, por ejemplo, helio e hidrógeno que transporta la muestra a través de una columna de fase estacionaria. La separación de los compuestos se realiza en función de su interacción con la fase estacionaria y su volatilidad. La temperatura es fundamental porque determina la tasa de elución y la resolución de los picos de la cromatografía [54] [55].

El tipo de columna seleccionada es determinante para la eficiencia de la separación. Las columnas capilares, cuyo diámetro interno puede variar entre 0,10 y 0,53 mm, son las más utilizadas por su elevada resolución y capacidad de separación. Por otro lado, el gas portador debe ser seco, inerte y libre de impurezas para evitar interferencias con la muestra [54].

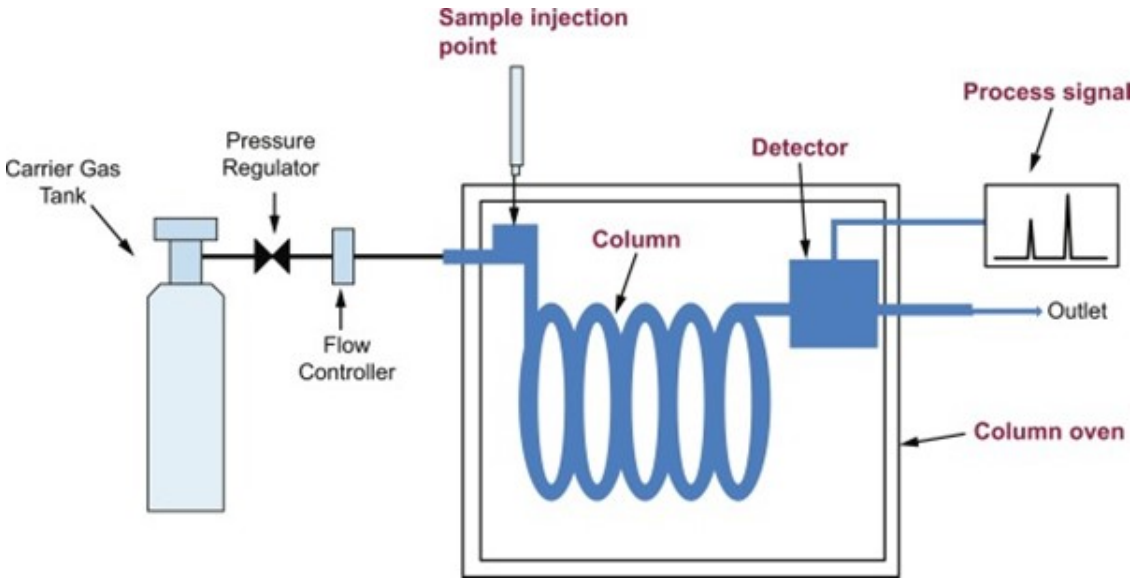

Figura 5.9: Diagrama simplificado de un sistema de cromatografía de gases. (Fuente: [8]

La GC se utiliza extensamente en análisis medioambientales, investigación en alimentos, control de calidad en la industria química y farmacéutica, así como detección de compuestos volátiles [56].

Espectrometría de Masas La espectrometría de masas (MS) es una técnica analítica utilizada para la identificación de compuestos mediante la determinación de la relación entre la masa y la carga (m/z) de sus iones. Su desarrollo data del siglo XX, pero su acoplamiento con la GC y la posterior optimización, han permitido conocer avances muy importantes en el campo de los análisis químicos tanto cualitativos como cuantitativos.

La MS genera iones a partir del bombardeo de moléculas con electrones de alta energía, fenómeno que se denomina ionización por impacto electrónico. Posteriormente, por medio de un analizador de tipo cuadrupolo, de trampa de iones o de tiempo de vuelo (TOF), dichos iones se separan en función de la relación m/z obtenida [55].

La espectrometría de masas produce un espectro exclusivo para cada compuesto, lo que permite la identificación de cada uno, esto también gracias a la comparación efectuada a través de bases de datos de los espectros, en el análisis cualitativo, y en el análisis cuantitativo sólo se contempla la detección de los iones relevantes asociados con los compuestos considerados, mediante lo que se denomina patrones internos [55] [56].

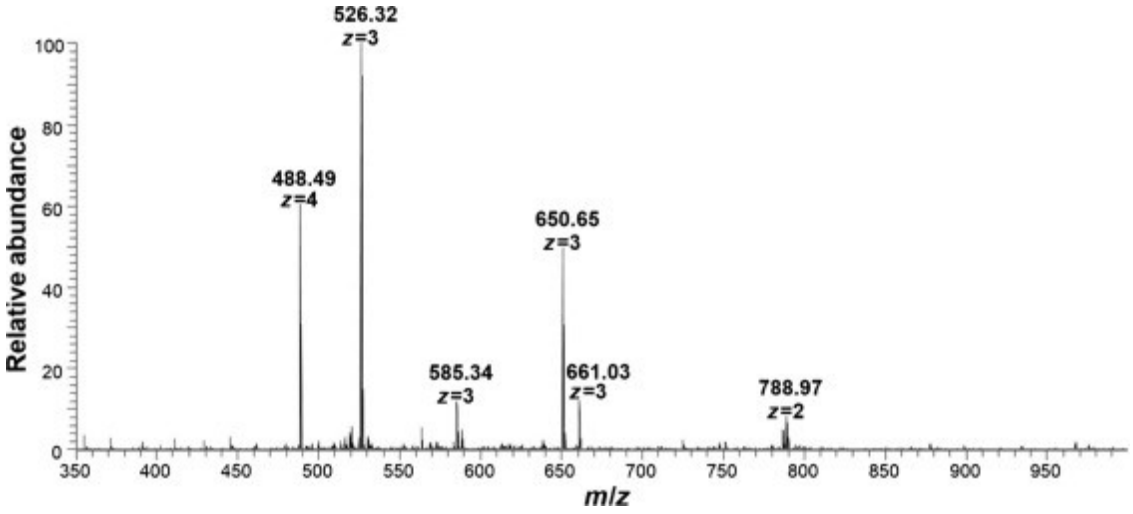

Figura 5.10: Espectro de masas obtenido por espectrometría de masas. (Fuente: [9])

A pesar de esa versatilidad, este método presenta limitaciones en algunos aspectos, como lo pueden llegar a ser la inestabilidad de los compuestos más volátiles, pues requieren un procedimiento de derivatización. Por otra parte, el método no tiene capacidad para distinguir unos isómeros de otros, basándose tan sólo en sus espectros de masas [55].

Acoplamiento Cromatografía de Gases - Espectrometría de Masas (GC-MS) El acoplamiento de estas dos técnicas se dio en los años 50 mientras se trataba de dar solución al problema de incompatibilidad de presión de ambas técnicas mediante el uso de separadores de flujo y el acoplamiento directo, lo que permite la combinación entre la capacidad de separación de la GC y la identificación de compuestos del MS [55].

Los sistemas de GC-MS permiten realizar un análisis de forma cualitativo y cuantitativo con una alta sensibilidad y especificidad. Su uso resulta fundamental en el ámbito de la metabolómica, el control de calidad y la detección de contaminantes en matrices complejas como los alimentos y el agua [56].

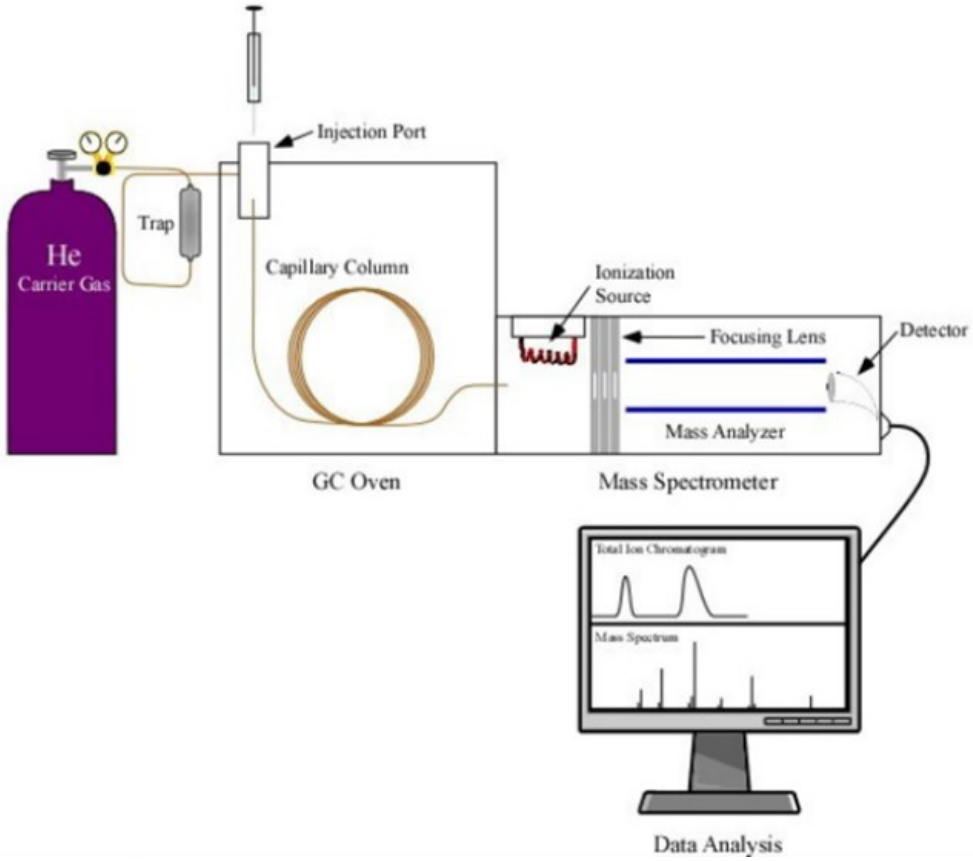

Figura 5.11: Diagrama de un sistema de cromatografía de gases acoplado a espectrometría de masas (GC-MS). (Fuente: [9])

Con el desarrollo de técnicas como el monitoreo de iones seleccionados (SIM) y espectrometría de alta resolución (HR-MS), la sensibilidad y la exactitud en el análisis por medio de la técnica GC-MS han mejorado notablemente. Se han desarrollado columnas de microdiámetro y se ha implementado un calentamiento resistivo, lo que ha permitido reducir significativamente los tiempos de análisis y alcanzar métodos en menos de un minuto [56].

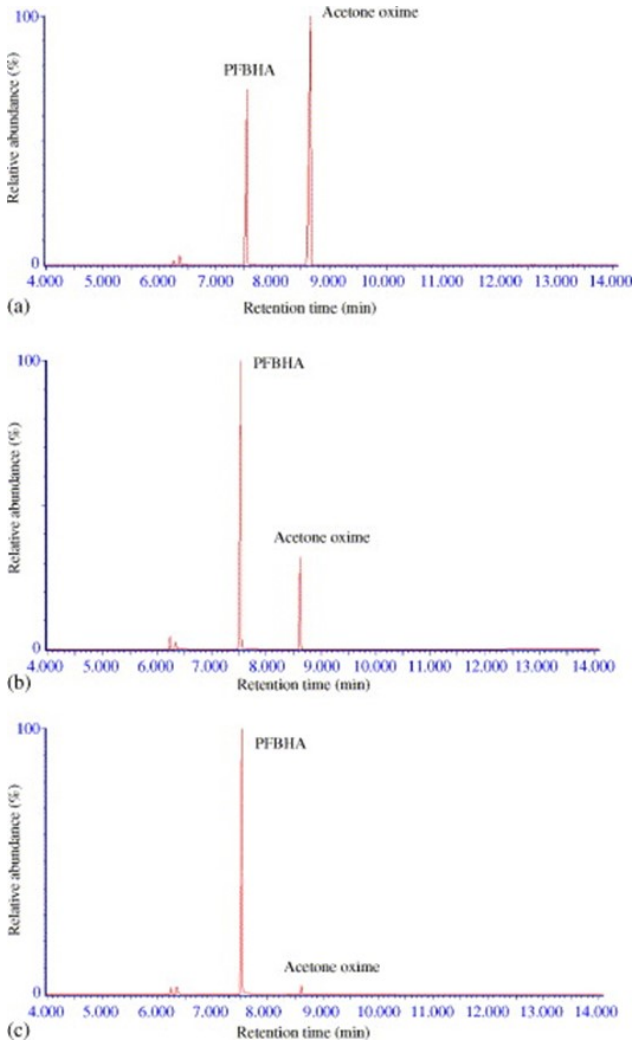

Figura 5.12: Cromatogramas obtenidos mediante GC-MS mostrando los picos correspondientes a PFBIHA y acetona oxima en diferentes condiciones experimentales. (Fuente: [10])

5.3.5. Microextracción en fase sólida SPME

La microextracción en fase sólida, SPME por sus siglas en inglés (Solid-Phase Microextraction), es una técnica moderna de preparación de muestras, rápida y efectiva, que se realiza en un único paso, combinando las etapas de muestreo, extracción y preconcentración de analitos. Esta técnica fue introducida por Pawliszyn et al. en la década de los años 90 del pasado siglo, ante la necesidad de la existencia de procedimientos de preparación de muestras más rápidos, libres de disolventes y que tuvieran una adaptabilidad a los preceptos recogidos en la química analítica verde. La microextracción en fase sólida ha cobrado popularidad en los últimos años debido a su sencillez, bajo coste, versatilidad y aptitud a la hora de trabajar matrices complejas y compuestos que se encuentren en

concentraciones traza [57].

La técnica de microextracción en fase sólida, SPME, se basa en el equilibrio que puede establecerse entre una fase de muestra y una fase extractora recubierta sobre una fibra de sílice fundida o sobre otras superficies, en función del montaje y ejecución de la técnica. La transferencia de masa queda determinada por el coeficiente de distribución entre la matriz de la muestra y el recubrimiento en la fibra, por lo que se permite la extracción de compuestos volátiles y semivolátiles. La fibra introducida en el medio que contiene los analitos se puede introducir, tras eliminar el exceso de muestra, en un instrumento analítico, como por ejemplo sistemas de cromatografía de gases o cromatografía líquida acoplados a espectrometría de masas para el análisis y la detección de componentes [57][11].

Modos de Operación La SPME puede realizarse en tres modalidades básicas: Direct Immersion (DI-SPME), donde la fibra se introduce en la muestra líquida o sólida haciendo que los analitos se adsorban en el extractor; Headspace SPME (HS-SPME), en el que la fibra se expone al espacio de cabeza que se genera sobre la muestra, siendo este un modo muy útil para la extracción de los analitos volátiles o semivolátiles; SPME con Membrana Protectora, utilizada para muestras que contienen interferentes de alto peso molecular como por ejemplo proteínas y que servirá para evitar que se dañe la fibra angostinando una buena parte de las proteínas.

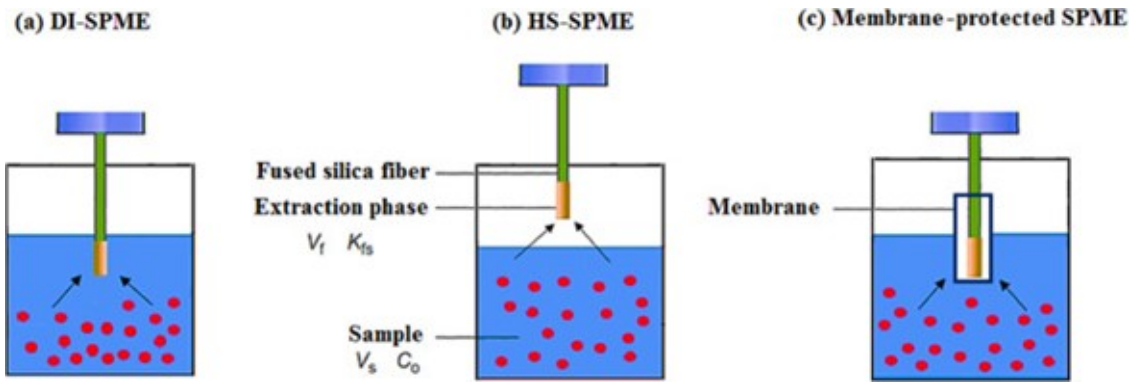

Figura 5.13: Esquema de diferentes técnicas de microextracción en fase sólida (SPME): (a) SPME en solución directa (DI-SPME), (b) SPME en espacio de cabeza (HS-SPME), y (c) SPME protegida por membrana. (Fuente: [11])

Aplicaciones en la Investigación Biomédica SPME ha demostrado ser una herramienta valiosa en el análisis de matrices biológicas complejas, como fluidos corporales (plasma, orina, saliva) y tejidos sólidos. Su capacidad para extraer compuestos en concentraciones traza lo hace ideal para estudios metabolómicos, farmacocinéticos y toxicológicos. Los recubrimientos biocompatibles han ampliado su aplicación para análisis in vivo, permitiendo la extracción directa de metabolitos y medicamentos en sistemas vivos con mínima invasividad. Esta característica es crítica para el monitoreo en tiempo real de biomarcadores y fármacos con ventanas terapéuticas estrechas, especialmente en áreas como la farmacología personalizada y la medicina de precisión [57].

5.4. Trabajos Relacionados

1. **Discovery of potential biomarkers in exhaled breath for diagnosis of type 2 diabetes mellitus based on GC-MS with metabolomics. [22]**

   Este artículo constituye el referente metodológico principal del presente trabajo. En él, los autores aplican un enfoque basado en cromatografía de gases acoplada a espectrometría de masas (GC-MS) junto con análisis metabolómico multivariado para diferenciar pacientes con diabetes mellitus tipo 2 (T2DM) de sujetos sanos a partir del aliento exhalado. El estudio empleó solid phase microextraction (SPME) como técnica de preconcentración, combinada con un análisis estadístico robusto que incluyó PCA, OPLS-DA, pruebas de Mann–Whitney y curvas ROC. Como resultado, se identificaron ocho metabolitos con potencial valor diagnóstico, entre ellos isopropanol, tridecano y acetona, alcanzando una sensibilidad del 97.9 % y una especificidad del 100 % en el modelo de regresión logística. Este trabajo no solo demuestra la viabilidad del análisis de COVs como herramienta no invasiva para el diagnóstico de T2DM, sino que también establece un protocolo integral desde la recolección de muestras hasta la validación estadística. Sin embargo, el enfoque de Yan et al. está orientado a pacientes que ya presentan la enfermedad en fase clínica, con el objetivo de confirmar el diagnóstico. En contraste, el presente proyecto se centra en identificar patrones de compuestos orgánicos volátiles asociados con un riesgo elevado de desarrollar diabetes mellitus tipo 2, antes de su manifestación clínica. Esta diferencia metodológica y temporal convierte nuestro estudio en una propuesta de tamizaje preventivo, con énfasis en la medicina predictiva y la detección temprana en poblaciones aparentemente sanas.

2. **Accuracy of breath test for diabetes mellitus diagnosis:A systematic review and meta-analysis. [58]**

   Este artículo habla sobre cómo detectar la diabetes mellitus tipo 2 analizando el aliento condensado que exhalamos. El objetivo principal es encontrar los compuestos químicos llamados compuestos orgánicos volátiles (COV) que puedan servir como señales claras para diagnosticar esta enfermedad; además, evalúa qué tan precisos pueden ser estos compuestos al usarlos para este fin. Esto es muy importante para este proyecto, ya que buscamos una forma no invasiva de diagnosticar la diabetes, algo que sería un gran avance en la medicina. También se mencionan ejemplos de otras pruebas similares, como las de aliento que ya se usan para detectar Helicobacter pylori, problemas gastrointestinales, asma o ictericia en bebés recién nacidos.

   La técnica más destacada en el artículo es la cromatografía de gases con espectrometría de masas (GC-MS), que se utiliza mucho para analizar el aliento de pacientes con diabetes. Esta técnica ayuda a identificar con precisión los compuestos relacionados con la enfermedad. El artículo nos ha servido para entender mejor la diabetes y cómo trabajar en su detección, pero no explica en detalle un punto clave: cómo recolectar correctamente las muestras de aliento.

   Recolectar bien las muestras es fundamental para que los resultados sean confiables. Si no se hace de manera adecuada, los compuestos detectados podrían no tener nada que ver con la enfermedad o podrían estar influenciados por cosas externas, como la contaminación del aire

o los hábitos del paciente. Este es un elemento esencial que necesitamos solucionar para el triunfo del proyecto.

3. **Exhaled breath condensate: Methodological recommendations and unresolved question. [59]**

   Este artículo se enfoca en la obtención de muestras de los pulmones mediante el condensado de aliento exhalado (EBC) como un método no invasivo. Describe la forma en que el EBC contiene varios mediadores como el amoníaco, el peróxido de hidrógeno y las citoquinas, que pueden ser señales de enfermedades respiratorias. Además, especifica los elementos conceptuales para la recolección de EBC, tales como la utilización de dispositivos particulares y las condiciones de almacenamiento, subrayando la importancia de regular los procesos para maximizar la capacidad de conservación de los datos. Este método es favorable para el proyecto, pues proporciona una base sólida para la recolección y estudio de compuestos orgánicos volátiles (COV) en el aliento exhalado, el cual es un componente esencial en el estudio sobre la diabetes mellitus tipo 2.

   Sin embargo, el trabajo del grupo que realizó este estudio, se basó en el intercambio de conocimientos entre expertos y se sigue pidiendo a todos los interesados en este campo que aporten datos relacionados con preguntas metodológicas para ayudar a comprender mejor las muestras y poder evaluar la utilidad de los datos obtenidos. Asimismo, es importante destacar que, a pesar de que el artículo se enfoca en gran parte del método de recolección, no discute en profundidad la relación concreta entre los compuestos orgánicos volátiles (COV) recogidos y enfermedades metabólicas como la diabetes, dejando un vacío en cómo relacionar estos biomarcadores con su relevancia en el diagnóstico.

4. **Management of Hyperglycemia in Type 2 Diabetes: A Consensus Report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD) [60].**

   Este estudio es un análisis detallado sobre la gestión de la hiperglucemia en la diabetes de tipo 2. Contiene sugerencias acerca de intervenciones con medicamentos, sin medicamentos y modificaciones en el estilo de vida, destacando la importancia de los agonistas del receptor GLP-1 y los inhibidores de SGLT2 para la defensa cardiorrenal. Además, muestra que es fundamental un enfoque personalizado que tenga en cuenta las preferencias y características de cada paciente. Este recurso es útil para nuestro proyecto porque amplía el marco clínico de la diabetes mellitus tipo 2, ofreciendo un esquema para entender la relevancia de los biomarcadores en el diagnóstico y control de la enfermedad. Sin embargo, el documento es muy bueno para hablar sobre la enfermedad e incluye un mayor enfoque en los determinantes sociales de la salud, el sistema de atención médica y los comportamientos relacionados con la actividad física, incluido el sueño. Así que nos ayuda a ampliar nuestros conocimientos sobre la diabetes mellitus tipo 2 pero no examina detalladamente otros métodos de diagnóstico, como la aplicación de compuestos orgánicos volátiles (COV) o la Cromatografía De Gases Acoplada A Espectrometría De Masas.

5. **A non-invasive method for the detection of glucose in human exhaled breath by condensation collection coupled with ion chromatography [61].**

   El estudio presenta un método analítico no invasivo para identificar glucosa en el aliento exhalado, empleando un dispositivo de recolección de condensado y cromatografía con detección amperométrica pulsada. También menciona que los métodos habituales para medir la glucosa en sangre implican la recolección de muestras sanguíneas, lo que causa dolor e incrementa la probabilidad de infecciones en los pacientes. Así pues, se desarrolló un protocolo simplificado que incorpora factores esenciales como la temperatura de refrigeración y el flujo de muestreo, alcanzando una correlación lineal entre los niveles de glucosa en sangre y en el aliento exhalado. Además, se observó que los pacientes con diabetes presentan niveles de glucosa en el aliento considerablemente superiores tras comer en comparación con personas saludables, lo que posiciona este método como una herramienta prometedora para el monitoreo de la enfermedad.

   El artículo está vinculado con la investigación de este proyecto, dado que se enfoca en la utilización del condensado de aliento exhalado (EBC) como un método para identificar biomarcadores asociados a la diabetes mellitus tipo 2. El uso de cromatografía para el análisis de la glucosa en el aliento mejora los objetivos del proyecto al investigar la relación entre los compuestos orgánicos volátiles (COV) y la enfermedad.Sin embargo, aunque este artículo es importante, no trata de forma completa la variedad de otros compuestos orgánicos volátiles presentes en el aliento exhalado y su posible relación con fases tempranas del desarrollo de la diabetes mellitus tipo 2. Además, a pesar de que se enfoca en la glucosa, no analiza la relación de otros biomarcadores que podrían ofrecer un diagnóstico más sólido y preciso.

6. **Gas chromatography-mass spectrometry-based analytical strategies for fatty acid analysis in biological samples [62].**

   El estudio explica cómo los ácidos grasos, esenciales en los sistemas biológicos, pueden estar vinculados con varias enfermedades cuando se encuentran desbalanceados. Sugiere métodos de análisis fundamentados en la cromatografía de gases-espectrometría de masas (GC-MS) para investigar ácidos grasos en diversas muestras biológicas, tales como plasma y tejidos. Resalta la relevancia de técnicas exactas de extracción, derivación y separación para entender a fondo estas moléculas. Se examinan diversas técnicas frecuentemente empleadas para la extracción de ácidos grasos, tales como la extracción líquido-líquido y la microextracción en fase sólida. Además, resalta que tanto los métodos de derivación ácida como básica, junto con otros métodos diseñados específicamente, fueron examinados de forma minuciosa, debatiendo sus fortalezas y limitaciones.

   La investigación de los compuestos orgánicos volátiles en el aliento exhalado condensado (AEC) comparte similitudes con el estudio de los ácidos grasos, dado que ambas intentan establecer vínculos entre elementos químicos particulares y enfermedades metabólicas como la diabetes de tipo 2. Las técnicas de extracción y evaluación mediante GC-MS detalladas en el documento pueden ser modificadas para describir los compuestos en el aliento exhalado condensado.

Además, la identificación de biomarcadores en el plasma puede proporcionar un fundamento para contrastar los niveles de volátiles en el aliento y en otros fluidos del cuerpo. No obstante, el artículo no se centra particularmente en el estudio de compuestos orgánicos volátiles en matrices de gases como el aliento exhalado condensado. A pesar de que describe la cromatografía de gases, se enfoca en los ácidos grasos y no toma en cuenta las adaptaciones requeridas para analizar volátiles más livianos y de reacción variada. Esto deja un hueco en el uso directo de estos métodos al estudio del aliento y su posible uso en la identificación temprana de diabetes tipo 2.

Capítulo 6

Materiales y Métodos

6.1. Desarrollo de protocolos

La primera etapa de este proyecto consistió en la caracterización de los participantes en Padilla, Cauca, y la elaboración de un protocolo de muestreo.

Para llevar a cabo esto, se diseñó y ejecutó una estrategia de recolección de datos sociales y clínicos mediante encuestas estructuradas, cuyo propósito fue identificar variables de riesgo asociadas a la DM2 en la población objetivo. Estas encuestas fueron digitalizadas en la plataforma RedCap, facilitando la sistematización y análisis de la información.

Uno de los hitos técnicos fue el desarrollo y validación del protocolo de muestreo de aliento. Este procedimiento describe con detalle las condiciones necesarias para asegurar la calidad de las muestras de AEC, incluyendo el diseño del sistema de recolección, tiempos óptimos de exhalación y condiciones térmicas adecuadas para preservar los VOCs. Para evaluar la eficacia del dispositivo de condensación, se realizaron réplicas experimentales con seis sujetos, aplicando ciclos de recolección de 5, 10 y 15 minutos. En las primeras pruebas, se empleó un tubo Falcon acoplado a un recipiente metálico como medio de enfriamiento; no obstante, se evidenció que tras 12 minutos, el sistema perdía capacidad térmica, comprometiendo la eficiencia de captura.

Como solución, en una tercera réplica se reemplazó el contenedor metálico por un recipiente de mayor tamaño con hielo, el cual mantuvo de forma estable la temperatura requerida, permitiendo obtener muestras de calidad durante períodos prolongados. Los resultados obtenidos demostraron que existe una relación positiva entre el tiempo de recolección y la masa condensada, aunque se observó una mayor variabilidad interindividual a los 15 minutos. Con base en estos hallazgos, se definió un tiempo estándar de 10 minutos como suficiente para garantizar un volumen mínimo de 500 microlitros por muestra sin comprometer la comodidad del voluntario.

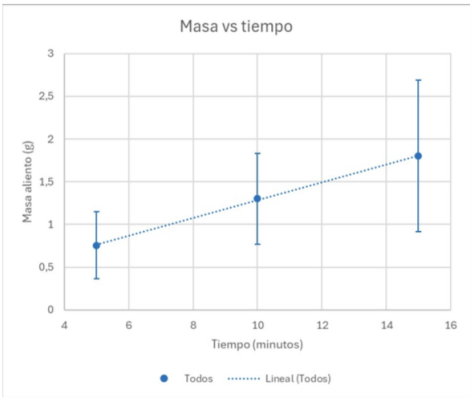

Figura 6.1: Relación entre el tiempo de recolección y la masa de aliento exhalado condensado

6.2. Evaluación del Riesgo de Diabetes Tipo 2

El riesgo de desarrollar diabetes mellitus tipo 2 (DM2) fue evaluado en una cohorte de 326 personas residentes en el municipio rural de Padilla, Cauca (Colombia), utilizando el cuestionario FINDRISC (Finnish Diabetes Risk Score), un instrumento validado internacionalmente para la predicción de riesgo a 10 años. Este cuestionario integra ocho ítems: edad, índice de masa corporal (IMC), circunferencia abdominal, nivel de actividad física, frecuencia de consumo de frutas y verduras, uso de medicamentos antihipertensivos, antecedentes personales de hiperglucemia y antecedentes familiares de diabetes. Cada ítem otorga una puntuación que, sumada, clasifica a los individuos en cinco categorías de riesgo [63].

Para esta investigación, se consideraron como grupo de alto riesgo (RA) aquellos participantes con puntuaciones ≥ 15 (riesgo alto o muy alto), y como grupo de bajo riesgo (RB) quienes presentaron puntuaciones ≤ 14. Los cuestionarios fueron diligenciados bajo supervisión directa para asegurar la validez de la información reportada.

6.3. Recolección de Muestras de Aliento Exhalado Condensado

Las muestras de aliento se recolectaron mediante un protocolo no invasivo cuidadosamente estandarizado, que garantizara la preservación de compuestos orgánicos volátiles (COVs) y evitara contaminaciones externas. Los participantes (n=46; RA=28, RB=18) exhalaron durante 10 minutos en tubos Falcon estériles de 50 mL, a través de un adaptador plástico, asegurando la recolección de al menos 500 µL de AEC por muestra. Las condiciones térmicas se mantuvieron a baja temperatura durante el proceso para evitar la degradación de compuestos volátiles.

Posteriormente, se añadieron 10 μL de una solución de control compuesta por hexano, tetracloruro de carbono, cloroformo y ciclohexano, a una concentración de 250 ppm. Las muestras fueron selladas herméticamente con *Parafilm*, etiquetadas y almacenadas a $-20\,^{\circ}$C hasta su análisis.

6.4. Preparación de Muestras

Previo al análisis instrumental, las muestras fueron descongeladas a temperatura ambiente, centrifugadas a 5000 rpm por 2 minutos para eliminar sedimentos o fases separadas, y finalmente homogeneizadas mediante agitación en vórtex durante 30 segundos. Este procedimiento garantizó la uniformidad y reproducibilidad de las muestras analizadas.

6.5. Justificación metodológica

6.5.1. GC-MS: metodología seleccionada

La cromatografía de gases acoplada a espectrometría de masas (GC-MS) representa la herramienta más robusta y confiable para la identificación de compuestos orgánicos volátiles (COVs), especialmente en el análisis de muestras de aliento condensado. Su funcionamiento se basa en una primera etapa donde los compuestos se separan mediante una columna cromatográfica en fase gaseosa, lo cual permite distinguir compuestos que pueden tener masas moleculares similares pero estructuras diferentes. Posteriormente, el sistema de espectrometría de masas fragmenta las moléculas y genera patrones únicos que se comparan con bases de datos estandarizadas, como la base del NIST (National Institute of Standards and Technology), considerada una de las más completas del mundo. La principal ventaja de utilizar GC-MS para nuestro proyecto radica en su altísima sensibilidad, capaz de detectar compuestos en concentraciones extremadamente bajas, incluso a niveles de partes por billón (ppb) o partes por trillón (ppt). Esta característica es fundamental porque los COVs presentes en el aliento humano suelen encontrarse en cantidades mínimas. Además, la especificidad que brinda GC-MS es insuperable, ya que no solo detecta, sino que también permite identificar con precisión cada compuesto presente en la muestra.

Otra gran fortaleza del GC-MS es su reproducibilidad. Gracias a la estabilidad de los tiempos de retención y a la estandarización de las bibliotecas de masas, se garantiza que los resultados sean comparables tanto entre diferentes muestras como entre diferentes ensayos. Este nivel de confiabilidad resulta esencial en investigaciones clínicas donde se requiere una validación estricta de los resultados obtenidos. Un beneficio adicional es que este equipo se encuentra disponible dentro del laboratorio en el cual se realizó nuestro proyecto, lo que nos permitió optimizar recursos económicos y humanos, evitando los sobrecostos asociados al envío de muestras a laboratorios externos y manteniendo un control total sobre los parámetros de análisis. Aunque el GC-MS requiere pasos adicionales como la derivatización de compuestos polares o la preconcentración mediante técnicas como SPME (microextracción en fase sólida), estos pasos son estándar en estudios científicos de metabolómica y no representan una desventaja significativa. Por el contrario, estas etapas aumentan la sensibilidad del análisis y permiten detectar compuestos que en otras técnicas pasarían desapercibidos. [64]

6.5.2. Comparación con metodologías alternativas

1. PTR-MS: La espectrometría de masas por reacción de transferencia protónica (PTR-MS) es una técnica avanzada y moderna que se utiliza principalmente para el análisis en tiempo

real de compuestos orgánicos volátiles en matrices gaseosas. Su funcionamiento se basa en un proceso de ionización blanda mediante la transferencia de protones desde un ion precursor, generalmente H_3O^+, hacia las moléculas de interés. Esta metodología ofrece ventajas indiscutibles, como la posibilidad de realizar análisis inmediatos sin necesidad de pretratamiento de la muestra, lo cual permite un monitoreo metabólico dinámico en tiempo real, algo que GC-MS no puede realizar. Además, la sensibilidad de PTR-MS es extremadamente alta, llegando a detectar compuestos en niveles de partes por trillón (ppt), lo que lo convierte en una opción interesante para estudios de aliento en contextos como pruebas de tolerancia a la glucosa o mediciones rápidas postprandiales.

Sin embargo, para los objetivos específicos de nuestro proyecto, el PTR-MS presenta desventajas muy relevantes. La principal limitación radica en que esta técnica no utiliza un sistema de separación cromatográfica previo a la detección, lo que significa que no es capaz de diferenciar isómeros o compuestos que compartan una misma relación masa/carga (m/z). Este aspecto es crítico para estudios de identificación de biomarcadores, ya que la especificidad en la identificación es fundamental para establecer relaciones clínicas entre ciertos metabolitos y el desarrollo de diabetes. Además, la técnica PTR-MS no proporciona una identificación completa del compuesto, sino que entrega señales basadas únicamente en la masa molecular, lo que puede derivar en ambigüedades y errores de interpretación. Otra limitación importante es que se requieren calibraciones frecuentes para corregir las respuestas del sistema y además no existen bibliotecas de espectros estandarizadas equivalentes a las del GC-MS, lo cual complica aún más la identificación precisa de compuestos poco estudiados. [65]

2. SIFT-MS: La espectrometría de masas mediante tubo de flujo de iones seleccionados (SIFT-MS) es otra alternativa emergente utilizada en la identificación y cuantificación directa de compuestos volátiles en tiempo real. Esta técnica funciona introduciendo iones precursores seleccionados (H_3O^+, NO^+, O_2^+) que reaccionan con los compuestos presentes en la muestra de gas, generando productos iónicos que son analizados directamente por el espectrómetro de masas. Una ventaja notable de SIFT-MS es que permite la cuantificación absoluta de compuestos sin necesidad de calibraciones externas muy complejas, además de contar con límites de detección también en rangos de ppb y ppt. Gracias a la posibilidad de utilizar diferentes iones precursores, ofrece una mayor cobertura de compuestos volátiles que el PTR-MS, lo cual puede ser útil en estudios exploratorios o en contextos clínicos donde se requieren respuestas rápidas y sin pretratamiento de muestras.

   Sin embargo, en el contexto de nuestro estudio, SIFT-MS también presenta desventajas significativas. Al igual que PTR-MS, no realiza una separación previa mediante cromatografía, por lo que no permite distinguir entre compuestos isoméricos o moléculas con masas similares. Esta falta de separación puede dar lugar a datos confusos o imprecisos, especialmente en un campo tan sensible como la identificación de biomarcadores metabólicos. Otro punto crítico es que, aunque SIFT-MS permite cuantificación absoluta, no posee bases de datos tan robustas como GC-MS, lo que dificulta aún más la identificación inequívoca de compuestos menos estudiados o de aparición reciente en el contexto biomédico. Además, la instrumentación es

considerablemente más costosa que la del GC-MS y su uso aún no está completamente estandarizado en aplicaciones clínicas, por lo que no existen protocolos ampliamente validados para estudios de prediabetes utilizando SIFT-MS. [66]

6.6. Optimización y Análisis por SPME-GC-MS

Complementariamente, se diseñó el protocolo de medición cromatográfica, el cual establece las condiciones experimentales para el análisis de los VOCs mediante GC-MS. Este protocolo fue elaborado por la asesora del proyecto, la estudiante de doctorado Viviana Ortega, siguiendo estándares internacionales y adaptándolo a las capacidades instrumentales del laboratorio. En él se detallan la preparación de las muestras, los parámetros de extracción en fase sólida (SPME) y las condiciones de operación del cromatógrafo y el espectrómetro.

6.6.1. Optimización del proceso

Antes del análisis de muestras biológicas, se desarrolló un diseño factorial completo con ayuda del software Statgraphics Centurion para optimizar los parámetros de extracción por microextracción en fase sólida (SPME).

En la siguiente tabla se muestran las tres variables independientes o factores evaluados. El diseño factorial permite estudiar simultáneamente los efectos individuales y combinados de estos factores sobre la eficiencia de extracción de compuestos volátiles.

Factor	**Límite inferior**	**Límite superior**	**Niveles**	**Unidades**
Oven extraction temperature (T)	35	75	3	°C
Fiber exposure time (tef)	5	240	3	Minutos
Fiber desorption time (tdf)	4	20	3	Minutos

Cuadro 6.1: Factores experimentales y niveles evaluados en el diseño del método SPME-GC-MS.

- Temperatura del horno de extracción (T): Controla la volatilización de los compuestos hacia el espacio de cabeza. Un aumento en la temperatura incrementa la presión de vapor de los analitos, pero un exceso puede degradarlos.
- Tiempo de exposición de la fibra (tef): Determina cuánto tiempo la fibra SPME está en contacto con el espacio de cabeza. Un mayor tiempo puede aumentar la captación de compuestos, pero también saturar la fibra.
- Tiempo de desorción de la fibra (tdf): Tiempo durante el cual la fibra libera los compuestos absorbidos dentro del inyector del GC. Un tiempo corto puede resultar en una desorción incompleta, y un tiempo excesivo puede dañar la fibra o introducir ruido.

Las soluciones estándar contenían acetona, 2-propanol, xileno y hexano, y se prepararon en matriz salina (NaCl al 3.6 en agua tipo 1). Los compuestos fueron seleccionados por su representa-

tividad en aliento humano y volatilidad. Se eligieron condiciones óptimas que maximizaban el área de los picos cromatográficos sin comprometer la resolución ni la reproducibilidad. Finalmente, se estableció una temperatura de extracción de 35°C, exposición de 40 min y desorción de 20 min.

Tras realizar el diseño factorial y analizar los resultados mediante modelos de superficie de respuesta, se identificaron rangos más estrechos y eficientes para los factores más relevantes. En este caso, se omitió el tiempo de desorción porque ya se estableció como óptimo (20 minutos), y se refinaron los rangos de temperatura y exposición.

Factor	**Límite inferior**	**Límite superior**	**Unidades**
Oven extraction temperature (T)	35°C	50°C	°C
Fiber exposure time (tef)	10 min	50 min	Minutos

Cuadro 6.2: Factores y rangos evaluados en la optimización del método SPME.

El análisis de superficie de respuesta mostró que fuera de estos rangos el rendimiento de extracción disminuía o se estabilizaba, por lo que se adoptaron estos límites para la etapa final del método. Esta optimización mejora la sensibilidad y reproducibilidad del análisis de aliento, evitando tiempos innecesarios o condiciones que puedan comprometer la estabilidad de los compuestos volátiles.

En cuanto a las condiciones instrumentales utilizadas durante el análisis de los COVs mediante GC-MS, utilizando SPME-HS, se realizó la siguiente tabla reuniendo las condiciones cruciales para garantizar la reproducibilidad, sensibilidad y resolución del análisis de aliento exhalado condensado.

Parámetro	Descripción
Headspace agitation	Se agitó la muestra durante la fase de extracción para mejorar la liberación de COVs hacia el espacio de cabeza.
Headspace oven time	Tiempo que la muestra permaneció en el horno previo a la extracción con la fibra SPME (15 minutos).
Injection port temperature	Temperatura del puerto de inyección donde la fibra libera los analitos al sistema GC-MS (250°C).
Split ratio	Modo de inyección en el sistema GC; se utilizó un modo dividido 1:1 para evitar sobrecarga del detector.
Column flow	Flujo de gas portador (helio) a través de la columna capilar (1.0mL/min).
Initial column temperature	Temperatura inicial del horno cromatográfico (40°C mantenida por 5 minutos).
Temperature ramp 1 y 2	Rampa de aumento progresivo de temperatura para lograr separación de compuestos volátiles de distinta polaridad y volatilidad.
Final stabilization	Etapa final donde la temperatura se mantiene constante (200°C por 5 minutos) para garantizar el arrastre de todos los compuestos.
Carrier gas	Se usó helio como gas portador por su alta pureza e inercia.
Solvent delay	Tiempo durante el cual el detector no registra señal, para evitar saturación por el disolvente (5 minutos).

Cuadro 6.3: Parámetros iniciales SPME-GC-MS.

6.6.2. Análisis de Muestras

El análisis se realizó con un equipo de cromatografía de gases acoplado a espectrometría de masas (GC-MS) utilizando una columna capilar Agilent DB-624 (60m × 0.25mm × 1.4µm). Se empleó una fibra SPME recubierta con polidimetilsiloxano y carbón de amplio rango (WR/PDMS, 95µm). La fase móvil fue helio ultrapuro (99.999), con un flujo constante de 1.0mL/min.

El detector de masas operó en modo scan entre 20–300m/z, comenzando en el minuto 5 del cromatograma (retardo de solvente). La inyección se realizó en modo split 1:1 a 250°C.

6.6.3. Reactivos y Estándares

Los solventes y compuestos utilizados como estándares internos incluyeron: hexano, tetracloruro de carbono, cloroformo, ciclohexano, acetona y 2-propanol, todos de grado analítico o LC-MS. Las soluciones se prepararon a 250 ppm y se añadieron a todas las muestras para permitir la estandarización relativa de áreas. Procesamiento y Análisis Estadístico de Datos El análisis de datos fue realizado mediante scripts desarrollados en lenguaje Python, los cuales incluyen las etapas de filtrado, normalización, pruebas estadísticas y visualización. Los códigos completos empleados para

dicho procesamiento se encuentran disponibles en el Anexo para fines de consulta y reproducción del análisis.

6.7. Procesamiento y Análisis Estadístico de Datos:

El análisis de datos fue realizado mediante scripts desarrollados en lenguaje Python, los cuales incluyen las etapas de filtrado, normalización, pruebas estadísticas y visualización. Los códigos completos empleados para dicho procesamiento se encuentran disponibles en el Anexo para fines de consulta y reproducción del análisis.

6.7.1. Preprocesamiento de datos en MS-DIAL

Los datos obtenidos por cromatografía de gases acoplada a espectrometría de masas (GC-MS) fueron preprocesados mediante el software MS-DIAL, una herramienta reconocida en el ámbito de la metabolómica no dirigida por su capacidad para realizar detección, alineamiento y anotación de señales provenientes de compuestos volátiles.

MS-DIAL opera mediante algoritmos de deconvolución espectral que permiten separar señales superpuestas y generar espectros individuales para cada compuesto detectado. Se configuró para modo de ionización electrónica (EI) con adquisición de datos en MS1 (centroid data), analizando un rango de masas entre 35 y 300 Da y un tiempo de retención entre 5 y 30 minutos. El umbral mínimo de intensidad para detección de picos se estableció en 500 unidades, con un nivel de suavizado de 3 escaneos y una anchura media de pico estimada en 20 escaneos.

La identificación de compuestos se realizó mediante comparación con bibliotecas espectrales (MSP) como la base de datos MoNA, usando un umbral mínimo de similitud espectral del 70 %. Solo se reportó el compuesto con mayor puntaje por espectro, y se aplicó una tolerancia de 0.1 minutos para el alineamiento basado en el tiempo de retención. Este procesamiento permitió generar una matriz de intensidades relativa a cada compuesto detectado en todas las muestras, la cual fue exportada para análisis estadístico.

Parámetro	Valor	Unidad / Comentario
Class ID	RA, RB, Standard, QC	Identificación de grupos de muestra.
Ionization type	Hard ionization	Se refiere al modo EI (Electron Ionization), típico en GC-MS.
Separation type	Chromatography	Tipo de técnica utilizada.
Collision type	Centroid Data	Tipo de datos generados (MS1, no espectros fragmentados).
Ion mode	Positive ion mode	Modo de detección de iones positivos.
Target omics	Metabolomics	Enfoque orientado al análisis de metabolitos.
Mass range	35–300 Dalton	Rango de masas buscadas en los espectros.
Retention time	5–30 Minutes	Ventana de tiempo cromatográfico a analizar.
Number of threads	4	Número de núcleos utilizados para procesamiento.
Minimum peak height	500	Umbral mínimo para considerar una señal como pico.
Smoothing level	3	Nivel de suavizado aplicado al perfil espectral.
Average peak width	20	Anchura promedio esperada para los picos.
Sigma window value	0.5	Valor relacionado con el algoritmo de suavizado.
EI spectra cut off	10	Umbral mínimo de intensidad para el espectro.
MSP file	Mona	Biblioteca espectral utilizada para comparación.
Only report the top hit / RT	Select	Se selecciona solo el compuesto con mejor coincidencia.
Retention time tolerance	0.1 Minutes	Tolerancia para el alineamiento de picos entre muestras.
EI similarity tolerance	70 %	Porcentaje mínimo de coincidencia espectral para identificación.

Cuadro 6.4: Parámetros utilizados en el preprocesamiento e identificación de compuestos en MS-DIAL.

6.7.2. Análisis de muestras blanco en MS-DIAL

De la misma forma se hizo un procesamiento en MS-Dial para 9 muestras blanco, esto con el fin de identificar posibles compuestos contaminantes, de este análisis se exportó una lista de los metabolitos encontrados.

6.7.3. Preparación y procesamiento de los datos

Con el fin de poder realizar un correcto análisis de los datos, usando el lenguaje de programación python, se llevaron a cabo los siguientes procesos:

- Filtro de muestras blanco: Se eliminaron los compuestos encontrados tras el análisis de las muestras blanco. Este filtro favorece el descarte de contaminantes ambientales, compuestos provenientes de los materiales de laboratorio, o interferencias del sistema instrumental.

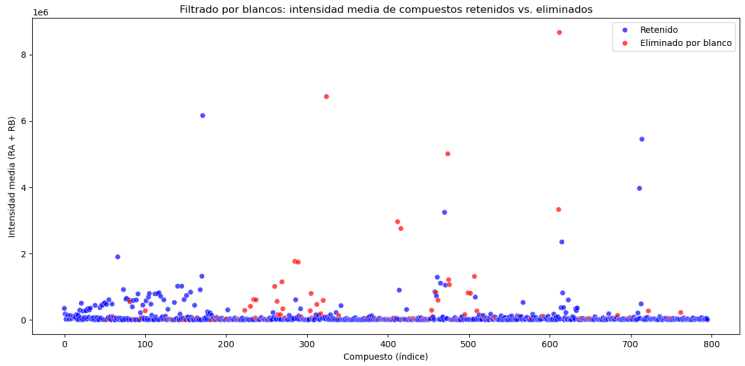

Figura 6.2: Scatter plot sobre la eliminación de compuestos detectados en muestras blanco.

- Filtro de compuestos ubicuos basado en percentil: Para reducir el sesgo asociado a compuestos con alta prevalencia pero escasa capacidad discriminativa entre grupos, se implementó un filtro basado tanto en la frecuencia de detección como en la intensidad relativa. Específicamente, se excluyeron aquellos metabolitos cuya intensidad superó el percentil 75 en más del 90 porciento de las muestras, bajo la premisa de que su presencia ubicua e intensidad consistentemente elevada podrían reflejar contaminantes sistemáticos, compuestos exógenos o señales de fondo inherentes al entorno analítico, más que verdaderos biomarcadores con valor diferencial.

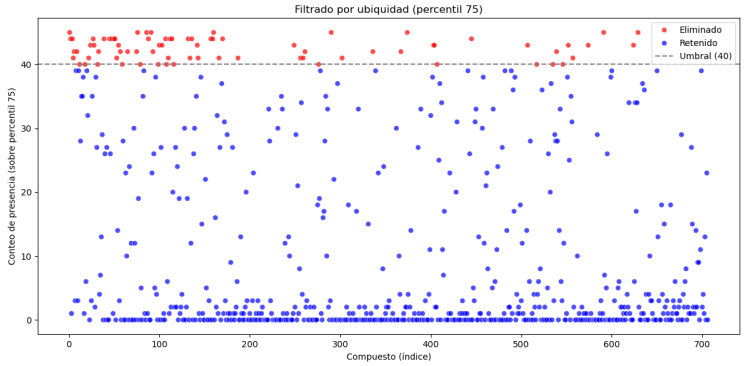

Figura 6.3: Scatter plot de metabolitos retenidos y eliminados tras el filtrado por ubiquidad.

- Normalización por suma total: Con el fin de mitigar el efecto de variaciones técnicas entre muestras y permitir una comparación más precisa de las abundancias relativas entre grupos, se aplicó una normalización por suma total. Esta técnica consiste en dividir el área de cada metabolito por la suma total de áreas de su respectiva muestra, transformando los valores absolutos en proporciones relativas. Esta estrategia es ampliamente utilizada en análisis metabolómicos para controlar diferencias sistemáticas en la carga total de compuestos detectados, que pueden surgir por variabilidad en la cantidad de muestra, eficiencia de extracción o sensibilidad instrumental, garantizando así que las comparaciones reflejen cambios biológicos reales más que artefactos técnicos.

6.8. Análisis estadístico:

Tras los procesos de filtrado y normalización, se procedió a la evaluación estadística de los metabolitos retenidos, con el objetivo de identificar aquellos compuestos cuya abundancia relativa difiere significativamente entre los grupos de alto riesgo (RA) y bajo riesgo (RB) de desarrollar diabetes mellitus tipo 2.

Desde una perspectiva metodológica, la selección de pruebas no paramétricas se fundamenta en dos aspectos clave de los estudios metabolómicos: la naturaleza frecuentemente asimétrica y heterogénea de los datos de abundancia relativa, que no suelen ajustarse a una distribución normal, y en segundo lugar, el tamaño muestral limitado, que restringe la robustez de supuestos paramétricos clásicos (por ejemplo, t-test).

En primera instancia, se calcularon las medianas de abundancia relativa para cada metabolito en los grupos RA y RB. El uso de la mediana como estadístico descriptivo central responde a la distribución no normal y a la presencia de valores atípicos, permitiendo una comparación robusta de tendencia central entre grupos.

A partir de estas medianas, se calculó el logaritmo en base 2 del cambio de pliegue (log_2FC) de cada metabolito, definido como:

$$\log_2 \left(\frac{\text{mediana}_{\text{RB}}}{\text{mediana}_{\text{RA}}} \right)$$

Este valor cuantifica la magnitud y dirección del cambio en la abundancia relativa del metabolito entre los grupos. Un log_2FC positivo indica una mayor abundancia en el grupo RB, mientras que un valor negativo refleja una predominancia en el grupo RA.

Posteriormente, se implementaron dos pruebas estadísticas para validar la significancia de estas diferencias:

- Prueba de Levene: aplicada a cada metabolito para evaluar la homogeneidad de varianzas entre grupos. Se utilizó un umbral de $p \geq 0{,}05$ para aceptar la hipótesis nula de igualdad de varianzas. Este paso se realiza para caracterizar la dispersión y descartar posibles sesgos por heterocedasticidad, lo cual, aunque no es requisito indispensable de la prueba de Mann-Whitney, permite describir mejor la calidad de los datos y sugiere estabilidad en la varianza.

- Prueba de Mann-Whitney-Wilcoxon (MWW): prueba no paramétrica utilizada para comparar la distribución de valores entre ambos grupos sin asumir normalidad. Para cada metabolito se calculó el estadístico de prueba y su p-valor asociado. Se consideró un umbral de significancia relajado ($p < 0.1$), común en estudios exploratorios de biomarcadores donde se prioriza la sensibilidad frente a la especificidad.

Los resultados de ambas pruebas se integraron para identificar aquellos metabolitos que presentaran diferencias significativas y homogeneidad de varianzas entre los grupos. Adicionalmente, se calcularon los valores de $-\log_{10}(p\text{-value})$ con el fin de facilitar su representación gráfica mediante diagramas tipo *volcano plot*.

A partir de este análisis, los compuestos se clasificaron en tres grupos:

- **Significativos con $\log_2$FC $>$0**: compuestos aumentados en el grupo RB.
- **Significativos con $\log_2$FC $<$0**: compuestos aumentados en el grupo RA.
- **No significativos**: compuestos con $p \geq 0{,}1$ o que no superaron el filtro de Levene.

Este enfoque combinó criterios estadísticos robustos con medidas de magnitud del efecto, lo que permitió una identificación más confiable de compuestos diferencialmente expresados. Los resultados fueron visualizados mediante un volcano plot, donde cada punto representa un metabolito y su posición refleja tanto la magnitud del cambio como su significancia estadística. Aquellos metabolitos clasificados como significativos fueron además exportados en archivos separados según su predominancia en RA o RB, facilitando su análisis posterior como posibles biomarcadores.

En conjunto, este esquema metodológico combina supuestos realistas sobre la naturaleza de los datos con prácticas estadísticas recomendadas para análisis de alta dimensionalidad, fortaleciendo la validez interna de los hallazgos y su potencial uso en investigaciones de validación clínica.

6.8.1. Correlación y análisis multivariado

Para explorar patrones de relación entre los metabolitos significativos, se calcularon coeficientes de correlación de Spearman entre todos los pares de compuestos. Esta prueba no paramétrica mide asociaciones monótonas, siendo adecuada para datos de distribución desconocida. Los coeficientes obtenidos fueron visualizados mediante mapas de calor, agrupando los compuestos mediante clústeres jerárquicos para facilitar la interpretación de posibles rutas metabólicas compartidas o asociaciones fisiopatológicas.

6.8.2. Representación gráfica de abundancia relativa

Finalmente, se construyeron diagramas de caja para cada uno de los metabolitos estadísticamente significativos, comparando su abundancia relativa entre RA y RB. Estas representaciones permiten observar la mediana, el rango intercuartílico y la presencia de valores atípicos, proporcionando una visión intuitiva sobre la variabilidad intra e intergrupal de los compuestos detectados.

Capítulo 7

Resultados y Discusión

De acuerdo con lo definido en la sección metodológica, que incluyó la recolección de muestras de aliento exhalado, análisis por GC-MS y procesamiento estadístico comparativo entre grupos, se presentan a continuación los resultados obtenidos, en donde se identificaron diferencias significativas en la abundancia relativa de compuestos entre los grupos RA y RB. Además de los resultados experimentales obtenidos, se realizó un análisis documental y teórico del contexto metabólico de los compuestos de interés.

1. Análisis de compuestos significativos
   Tras aplicar la prueba de Mann-Whitney-Wilcoxon para comparar las intensidades relativas de cada compuesto entre los grupos RA (alto riesgo) y RB (bajo riesgo), se generó un gráfico de volcán para visualizar los resultados.

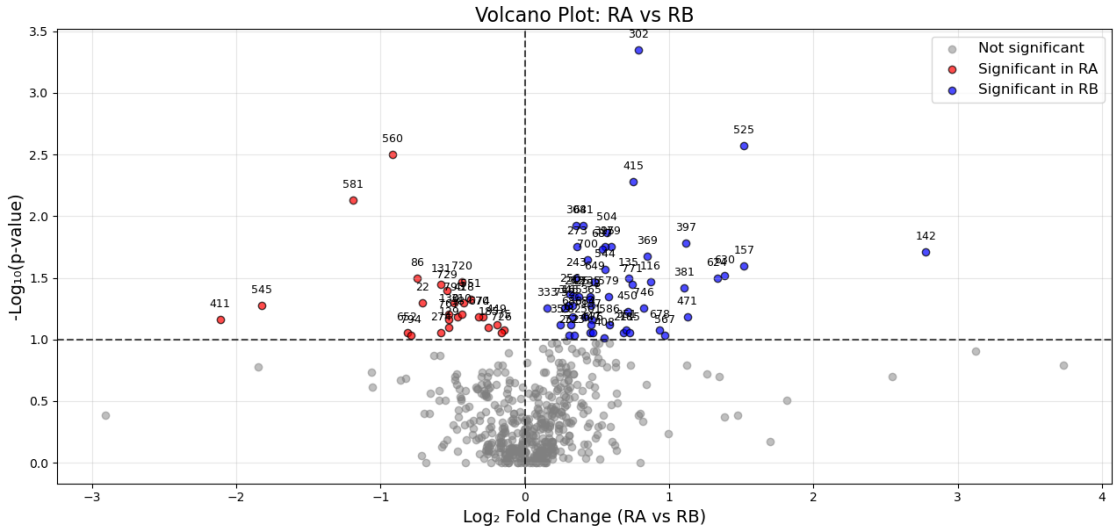

Figura 7.1: Gráfico de volcán de los compuestos diferenciadores entre RA y RB.

En el gráfico se pueden ver dos grupos, los compuestos en rojo, a la izquierda, son los compuestos significativos para el grupo de alto riesgo de desarrollar diabetes. Los azules, a la derecha,

son los significativos en el grupo de bajo riesgo de desarrollar diabetes. En la siguiente tabla se muestra el grupo de compuestos significativos en RA:

Alignment ID	**Metabolite name**
18	Unknown
22	Unknown
74	Unknown
75	Unknown
86	Unknown
129	2-PENTANOL
130	Unknown
131	Unknown
185	2-HEXANONE
274	octanone (2-)
310	Unknown
411	Unknown
418	BUTANEDIOIC ACID DIMETHYL ESTER
449	3,5-DIMETHYLPYRIDINE
545	Unknown
551	Unknown
560	Unknown
581	Unknown
652	Unknown
670	3(S)-CITRONELOL
720	2-PHENYLPROPIONALDEHYDE
726	Unknown
729	ETYLENEGLYCOL MONOPHENYL ETHER
767	Unknown
793	Unknown
794	Unknown

Cuadro 7.1: Compuestos significativos identificados en el grupo RA

En la tabla se observan 26 compuestos significativos encontrados en el grupo RA, de los cuales 8 fueron identificados como metabolitos conocidos y 18 permanecen como compuestos desconocidos.

2. Identificación de compuestos significativos más relevantes
   Para priorizar compuestos de interés, se seleccionaron los 4 compuestos desconocidos más

significativos (ID 581, 560, 411, 545), y se realizó su identificación con MassHunter y la librería NIST. Los compuestos identificados se pueden ver en la siguiente tabla:

ID	Sample	RT MSdial	RT Mass Hunter	Compound	Probability (%)
581	325	21.767	21.768	Octanal	87.2
560	177	21.450	21.448	5-Hepten-2-One, 6 methyl-	80.8
411	193	20.0967	20.096	Cyclotetrasiloxane, octamethyl-	87.1
545	283	20.084	20.089	Cyclotetrasiloxane, octamethyl-	85.5

Cuadro 7.2: Compuestos significativos identificados en el grupo RA con Mass Hunter y la librería NIST

Para poder continuar se revisaron los compuestos filtrados por el análisis de las muestras blanco, y se encontró que el Cyclotetrasiloxane, octamethyl- fue uno de los compuestos eliminados por el filtro. Por ello, se conservaron únicamente los compuestos 581 y 560 para el análisis posterior, que corresponden a los compuestos octanal y 6 methyl-5-Hepten-2-One respectivamente.

3. Estudio de correlación
   Se integraron los 8 metabolitos conocidos y los 2 identificados y aceptados (octanal y 6 methyl-5-Hepten-2-One) para generar un mapa de calor, el cual permitió observar patrones de coexpresión y correlación entre compuestos.

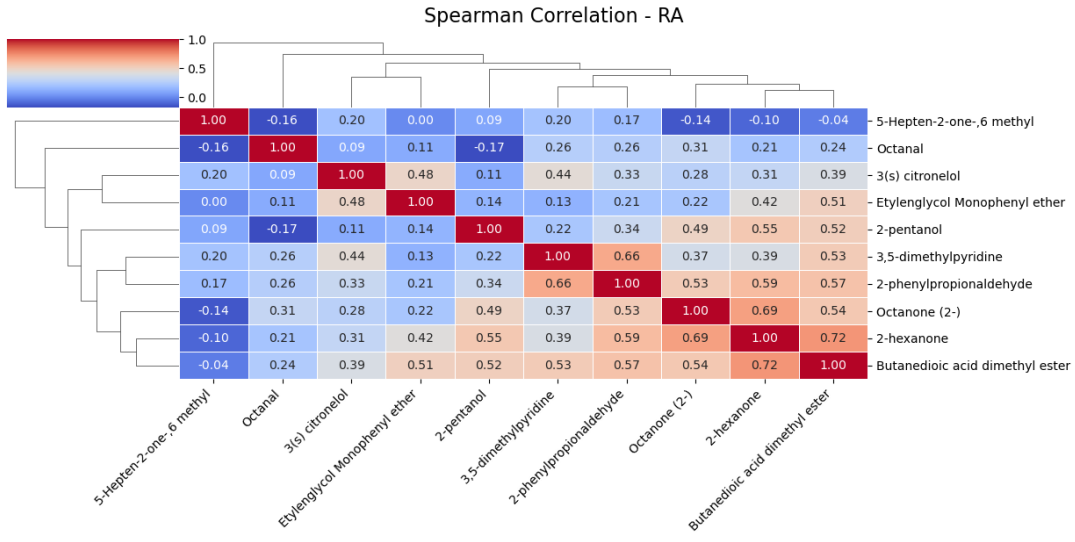

Figura 7.2: Mapa de calor de metabolitos significativos relevantes en RA.

En el mapa de calor se observa que los compuestos 2-phenylpropionaldehyde, octanone (2-), 2-hexanone y butanedioic acid dimethyl ester presentan una correlación positiva en sus patrones de expresión, lo que indica que tienden a aumentar o disminuir de forma conjunta en las muestras analizadas. Esta coexpresión sugiere que podrían estar modulados por factores comunes, lo cual abre la posibilidad de una exposición compartida, una respuesta fisiopatológica convergente, o una combinación de ambas.

4. Diagramas de caja
   Se construyeron diagramas de caja (boxplots) para comparar visualmente la distribución de las intensidades relativas entre RA y RB. En la mayoría de los casos, los compuestos mostraron una tendencia clara de sobreexpresión en RA, lo que refuerza su posible rol como biomarcadores respiratorios en etapas tempranas del riesgo de DM2.

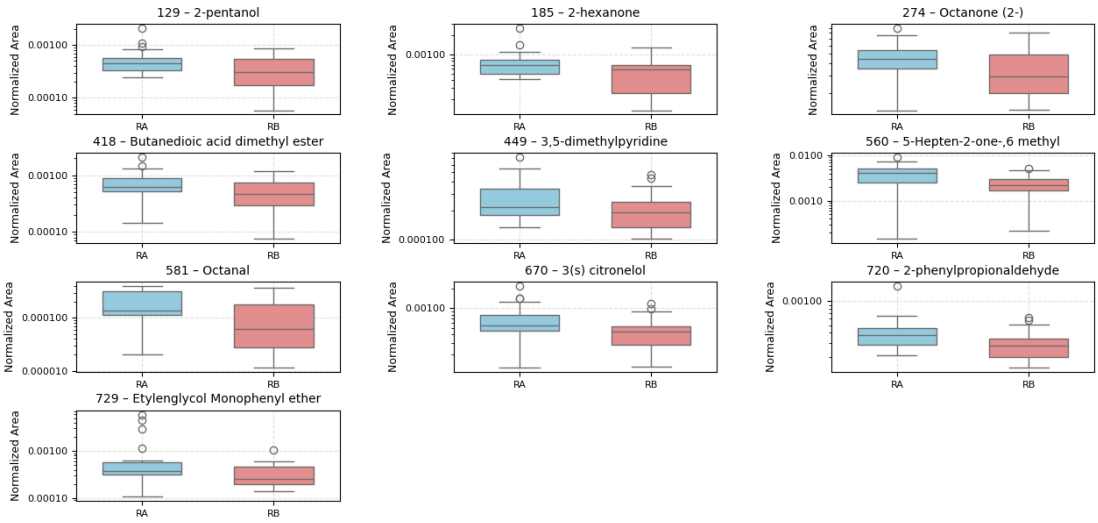

Figura 7.3: Diagramas de caja de los compuestos significativos relevantes en RA.

La distribución diferencial de los metabolitos en los diagramas de caja refuerza su potencial como discriminadores metabólicos para personas de riesgo alto de desarrollar diabetes, se observa que las medianas son consistentemente más altas en RA que en RB, ilustrando las diferencias significativas entre ambos grupos para cada uno de los compuestos.

5. Análisis del contexto metabólico de los compuestos

Con el fin de comprender la relevancia fisiopatológica de los compuestos identificados, se realizó un análisis del contexto metabólico de cada uno considerando su origen, ya sea endógeno o exógeno. Este análisis se sustentó en literatura científica, enfocándose en rutas metabólicas, condiciones fisiopatológicas asociadas y posibles mecanismos de producción o acumulación en el organismo humano.

- 2-Pentanol:
  El 2-pentanol es un alcohol secundario de producción exógena. Un estudio reciente evaluó la conversión metabólica de 2-pentanona a 2-pentanol en hepatocitos humanos y en voluntarios sanos y pacientes con cirrosis hepática. Tras la administración oral de 100 mg de 2-pentanona, se observó un aumento significativo de 2-pentanol en el aliento exhalado, alcanzando un incremento superior a 100 veces en los primeros 10 minutos. Los pacientes con cirrosis mostraron niveles más elevados de 2-pentanol en comparación con los controles sanos, lo que sugiere una alteración en la actividad de la alcohol deshidrogenasa hepática [67].
  Este hallazgo indica que el 2-pentanol puede ser detectado en el aliento humano tras la ingestión de su precursor, y su concentración puede reflejar el estado funcional del hígado. Sin embargo, no existen evidencias sólidas de que el 2-pentanol se produzca y exhale de manera significativa en condiciones fisiológicas normales o en ausencia de administración exógena de precursores [68].
  Por otro lado, un estudio publicado por Fernández del Río et al. (2015) identificó a la 2-pentanona como uno de los compuestos volátiles significativamente elevados en el aliento de pacientes con cirrosis hepática, en comparación con sujetos sanos. Tras el trasplante de hígado, los niveles de este compuesto disminuyeron notablemente, lo que respalda su relación con la disfunción hepática. Los autores proponen a la 2-pentanona, junto con metanol y limoneno, como biomarcadores respiratorios potenciales para el diagnóstico no invasivo de enfermedades hepáticas, e incluso para su detección en etapas tempranas [69].
  En este contexto, la detección de 2-pentanol en individuos con riesgo metabólico elevado, como los participantes del presente estudio, podría interpretarse como un marcador indirecto de una mayor producción endógena de 2-pentanona, reflejando posibles alteraciones subclínicas en el metabolismo hepático o en el manejo de cetonas. Dado que condiciones como la resistencia a la insulina y el síndrome metabólico afectan las rutas de oxidación de ácidos grasos y la eficiencia enzimática hepática, la presencia de 2-pentanol podría tener relevancia fisiopatológica como indicador de desregulación metabólica en estados prediabéticos.
- 3(S)-Citronelol:
  El citronelol es un monoterpeno característico de diversos aceites esenciales de origen vegetal. La presencia de citronelol en el cuerpo humano se atribuye exclusivamente a

fuentes exógenas, tales como la ingestión de alimentos o productos herbales, la inhalación de vapores aromáticos o la aplicación tópica de aceites esenciales, y no a procesos de biosíntesis interna [70].

- Etylenglycol Monophenyl ether:
  Es un compuesto químico sintético utilizado como conservante en cosméticos, fármacos y productos industriales. No forma parte de ninguna ruta metabólica humana natural. Su presencia en el organismo únicamente ocurre tras exposición exógena, como por ejemplo, absorción cutánea o ingestión accidental desde productos que lo contienen [71].
- Octanone (2-):
  La 2-octanona es una cetona alifática que pertenece al grupo de las metil-cetonas. Desde un enfoque bioquímico, la aparición de 2-octanona en el organismo puede explicarse por rutas alternativas. Una de ellas es la descarboxilación de β-cetoácidos derivados de ácidos grasos, proceso que puede ocurrir en determinados contextos metabólicos alterados o por acción de la microbiota [72]. De hecho, se ha propuesto que ciertas bacterias intestinales o presentes en otras superficies corporales son capaces de transformar ácidos grasos en metil-cetonas como la 2-octanona a través de procesos fermentativos o mediante la acción de enzimas específicas como las monooxigenasas [73].
  El estudio de Bishop et al. (2018) aporta una evidencia importante sobre la relevancia de la 2-octanona como biomarcador metabólico. En su modelo experimental con babuinos, la 2-octanona fue uno de los compuestos volátiles significativamente elevados en los animales que habían experimentado restricción de crecimiento intrauterino (IUGR), una condición que se asocia directamente con mayor riesgo de desarrollar enfermedades cardio-metabólicas en la vida adulta, incluyendo la diabetes tipo 2. Los investigadores observaron un aumento de hasta 15 veces en los niveles de 2-octanona en estos animales, coincidiendo con la presencia de resistencia a la insulina y un fenotipo prediabético incipiente. En su discusión, los autores plantean que este incremento puede estar reflejando un estado inflamatorio y de estrés oxidativo, elementos reconocidos en la progresión hacia la disfunción metabólica [74].
- 2-Hexanone:
  La 2-hexanona es una cetona volátil presente en lacas, pinturas y otros productos químicos. Sin embargo, existe cierta evidencia de que la 2-hexanona puede aparecer endógenamente. Su presencia ha sido confirmada en orina humana [75], en sangre periférica [76] y en aliento exhalado bajo hipoxia [77], lo que sugiere una producción metabólica interna. En un estudio reciente, Cozzolino et al. (2022) identificaron la 2-hexanona en muestras de orina de pacientes con diabetes mellitus tipo 2 (T2DM), NAFLD, y su comorbilidad. Se observaron niveles significativamente más altos de 2-hexanona en pacientes con T2DM en comparación con aquellos que presentaban simultáneamente NAFLD y T2DM, lo cual podría reflejar una regulación diferencial de las vías de oxidación de lípidos y de estrés oxidativo. Además, se destaca que su metabolito 2,5-hexanodiona puede interferir con la homeostasis hormonal y activar procesos inflamatorios como la vía del inflamasoma NLRP3, implicando un papel patológico relevante [78]. En conjunto, estos hallazgos

consolidan a la 2-hexanona como un biomarcador potencialmente útil en el contexto de trastornos metabólicos como la T2DM, con implicaciones en inflamación, disfunción hepática y daño oxidativo.

- Butanedioic acid dimethyl ester:
  También conocido como dimetil succinato, es un derivado esterificado del ácido succínico. Mientras que el ácido succínico tiene un rol fundamental en el metabolismo energético celular, el dimetil succinato es principalmente un compuesto de uso industrial y no forma parte de la bioquímica humana normal [79].
- 3,5-dimethylpyridine:
  También conocida como 3,5-lutidina, es un compuesto usado como producto químico, no se produce de forma endógena en el humano. Su presencia solo se explicaría por contaminación externa [80].
- 2-Phenylpropionaldehyde:
  También conocido como 2-fenilpropanal, es un aldehído aromático de cadena corta cuya presencia en organismos humanos se ha asociado principalmente con fuentes exógenas. Hasta la fecha, no se ha descrito una ruta metabólica endógena en humanos que justifique su biosíntesis, por lo que su detección en matrices biológicas suele atribuirse a exposición ambiental, dietaria o a procesos metabólicos microbianos [81].
- Octanal:
  Es un aldehído de 8 carbonos, también llamado aldehído caprílico, es un compuesto de producción endógena en humanos, principalmente como resultado de la oxidación de lípidos. El octanal ha sido identificado en diversas muestras biológicas, en especial en el aliento exhalado y en tejidos sometidos a estrés oxidativo [82]. Estudios de biomarcadores de cáncer, por ejemplo, han mostrado que ciertas enfermedades se asocian con niveles elevados de octanal originados internamente. En pacientes con cáncer de pulmón, las concentraciones exhaladas de octanal son significativamente más altas que en individuos sanos o fumadores control, lo que sugiere que el tumor o el microambiente oxidativo del cáncer generan este aldehído [83]. Una de las fuentes endógenas más importantes del octanal es la peroxidación lipídica de ácidos grasos insaturados presentes en los fosfolípidos pulmonares, especialmente el ácido oleico (ω-9) y el ácido palmitoleico (ω-7). Bajo condiciones de estrés oxidativo, especies reactivas de oxígeno (ROS) inducen la abstracción de hidrógenos en posiciones bis-alílicas de estas cadenas, iniciando un proceso de radicales libres que lleva a la formación de hidroperóxidos lipídicos. Estos intermedios pueden ser degradados por metales redox-activos o enzimas específicas, como las glutatión peroxidasas o citocromos P450, produciendo escisiones de enlaces carbono-carbono. Esta degradación da lugar a fragmentos aldehídicos volátiles, entre los cuales el octanal es uno de los productos predominantes cuando la escisión ocurre en ácidos grasos con dobles enlaces en posición 9, como el oleico [84].

  En enfermedades metabólicas como la diabetes mellitus tipo 2, la hiperglucemia crónica y la resistencia a la insulina generan un ambiente de estrés oxidativo persistente, promo-

viendo la peroxidación lipídica. Si bien la literatura disponible no posiciona al octanal como marcador principal de diabetes, su formación a partir de lípidos insaturados bajo condiciones oxidativas sugiere que podría encontrarse en concentraciones alteradas en pacientes con dismetabolismo, lo cual justifica su inclusión en análisis de VOCs asociados a este contexto.

- 5-Hepten-2-one-,6 methyl:
  Conocida como sulcatona, se produce de forma endógena en los seres humanos, particularmente en la piel. Este compuesto es un volátil característico del olor humano, identificado en el sudor y secreciones cutáneas. Un estudio clásico fraccionó químicamente el sudor humano y halló que la fracción neutra (no ácida) contenía, entre otros componentes, 6-metil-5-hepten-2-ona en cantidades significativas [85]. Además, evidencia más reciente ha confirmado la presencia de este compuesto en el aliento exhalado humano. En un estudio de Kononova et al. (2024), que analizó el perfil de compuestos orgánicos volátiles (VOCs) en muestras de aliento mediante cromatografía de gases acoplada a espectrometría de masas (GC-MS), la 6-metil-5-hepten-2-ona fue identificada como uno de los VOCs significativamente aumentados en pacientes con cáncer colorrectal frente a sujetos sanos [86].

6. Concentración semicuantitativa
   A continuación, se presentan las concentraciones semicuantitativas estimadas para los 10 compuestos orgánicos volátiles (COVs) más relevantes identificados en el grupo de alto riesgo (RA), calculadas con base en la relación de áreas entre cada metabolito y su estándar asignado. Para ello, se empleó una solución de control que contenía 250 µg/mL de tetracloruro de carbono, hexano, cloroformo y ciclohexano, de la cual se adicionaron 20 µL a cada muestra tras la recolección. El volumen final de cada muestra fue registrado individualmente, lo que permitió estandarizar las concentraciones en unidades de µg/mL.

   Para asignar el estándar correspondiente a cada COV, se mapeó el rango de masas (m/z) de los cuatro estándares disponibles, y se eligió el más cercano en masa a cada metabolito, como se muestra en las siguientes tablas:

Estándar	**m/z**
Chloroform	82.94
Cyclohexane	84.08
Carbon tetrachloride	116.89
Hexane	57.08

Cuadro 7.3: Mapeo del rango de masas de los 4 estándares

Metabolito	m/z	Estándar asignado	m/z estándar	Diferencia m/z
2-pentanol	73.05	Chloroform	82.94	9.89
3(s)-citronelol	69.04	Hexane	57.08	11.96
Ethyleneglycol monophenyl ether	94.04	Cyclohexane	84.08	9.96
2-octanone	43.04	Hexane	57.08	14.04
2-hexanone	43.04	Hexane	57.08	14.04
Butanedioic acid dimethyl ester	115.03	Carbon tetrachloride	116.89	1.86
3,5-dimethylpyridine	54.03	Hexane	57.08	3.05
2-phenylpropionaldehyde	57.03	Hexane	57.08	0.05
Octanal	142.15	Carbon tetrachloride	116.89	25.26
5-Hepten-2-one, 6-methyl	39.99	Hexane	57.08	17.09

Cuadro 7.4: Asignación de estándares para los metabolitos según la diferencia de masas

La concentración semicuantitativa (C_{semi}) se calculó mediante la siguiente expresión:

$$C_{\text{semi}} = C_{\text{estándar}} \times \left(\frac{A_{\text{metabolito}}}{A_{\text{estándar}}} \right)$$

donde $C_{\text{estándar}}$ representa la concentración del estándar en la muestra, determinada a partir del volumen añadido y el volumen final de la muestra. Esta aproximación permitió convertir las áreas relativas en concentraciones absolutas aproximadas (µg/mL), facilitando la comparación directa entre muestras y entre grupos (RA vs RB).

Las siguientes tablas presentan las concentraciones semicuantitativas individuales por muestra, agrupadas en tres series:

Metabolito	Estándar	001 RA	007 RB	015 RA	016 RA	035 RA	038 RB	039 RA	042 RA	056 RA	061 RA	064 RB	071 RA	078 RB	094 RA	099 RA
2-pentanol	Chloroform	0.0201	0.0943	0.0256	0.0873	2.3367	0.1143	0.0215	0.2613	0.2477	0.1139	1.2733	0.3262	0.1978	0.2972	0.2493
3(s)-citronelol	Hexane	1.1074	2.8456	3.0832	1.4269	164.603	18.8309	10.2834	207.4035	42.6106	19.8962	28.3942	105.4995	45.6583	21.9499	127.4503
Ethyleneglycol Monophenyl ether	Cyclohexane	0.1247	11.2791	0.1372	0.0965	30.3952	8.7375	0.7130	7.9545	7.2909	12.8800	19.8692	20.0010	16.0713	5.5270	6.2509
2-octanone	Hexane	2.9503	4.5817	0.4972	1.0417	182.3002	10.5535	5.4144	76.9672	12.4532	9.2027	17.6577	41.3063	75.2594	38.8230	45.5709
2-hexanone	Hexane	4.0262	7.1022	1.0865	1.1702	298.1521	29.1981	7.0998	89.9807	41.6399	27.5689	50.7355	74.8538	39.2667	39.7753	58.2618
Butanedioic acid dimethyl ester	Carbon tetrachloride	0.0284	0.9833	0.0538	0.0188	7.9657	3.1435	0.0905	1.4857	1.1107	0.5664	18.4676	1.9057	1.3728	1.5154	0.4893
3,5-dimethylpyridine	Hexane	1.4442	5.5801	0.8726	0.5027	177.2258	9.6900	2.2608	81.0317	18.2653	12.3166	6.9817	69.4336	27.3966	26.4498	44.1736
2-phenylpropionaldehyde	Hexane	1.8583	4.0210	0.7421	0.7978	206.9276	11.8326	3.8975	54.7474	22.7934	15.2996	28.3524	44.4794	54.9001	30.2547	49.6480
Octanal	Carbon tetrachloride	0.0133	0.2057	0.0127	0.0055	0.1205	0.8503	0.0469	0.2726	0.1462	0.2547	2.7458	0.5044	2.0217	0.6503	0.1939
5-Hepten-2-one, 6-methyl	Hexane	20.2360	2.8384	5.3455	20.1750	1689.9562	25.2672	29.6445	1004.8555	372.6196	128.0374	103.1532	616.1820	192.8867	316.0467	523.5783

Cuadro 7.5: Concentración semicuantitativa ($\mu g/mL$) de cada metabolito en las muestras analizadas (primer serie de muestras)

Metabolito	Estándar	105 RA	108 RB	112 RA	116 RB	127 RA	129 RA	146 RB	147 RA	150 RB	173 RA	174 RB	175 RA	176 RA	177 RA	179 RB
2-pentanol	Chloroform	0.0247	0.3667	0.2659	0.1344	3.3939	0.1369	0.1758	0.2038	0.0495	0.1423	0.4101	0.2814	0.4256	1.3587	0.0427
3(s)-citronelol	Hexane	4.2526	5.9387	58.9142	31.0704	179.7516	861.3912	3.0344	52.1971	5.8870	14.3713	141.9393	98.5794	24.2091	30.7297	1.2482
Ethyleneglycol Monophenyl ether	Cyclohexane	3.1134	3.1320	101.2390	17.5505	39.7182	584.3496	5.9525	8.6850	2.8254	11.3726	19.4500	55.8483	31.6701	320.4606	1.6979
2-octanone	Hexane	4.6330	2.0412	35.8839	66.9068	464.5558	168.3402	0.6978	57.4980	3.6726	4.2048	39.1173	130.5727	23.3138	45.2387	1.3616
2-hexanone	Hexane	6.1698	10.2915	121.3214	106.1036	749.5126	199.4905	0.4636	90.5910	3.9569	14.8795	85.8859	199.2428	37.6233	59.0719	0.3476
Butanedioic acid dimethyl ester	Carbon tetrachloride	0.0140	0.1955	7.7997	1.4633	8.1721	4.3429	0.0721	0.8050	0.6221	0.8470	2.6022	6.4478	9.0950	8.5205	0.3943
3,5-dimethylpyridine	Hexane	1.8817	1.7889	38.8790	38.6399	101.5431	57.8492	0.7311	27.5887	2.9846	3.6374	51.8907	51.0805	10.4141	28.2738	0.2280
2-phenylpropionaldehyde	Hexane	2.9153	2.4039	80.0920	51.4912	301.9061	98.9481	0.7353	24.8178	3.5808	7.1365	82.6876	88.7533	17.1722	30.5193	0.3507
Octanal	Carbon tetrachloride	0.0092	0.0229	5.7533	0.0689	0.4603	1.4705	0.0262	0.5333	0.0241	0.0289	0.1781	1.8866	2.2014	2.2535	0.0243
5-Hepten-2-one, 6-methyl	Hexane	100.1913	27.1009	1389.9240	206.9030	870.4817	1566.8134	9.7108	84.5653	54.5225	88.9889	456.9150	521.6057	152.0415	660.9238	4.7353

Cuadro 7.6: Concentración semicuantitativa ($\mu g/mL$) de cada metabolito en las muestras analizadas (segunda serie de muestras)

Metabolito	Estándar	184 RB	186 RB	192 RB	193 RB	195 RA	200 RB	258 RB	269 RA	283 RB	285 RB	307 RA	310 RA	312 RA	325 RA	331 RA
2-pentanol	Chloroform	0.3193	0.0747	0.6973	0.1952	6.8066	0.0787	0.1450	0.7177	0.2132	0.2267	0.7543	0.5097	0.6806	0.1081	0.1310
3(s)-citronelol	Hexane	77.1795	141.9783	11.2572	16.6213	187.2364	28.9102	62.9188	16.2868	28.8442	3.5969	44.0214	12.3507	113.6713	26.6472	146.3062
Ethyleneglycol Monophenyl ether	Cyclohexane	8.8599	6.5125	3.8378	2.1534	867.3772	7.2334	4.0334	87.0399	11.9966	52.9338	1282.7727	101.9601	6750.3796	6.6660	8.9630
2-octanone	Hexane	49.8628	23.5481	14.0731	18.2167	177.1683	53.5536	10.5904	12.2958	30.8986	8.8523	26.5673	17.7762	101.0149	31.3990	95.7546
2-hexanone	Hexane	126.2694	37.1919	43.8627	24.1211	388.1840	22.1406	32.8794	17.3365	40.1800	13.6834	48.8259	52.0180	170.8075	36.6191	137.4819
Butanedioic acid dimethyl ester	Carbon tetrachloride	0.8618	0.5614	0.5165	0.7051	68.2822	0.8864	0.2886	15.1138	0.3924	1.4079	13.1925	4.5781	15.8952	0.4823	0.5071
3,5-dimethylpyridine	Hexane	54.3453	24.0418	24.0599	6.4899	88.8034	24.2246	10.7907	4.6925	7.1371	2.0817	17.1092	4.0852	44.2202	9.3063	44.2837
2-phenylpropionaldehyde	Hexane	77.9945	24.6845	21.2219	10.6408	126.4053	32.1519	14.5670	7.9550	15.8360	3.2389	22.8313	7.3306	83.5599	76.4761	58.6582
Octanal	Carbon tetrachloride	0.0653	0.0363	0.2223	0.2768	19.6641	0.1513	0.0160	4.0452	0.0180	0.5355	3.4023	0.2373	2.7607	0.3174	0.0990
5-Hepten-2-one, 6-methyl	Hexane	686.8588	384.7530	153.6506	194.2832	54.3502	431.9192	60.6293	61.4852	268.9696	6.1728	147.4638	83.3321	816.8921	120.9614	1049.6677

Cuadro 7.7: Concentración semicuantitativa ($\mu g/mL$) de cada metabolito en las muestras analizadas (tercera serie de muestras)

En la siguiente tabla se presenta la concentración promedio para los grupos RA y RB, respectivamente.

Metabolito	Estándar asignado	Promedio RA ($\mu g/mL$)	Promedio RB ($\mu g/mL$)
2-pentanol	Chloroform	0.7381	0.2672
3(s)-citronelol	Hexane	95.4159	36.4530
Ethyleneglycol Monophenyl ether	Cyclohexane	383.4439	11.3404
2-octanone	Hexane	67.1387	23.9692
2-hexanone	Hexane	110.1034	37.4267
Butanedioic acid dimethyl ester	Carbon tetrachloride	6.6417	1.9409
3,5-dimethylpyridine	Hexane	35.8379	16.6157
2-phenylpropionaldehyde	Hexane	54.3305	24.4828
Octanal	Carbon tetrachloride	1.7535	0.4161
5-Hepten-2-one, 6-methyl	Hexane	462.8283	181.7372

Cuadro 7.8: Concentración semicuantitativa promedio ($\mu g/mL$) de cada metabolito según tipo de muestra RA y RB

En el siguiente gráfico se visualiza la comparación gráfica de las concentraciones semicuantitativas promedio (en $\mu g/mL$L) de los 10 compuestos significativos entre los grupos RA y RB. Este gráfico permite observar de forma clara las diferencias en abundancia relativa entre los dos grupos. Estas diferencias son consistentes con los hallazgos del análisis estadístico y refuerzan la posible utilidad de estos compuestos como marcadores metabólicos tempranos asociados al riesgo de desarrollar diabetes mellitus tipo 2.

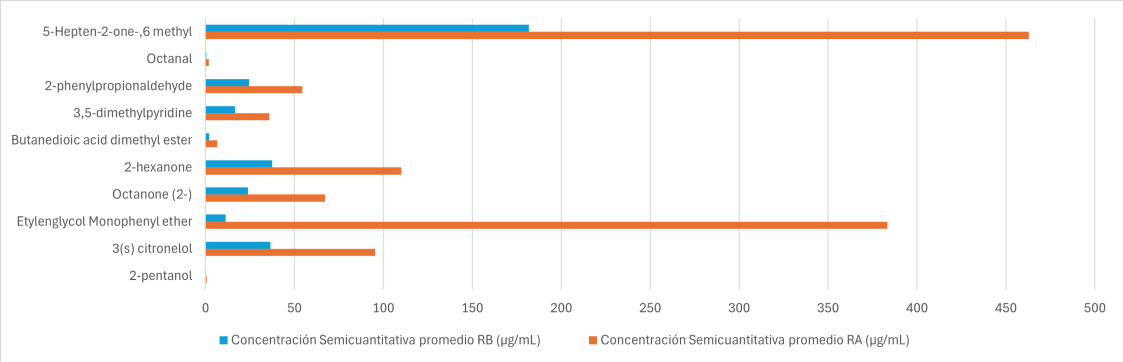

Figura 7.4: Comparación de concentraciones semicuantitativas promedio ($\mu g/mL$) de los metabolitos significativos entre los grupos RA y RB.

7. Resumen de los resultados
   En las siguientes tablas se muestra un resumen de los 10 COVs resultados y algunos datos de interés para su consideración.

ID	RT (min)	Metabolito	MWW p-Value	Levene p-Value	Log_2FC	Producción Endógena Reportada
129	14.747	2-pentanol	0.0802	0.8887	-0.5256	Posible (desde 2-pentanona vía metabolismo hepático)
670	27.484	3(s) citronelol	0.0655	0.2628	-0.3218	No
729	28.699	Ethylene glycol monophenyl ether	0.0403	0.1950	-0.5417	No
274	19.247	Octanone (2-)	0.0886	0.4487	-0.5839	Sí
185	16.951	2-hexanone	0.0802	0.7640	-0.2520	Sí
418	22.541	Butanedioic acid dimethyl ester	0.0503	0.6025	-0.4218	No
449	23.279	3,5-dimethylpyridine	0.0763	0.4192	-0.1958	No
720	28.541	2-phenylpropionaldehyde	0.0340	0.7006	-0.4383	No
581	25.861	Octanal	0.0074	0.3970	-1.1919E+16	Sí
560	25.489	5-Hepten-2-one, 6 methyl-	0.0031	0.0733	-0.9184	Sí

Cuadro 7.9: Resumen estadístico de los 10 compuestos orgánicos volátiles (COVs) significativos más relevantes diferenciadores en RA.

Capítulo 8

Conclusiones

Conclusiones

Este trabajo permitió realizar una exploración de los compuestos orgánicos volátiles (COVs) presentes en el aliento exhalado condensado (AEC), diferenciando entre individuos con alto y bajo riesgo de desarrollar diabetes mellitus tipo 2 (DM2), y aportando evidencia sobre su potencial como herramienta de tamizaje metabólico no invasiva. A través de un análisis estadístico riguroso, basado en pruebas no paramétricas, visualizaciones comparativas y validación documental, se logró identificar un conjunto de compuestos significativamente asociados con un mayor riesgo metabólico.

En total, se identificaron 26 compuestos diferenciadores en el grupo de alto riesgo (RA), de los cuales 8 fueron reconocidos como metabolitos conocidos. Adicionalmente, entre los compuestos inicialmente clasificados como desconocidos, se seleccionaron los cuatro más relevantes para su identificación estructural mediante el software *MassHunter* y la librería *NIST*. Dos de ellos fueron descartados posteriormente tras ser identificados como contaminantes mediante análisis de muestras blanco. Entre los compuestos identificados y confirmados como de posible origen endógeno se destacan el octanal, la 2-octanona, la 2-hexanona y la sulcatona, que resultan fisiopatológicamente relevantes por su implicación en procesos como la peroxidación lipídica, el estrés oxidativo, la disfunción hepática y otras alteraciones metabólicas. A este grupo se suma el 2-pentanol, cuya producción endógena ha sido reportada experimentalmente en condiciones de daño hepático, lo que sugiere su potencial como marcador indirecto de alteraciones metabólicas incipientes.

Más allá de los hallazgos analíticos, este estudio representa una contribución metodológica al campo de la metabolómica respiratoria, al demostrar que, mediante una combinación de técnicas accesibles como GC-MS, herramientas de identificación espectral y análisis estadístico adecuado, es posible generar información clínicamente relevante. La integración de procedimientos cuantitativos, controles de calidad y evaluación fisiopatológica fortalece la propuesta de los COVs como indicadores exploratorios en etapas tempranas de disfunción metabólica.

A pesar de los resultados prometedores, se reconocen limitaciones importantes, entre ellas el tamaño reducido de la muestra, la variabilidad intraindividual del volatiloma y la necesidad de validación externa en poblaciones independientes. Asimismo, es fundamental avanzar en la estandarización de protocolos de recolección y en el control de factores exógenos (como alimentación, microbiota y ambiente) que puedan influir en la composición del aliento. Otra limitación metodológica relevante fue la ausencia de una muestra basal ambiental (aire del entorno), que hubiese

permitido distinguir con mayor claridad los compuestos derivados del ambiente respecto a los propios del aliento exhalado. Aunque se implementó un filtrado de blancos, este abordaje no permite establecer un umbral absoluto de señal ambiental ni identificar compuestos que estén presentes simultáneamente en ambiente y muestra biológica.

Para abordar esta limitación, se recomienda que futuros estudios incluyan blancos de campo recolectados simultáneamente durante la toma de muestras, lo cual permitiría identificar compuestos volátiles ambientales con mayor precisión. En cuanto al tamaño muestral, se realizó una estimación retrospectiva basada en poder estadístico para detectar diferencias con un tamaño del efecto medio ($d = 0{,}5$), una significancia de $\alpha = 0{,}05$ y un poder estadístico del 80 %. Bajo estos parámetros, se requieren aproximadamente 64 participantes por grupo, es decir, un total de 128 sujetos para lograr una robustez estadística adecuada.

En conjunto, estos resultados sientan las bases para futuros desarrollos en medicina preventiva personalizada, mediante herramientas diagnósticas basadas en sensores de aliento, dispositivos portátiles y algoritmos de predicción multivariada. El enfoque propuesto no solo tiene implicaciones para el diagnóstico temprano de la DM2, sino también para el monitoreo continuo del riesgo metabólico y la intervención oportuna en poblaciones vulnerables.

Finalmente, este estudio reafirma el valor del enfoque interdisciplinario que articula ingeniería biomédica, química analítica, estadística y salud pública como vía para abordar desafíos complejos como la diabetes tipo 2 desde una perspectiva innovadora, tecnológica y socialmente pertinente.

Capítulo 9

Trabajos futuros

Trabajos Futuros

Los resultados obtenidos en esta investigación abren múltiples posibilidades para profundizar en el estudio de los compuestos orgánicos volátiles (COVs) como herramientas diagnósticas no invasivas en el contexto de la diabetes mellitus tipo 2 (DM2). A continuación, se proponen algunas líneas de trabajo futuro que podrían fortalecer, complementar o extender los hallazgos de este estudio:

1. **Ampliación del tamaño muestral y validación poblacional**: Realizar estudios con un mayor número de participantes permitiría aumentar la potencia estadística y confirmar la reproducibilidad de los compuestos identificados como diferenciadores. Asimismo, la inclusión de individuos con diagnóstico clínico confirmado de DM2 podría permitir validar los compuestos no solo como indicadores de riesgo, sino también como posibles marcadores diagnósticos o de progresión.

2. **Monitoreo longitudinal del perfil de VOCs**: Implementar un seguimiento temporal de los participantes clasificados como de alto riesgo permitiría evaluar si los compuestos identificados presentan estabilidad, evolución o anticipación respecto a la aparición clínica de la enfermedad. Esto aportaría evidencia clave para su uso en programas de tamizaje predictivo.

3. **Control ambiental riguroso y muestreo del entorno**: Incluir en futuros protocolos la recolección sistemática de muestras de aire ambiental en los lugares donde se realice la toma de aliento, con el fin de distinguir de manera más precisa entre VOCs de origen endógeno y aquellos derivados de exposición externa.

4. **Diseño de sensores selectivos o dispositivos portátiles**: A largo plazo, los compuestos más robustos identificados podrían ser la base para el diseño de biosensores o dispositivos de aliento portátiles, con aplicaciones potenciales en tamizaje comunitario, atención primaria o monitoreo domiciliario en poblaciones vulnerables.

5. **Estudios multicéntricos en diferentes contextos geográficos y étnicos**: Validar los perfiles de VOCs en diversas poblaciones permitiría evaluar la robustez y generalizabilidad de los hallazgos, así como identificar factores ambientales o genéticos que modulan su expresión.

Estas líneas de trabajo futuro permitirían avanzar hacia la consolidación del aliento exhalado como una matriz diagnóstica alternativa, y hacia el uso de la metabolómica respiratoria como herramienta estratégica en la detección temprana y no invasiva de enfermedades metabólicas crónicas como la diabetes tipo 2.

Capítulo 10

Anexos

Anexo 1 – Código para filtrado, normalización y análisis estadístico

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from scipy.stats import mannwhitneyu, levene
from sklearn.preprocessing import StandardScaler

file_path = "input/Area_Alignment.csv"
df = pd.read_csv(file_path, skiprows=3)

df.columns = df.iloc[0]
df = df[1:].reset_index(drop=True)

df = df.iloc[:, :75]

ra_cols = [col for col in df.columns if 'RA' in col]
rb_cols = [col for col in df.columns if 'RB' in col]
all_sample_cols = ra_cols + rb_cols

df[all_sample_cols] = df[all_sample_cols].apply(pd.to_numeric, errors='coerce')

print(f"\nTotal compuestos antes del filtrado: {df.shape[0]}")

compuestos_excluir = [
    "CYCLOHEPTATRIENE", "Putative: HEXANE", "HEXANE", "TRIMETHYL ORTHOFORMATE",
        "CHLOROFORM",
    "Putative: CYCLOHEXANE", "CYCLOHEXANE", "CARBON TETRACHLORIDE", "BENZENE",
    "2,4-DI-TERT-BUTYLPHENOL", "ORTHO XYLENE", "Putative: META-XYLENE", "BUTYL ACETATE",
    "Putative: OCTAMETHYLCYCLOTETRASILOXANE", "OCTAMETHYLCYCLOTETRASILOXANE",
        "HEPTYLALDEHYDE",
    "2-BUTOXYETHANOL", "ISOPROPYLBENZENE", "DOCOSANE", "4-ETHYLPYRIDINE", "HEPTANOL",
    "dihydromyrcenol", "Putative: 3-ETHYL-2,6,10-TRIMETHYLUNDECANE", "heptacosane",
    "Putative: heptacosane", "2,3,6-TRIMETHYLPYRIDINE", "OCTYLALDEHYDE", "Putative:
        2-ETHYLHEXANOL",
    "TRIDECANE", "propanoic acid, 2-methyl-, 3-hydroxy-2,4,4-trimethylpentyl ester",
    "2-ETHYLHEXANOL", "1-OCTANOL", "nonacosane", "DIETHYLENE GLYCOL DIBUTYL ETHER",
    "Putative: DODECANE", "Putative: eicosane", "4-HydroxyMandelate", "2-BUTOXYETHYL
        ACETATE",
    "4-HydroxyMandelic acid", "Putative: 4-Hydroxyphenylpyruvic acid",
        "4-Hydroxyphenylpyruvic acid",
    "undecanol (n-)", "Putative: NONANAL", "P-(1-PROPENYL)-TOLUENE", "NONANAL",
    "Putative: PELARGONALDEHYDE", "1-METHYL-2-PYRROLIDINONE",
    "3-METHYL-2-(TRIMETHYLSILYL)OXYBENZOIC ACIDTRIMETHYLSILYL ESTER",
    "tetracosane", "TETRAHYDRO-BETA-NAPHTHYLACETATE", "2-ETHYLHEXYL ACETATE", "1-NONANOL",
    "CROTONIC ACID BUTYL ESTER", "EICOSANE", "Putative: 1-NONANOL", "GERANYL ACETONE",
    "4,8-DIMETHYL-4-NONANOL", "Putative: NONANOL", "NONANOL", "Putative:
        2-methyl-nonadecane",
    "3,5,5-TRIMETHYL-2-CYCLOHEXENONE",
    "propanoic acid, 2-methyl-, 1-(1,1-dimethylethyl)-2-methyl-1,3-propanediyl ester",
    "DODECANOL", "P-ETHYLBENZALDEHYDE", "L-MENTHOL", "DECYLALDEHYDE",
    "ACRYLIC ACID 2-ETHYLHEXYL ESTER", "TETRAHYDROLAVANDULOL", "ALPHA-TERPINEOL",
    "n-Eicosane", "TRANS-3-PHENYL-2-PROPENAL", "NONADECANE", "1,3-DI-TERT-BUTYLBENZENE",
    "Putative: tetracosane", "OCTADECANE", "ISOAMYL BENZOATE", "1-DECANOL", "octadecanol
        acetate",
    "pentadecanol (n-)", "OCTYL BENZOATE", "PENTACOSANE", "1,2-BENZENEDICARBOXYLIC ACID
        DIETHYL ESTER"
```

```
]

metabolite_col = "Metabolite␣name"

df = df[~df[metabolite_col].isin(compuestos_excluir)].copy()
print(f"Compuestos␣despu s␣de␣eliminaci n␣manual:␣{df.shape[0]}")


percentiles = df[all_sample_cols].quantile(0.75)
presente = df[all_sample_cols] >= percentiles
conteo_presencia = presente.sum(axis=1)
porcentaje_umbral = 0.90
limite = int(len(all_sample_cols) * porcentaje_umbral)
df_filtrado = df[conteo_presencia < limite]
print(f"Compuestos␣totales:␣{df.shape[0]},␣luego␣del␣filtrado␣ubicuo:␣
    {df_filtrado.shape[0]}")

compuestos_eliminados = df[conteo_presencia >= limite].copy()
compuestos_eliminados.to_csv("Compuestos_eliminados_por_filtro.csv", index=False)
print(f"Compuestos␣eliminados␣por␣filtro␣de␣percentil␣y␣presencia␣ubicua:␣
    {compuestos_eliminados.shape[0]}")


df = df_filtrado.copy()
df[all_sample_cols] = df[all_sample_cols].div(df[all_sample_cols].sum(axis=0), axis=1)
df.to_csv("AreaAlignment_Normalizado.csv", index=False)

scaler = StandardScaler()
autoscaled_data = pd.DataFrame(scaler.fit_transform(df[all_sample_cols]),
    columns=all_sample_cols)

if df.columns[0] not in all_sample_cols:
    autoscaled_data.insert(0, df.columns[0], df.iloc[:, 0])

autoscaled_data.to_csv("AreaAlignment_Autoscaled.csv", index=False)

df['RA_median'] = df[ra_cols].median(axis=1)
df['RB_median'] = df[rb_cols].median(axis=1)

df_filtered = df[(df['RA_median'] > 1e-6) & (df['RB_median'] > 1e-6)].copy()
df_filtered.reset_index(drop=True, inplace=True)
df_filtered['log2FC'] = np.log2(df_filtered['RB_median'] / df_filtered['RA_median'])

p_values = []
wmw_statistic = []
levene_p_values = []

for i in df_filtered.index:
    ra_values = df_filtered.loc[i, ra_cols].dropna()
    rb_values = df_filtered.loc[i, rb_cols].dropna()

    levene_p = levene(ra_values, rb_values).pvalue if len(ra_values) > 1 and len(rb_values)
        > 1 else np.nan

    try:
        ra_values = ra_values.astype(float)
        rb_values = rb_values.astype(float)

```

```
        if len(ra_values) > 0 and len(rb_values) > 0:
            stat, p_value = mannwhitneyu(ra_values, rb_values, alternative='two-sided')
        else:
            stat, p_value = np.nan, np.nan
    except Exception as e:
        print(f"Error en fila {i}: {e}")
        stat, p_value = np.nan, np.nan

    p_values.append(p_value)
    wmw_statistic.append(stat)
    levene_p_values.append(levene_p)

df_filtered['WMW_statistic'] = wmw_statistic
df_filtered['p-value'] = p_values
df_filtered['Levene_p-value'] = levene_p_values
df_filtered['-log10(p-value)'] = -np.log10(df_filtered['p-value'].fillna(1))

significance_threshold = 0.1
df_valid = df_filtered[df_filtered['Levene_p-value'] >= 0.05].copy()
df_valid.reset_index(drop=True, inplace=True)
df_valid['Significant'] = df_valid['p-value'] < significance_threshold

print("\n Diagn stico de filtrado final:")
print("Compuestos tras Levene (>= 0.05):", df_valid.shape[0])
print("Significativos (p < 0.05):", df_valid['Significant'].sum())

significant_positive = df_valid[(df_valid['Significant']) & (df_valid['log2FC'] > 0)]
significant_negative = df_valid[(df_valid['Significant']) & (df_valid['log2FC'] < 0)]

print("   Aumentados en RB:", significant_positive.shape[0])
print("   Aumentados en RA:", significant_negative.shape[0])

print("\nResumen log2FC de significativos:")
print(df_valid[df_valid['Significant']]['log2FC'].describe())

no_sig = df_valid[~df_valid['Significant']]
sig_RA = df_valid[(df_valid['Significant']) & (df_valid['log2FC'] < 0)]
sig_RB = df_valid[(df_valid['Significant']) & (df_valid['log2FC'] > 0)]

plt.figure(figsize=(14, 10))

plt.scatter(
    no_sig['log2FC'],
    no_sig['-log10(p-value)'],
    color='gray',
    alpha=0.5,
    label='Not significant'
)

plt.scatter(
    sig_RA['log2FC'],
    sig_RA['-log10(p-value)'],
    color='red',
    edgecolor='black',
    alpha=0.7,
    label='Significant in RA'
)

```

```
plt.scatter(
    sig_RB['log2FC'],
    sig_RB['-log10(p-value)'],
    color='blue',
    edgecolor='black',
    alpha=0.7,
    label='Significant in RB'
)

for i, row in pd.concat([sig_RA, sig_RB]).iterrows():
    plt.text(
        x=row['log2FC'],
        y=row['-log10(p-value)'] + 0.1,
        s=row[df.columns[0]],
        fontsize=9,
        ha='center'
    )

plt.axhline(y=-np.log10(significance_threshold),
            linestyle='--', color='black', alpha=0.7)
plt.axvline(x=0,
            linestyle='--', color='black', alpha=0.7)

plt.xlabel('L o g  Fold Change (RA vs RB)', fontsize=14)
plt.ylabel('- L o g  (p-value)', fontsize=14)
plt.title('Volcano Plot: RA vs RB', fontsize=16)
plt.legend(fontsize=12)
plt.grid(alpha=0.3)
plt.tight_layout()
plt.show()


df_valid['Grupo_predominante'] = np.where(
    df_valid['Significant'],
    np.where(df_valid['log2FC'] < 0, 'RA', 'RB'),
    'No significativo'
)
df_valid.to_csv("Tabla_maestra_volcano.csv", index=False, sep=';')

compounds_ra_final = significant_negative.iloc[:, [0, 4]].copy()
compounds_ra_final['p-value'] = significant_negative['p-value']
compounds_ra_final['log2FC'] = significant_negative['log2FC']
compounds_ra_final['-log10(p-value)'] = significant_negative['-log10(p-value)']
compounds_ra_final['WMW_statistic'] = significant_negative['WMW_statistic']
compounds_ra_final['Levene_p-value'] = significant_negative['Levene_p-value']

compounds_rb_final = significant_positive.iloc[:, [0, 4]].copy()
compounds_rb_final['log2FC'] = significant_positive['log2FC']
compounds_rb_final['-log10(p-value)'] = significant_positive['-log10(p-value)']
compounds_rb_final['WMW_statistic'] = significant_positive['WMW_statistic']
compounds_rb_final['Levene_p-value'] = significant_positive['Levene_p-value']

compounds_ra_final.to_csv("wilcoxon_RA_aumentados.csv", index=False, sep=';')
compounds_rb_final.to_csv("wilcoxon_RB_aumentados.csv", index=False, sep=';')

print("\n Guardado: 'wilcoxon_RA_aumentados.csv'")
print("Guardado: 'wilcoxon_RB_aumentados.csv'")
```

Listing 10.1: Script en Python para el procesamiento estadístico de compuestos orgánicos volátiles (COVs)

Anexo B. Código para análsis de correlación

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

file_path_areas = "input/Area_Alignment.csv"
df_areas = pd.read_csv(file_path_areas, skiprows=3)

df_areas.columns = df_areas.iloc[0]
df_areas = df_areas[1:].reset_index(drop=True)
df_areas.rename(
    columns={
        df_areas.columns[0]: "ID",
        df_areas.columns[4]: "Metabolite name"
    },
    inplace=True
)

df_areas["ID"] = pd.to_numeric(df_areas["ID"], errors='coerce')

ra_cols = [col for col in df_areas.columns if 'RA' in str(col)]
print("Columnas RA detectadas:", ra_cols)

id_nombre_dict = {
    129: "2-pentanol",
    670: "3(s) citronelol",
    729: "Etylenglycol Monophenyl ether",
    274: "Octanone (2-)",
    185: "2-hexanone",
    418: "Butanedioic acid dimethyl ester",
    449: "3,5-dimethylpyridine",
    720: "2-phenylpropionaldehyde",
    581: "Octanal",
    560: "5-Hepten-2-one-,6 methyl"
}

df_areas_filtered = df_areas[df_areas['ID'].isin(id_nombre_dict.keys())].copy()

df_areas_filtered['Nombre_final'] = df_areas_filtered['ID'].map(id_nombre_dict)

df_RA = df_areas_filtered[['Nombre_final'] + ra_cols].copy()
df_RA.set_index("Nombre_final", inplace=True)
df_RA = df_RA.apply(pd.to_numeric, errors='coerce')

corr_RA = df_RA.T.corr(method='spearman')

g_RA = sns.clustermap(
    corr_RA,
```

```
    cmap="coolwarm",
    annot=True,
    fmt=".2f",
    linewidths=0.5,
    figsize=(12, 10)
)
g_RA.fig.suptitle("Spearman Correlation - RA", fontsize=16)
g_RA.fig.subplots_adjust(top=0.92)

g_RA.ax_heatmap.set_xlabel("")
g_RA.ax_heatmap.set_ylabel("")

plt.setp(g_RA.ax_heatmap.xaxis.get_majorticklabels(), rotation=45, ha="right")
plt.setp(g_RA.ax_heatmap.yaxis.get_majorticklabels(), rotation=0)

plt.show()
```

Listing 10.2: Script en Python para el análisis de correlación entre los COVs significativos relevantes en RA

Anexo C. Código para generación de diagramas de caja

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import matplotlib.ticker as ticker
import math

df = pd.read_csv("AreaAlignment_Normalizado.csv", sep=",")
df.rename(columns={df.columns[0]: "ID"}, inplace=True)

id_nombre_dict = {
    129: "2-pentanol",
    670: "3(s) citronelol",
    729: "Etylenglycol Monophenyl ether",
    274: "Octanone (2-)",
    185: "2-hexanone",
    418: "Butanedioic acid dimethyl ester",
    449: "3,5-dimethylpyridine",
    720: "2-phenylpropionaldehyde",
    581: "Octanal",
    560: "5-Hepten-2-one-,6 methyl"
}

df["ID"] = pd.to_numeric(df["ID"], errors="coerce")
df = df[df["ID"].isin(id_nombre_dict.keys())].copy()

df["Metabolite name"] = df["ID"].map(id_nombre_dict)
df["Compound"] = df["ID"].astype(int).astype(str) + "     " + df["Metabolite name"]

ra_cols = [c for c in df.columns if " RA " in c]
rb_cols = [c for c in df.columns if " RB " in c]

df[ra_cols + rb_cols] = df[ra_cols + rb_cols].apply(pd.to_numeric, errors="coerce")
```

```
34
35 df_long = df.melt(
36     id_vars=["Compound"],
37     value_vars=ra_cols + rb_cols,
38     var_name="Sample",
39     value_name="Area"
40 ).dropna(subset=["Area"])
41 df_long["Group"] = np.where(df_long["Sample"].str.contains("␣RA␣"), "RA", "RB")
42
43 compounds = df_long["Compound"].unique()
44 n = len(compounds)
45 cols = 3
46 rows = math.ceil(n / cols)
47
48 fig, axes = plt.subplots(rows, cols, figsize=(cols * 5, rows * 4))
49 axes = axes.flatten()
50
51 for i, comp in enumerate(compounds):
52     sub = df_long[df_long["Compound"] == comp]
53     ax = axes[i]
54
55     sns.boxplot(
56         data=sub,
57         x="Group",
58         y="Area",
59         palette={"RA": "skyblue", "RB": "lightcoral"},
60         width=0.6,
61         ax=ax
62     )
63
64     if (sub["Area"] > 0).all():
65         ax.set_yscale("log")
66         ax.yaxis.set_major_formatter(ticker.ScalarFormatter())
67         ax.yaxis.set_minor_formatter(ticker.NullFormatter())
68
69     ax.set_title(comp, fontsize=10)
70     ax.set_xlabel("")
71     ax.set_ylabel("Normalized␣Area", fontsize=9)
72     ax.tick_params(axis='x', labelsize=8)
73     ax.tick_params(axis='y', labelsize=8)
74     ax.grid(True, linestyle="--", alpha=0.4)
75
76 for j in range(i + 1, len(axes)):
77     fig.delaxes(axes[j])
78
79 plt.tight_layout()
80 plt.show()
```

Listing 10.3: Script en Python para la visualización de Boxplots de los compuestos significativos relevantes en RA

Bibliografía

[1] Frontiers, "Frontiers | Volatilomes of Bacterial Infections in Humans," https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2020.00257/full, 2020, accedido el 19 de noviembre de 2024. [En línea].

[2] A. H. Jalal, F. Alam, S. Roychoudhury, Y. Umasankar, N. Pala, and S. Bhansali, "Prospects and Challenges of Volatile Organic Compound Sensors in Human Healthcare," *Sensors*, vol. 18, no. 6, 2018, review, June 8, 2018.

[3] M. Woollam, A. P. Siegel, P. Grocki, J. L. Saunders, D. B. Sanders, M. Agarwal, and M. D. Davis, "Preliminary method for profiling volatile organic compounds in breath that correlate with pulmonary function and other clinical traits of subjects diagnosed with cystic fibrosis: a pilot study," *Journal of Breath Research*, vol. 16, no. 2, 2022. [Online]. Available: https://doi.org/10.1088/1752-7163/ac4536

[4] M. Mohd Ali, N. Hashim, S. Abd Aziz, and O. Lasekan, "Principles and recent advances in electronic nose for quality inspection of agricultural and food products," *Trends in Food Science & Technology*, vol. 99, pp. 1–10, 2020. [Online]. Available: https://doi.org/10.1016/j.tifs.2020.02.028

[5] (2024) Datos y cifras. Federación Internacional de Diabetes. Accessed: November 19, 2024. [Online]. Available: https://idf.org/es/about-diabetes/diabetes-facts-figures/

[6] R. Ruze, T. Liu, X. Zou, J. Song, Y. Chen, R. Xu, X. Yin, and Q. Xu, "Obesity and type 2 diabetes mellitus: connections in epidemiology, pathogenesis, and treatments," *Frontiers in Endocrinology*, vol. 14, p. 1161521, 2023.

[7] U. Galicia-Garcia, A. Benito-Vicente, S. Jebari, A. Larrea-Sebal, H. Siddiqi, K. B. Uribe, H. Ostolaza, and C. Martín, "Pathophysiology of type 2 diabetes mellitus," *International Journal of Molecular Sciences*, vol. 21, no. 17, p. 6275, 2020.

[8] I. Ali, M. Suhail, Z. A. Alothman, A. Abdulrahman, and H. Y. Aboul-Enein, "Chapter 2 - drug analyses in human plasma by chromatography," in *Methods of Therapeutic Drug Monitoring Including Pharmacogenetics*, ser. Handbook of Analytical Separations, G. Hempel, Ed. Elsevier Science B.V., 2020, vol. 7, pp. 15–46. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780444640666000022

[9] S. N. Thomas, "Chapter 10 - mass spectrometry," in *Contemporary Practice in Clinical Chemistry (Fourth Edition)*, fourth edition ed., W. Clarke and M. A. Marzinke, Eds. Academic Press, 2019, pp. 171–185. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128154991000107

[10] C. Deng, J. Zhang, X. Yu, W. Zhang, and X. Zhang, "Determination of acetone in human breath by gas chromatography–mass spectrometry and solid-phase microextraction with on-fiber derivatization," *Journal of Chromatography B*, vol. 810, no. 2, pp. 269–275, 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1570023204006579

[11] V. Jalili, A. Barkhordari, and A. Ghiasvand, "A comprehensive look at solid-phase microextraction technique: A review of reviews," *Microchemical Journal*, vol. 152, p. 104319, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0026265X19319186

[12] U. EPA, "Technical Overview of Volatile Organic Compounds | US EPA," https://www.epa.gov/indoor-air-quality-iaq/technical-overview-volatile-organic-compounds#definition, 2024, accedido el 19 de noviembre de 2024. [En línea].

[13] Mayo Clinic, "Diabetes de tipo 2 - Síntomas y causas - Mayo Clinic," https://www.mayoclinic.org/es/diseases-conditions/type-2-diabetes/symptoms-causes/syc-20351193#:~:text=La%20diabetes%20tipo%202%20es,de%20azÃžcar%20en%20la%20sangre, 2024, top-ranked Hospital in the Nation - Mayo Clinic.

[14] MedlinePlus, "Prediabetes | Hiperglicemia | MedlinePlus en español," https://medlineplus.gov/spanish/prediabetes.html#:~:text=Â¿QuÃľ%20causa%20la%20prediabetes?,las%20cÃľlulas%20para%20darles%20energÃŋa, 2024, medlinePlus - Health Information from the National Library of Medicine.

[15] Kaiser Permanente, "Kaiser Permanente: Atención y cobertura personalizadas justo a su medida," https://espanol.kaiserpermanente.org/es/health-wellness/health-encyclopedia/he.diabetes-diferencias-entre-tipo-1-y-tipo-2.uq1217abc#:~:text=Por%20lo%20general,%20las%20personas,se%20llama%20deficiencia%20de%20insulina, 2024.

[16] Manual MSD, "Complicaciones de la diabetes mellitus - Manual MSD versión para público general," https://www.msdmanuals.com/es/hogar/trastornos-hormonales-y-metabÃşlicos/diabetes-mellitus-y-otros-trastornos-del-metabolismo-de-la-glucosa-sanguÃŋnea/complicaciones-de-la-diabetes-mellitus, 2024, manual MSD versión para público general.

[17] MedlinePlus, "Prueba de glucosa en la sangre: Prueba de laboratorio de MedlinePlus," https://medlineplus.gov/spanish/pruebas-de-laboratorio/prueba-de-glucosa-en-la-sangre/#:~:text=Los%20riesgos%20de%20un%, 2024, medlinePlus - Health Information from the National Library of Medicine.

[18] X. Duan, Z. Chen, C. Xia, R. Zhong, L. Liu, and L. Long, "Increased levels of urine volatile organic compounds are associated with diabetes risk and impaired glucose homeostasis," *The Journal of Clinical Endocrinology & Metabolism*, vol. 109, no. 2, pp. e531–e542, February 2024.

[19] T. Issitt, L. Wiggins, M. Veysey, S. T. Sweeney, W. J. Brackenbury, and K. Redeker, "Volatile compounds in human breath: critical review and meta-analysis," *Journal*

of Breath Research, vol. 16, no. 2, p. 024001, February 2022. [Online]. Available: https://doi.org/10.1088/1752-7163/ac5230

[20] *Breathborne Biomarkers and the Human Volatilome*. Elsevier, 2020. [Online]. Available: https://doi.org/10.1016/c2018-0-04980-4

[21] F. Buljubasic and G. Buchbauer, "The scent of human diseases: a review on specific volatile organic compounds as diagnostic biomarkers," *Flavour & Fragrance Journal*, vol. 30, no. 1, pp. 5–25, 2015. [Online]. Available: https://search.ebscohost.com/login.aspx?direct=true&AuthType=sso&db=asn&AN=100066402&lang=es&site=ehost-live&scope=site

[22] Y. Yan, Q. Wang, W. Li, Z. Zhao, X. Yuan, Y. Huang, and Y. Duan, "Discovery of potential biomarkers in exhaled breath for diagnosis of type 2 diabetes mellitus based on gc-ms with metabolomics," *RSC Adv.*, vol. 4, pp. 25 430–25 439, 2014. [Online]. Available: http://dx.doi.org/10.1039/C4RA01422G

[23] (2024) Diabetes de tipo 2 - diagnóstico y tratamiento - mayo clinic. Top-ranked Hospital in the Nation - Mayo Clinic. [Online]. Available: https://www.mayoclinic.org/es/diseases-conditions/type-2-diabetes/diagnosis-treatment/drc-20351199

[24] Y. Qiang, X. Lu, and Y. Zhang, "Association between dietary patterns and glycemic control in type II diabetes mellitus patients," *Department of Endocrinology, Qingdao Hospital, University of Health and Rehabilitation Sciences (Qingdao Municipal Hospital)*, Jun. 2024, received 12 March 2024; accepted 27 June 2024.

[25] J. H. Medina-Chávez, M. Vázquez-Parrodi, P. Mendoza-Martínez, E. D. Ríos-Mejía, J. C. de Anda-Garay, and D. A. Balandrán-Duarte, "Protocolo de Atención Integral: prevención, diagnóstico y tratamiento de diabetes mellitus 2," *Revista Médica del Instituto Mexicano del Seguro Social*, vol. 60, no. Supl 1, pp. S4–S18, 2022.

[26] M. J. Hossain, M. Al-Mamun, and M. R. Islam, "Diabetes mellitus, the fastest growing global public health concern: Early detection should be focused," *Health Science Reports*, vol. 7, no. 3, p. e2004, March 22 2024.

[27] L. X. Acosta Ruiza, M. Angarita Merchán, and L. Orjuela Vargas, "Diabetes mellitus tipo 2: Latinoamérica y colombia, análisis del último quinquenio," *Revista Med*, vol. 31, no. 2, pp. 35–46, 2024.

[28] (2024) Prueba de hemoglobina a1c: Prueba de laboratorio de medlineplus. MedlinePlus - Health Information from the National Library of Medicine. [Online]. Available: https://medlineplus.gov/spanish/pruebas-de-laboratorio/prueba-de-hemoglobina-a1c/#:~:text=Una%20prueba%20de%20hemoglobina%20A1c%20(HbA1c)%20es%20una%20prueba%20de,la%20glucosa%20para%20obtener%20energÃŋa

[29] K. Ogurtsova, L. Guariguata, N. C. Barengo, P. L.-D. Ruiz, J. W. Sacre, S. Karuranga, H. Sun, E. J. Boyko, and D. J. Magliano, "Idf diabetes atlas: Global estimates of undiagnosed diabetes in adults for 2021," *Diabetes Research and Clinical Practice*, vol. 183, p. 109118, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168822721004770

[30] K. Dixit, S. Fardindoost, A. Ravishankara, N. Tasnim, and M. Hoorfar, "Exhaled breath analysis for diabetes diagnosis and monitoring: Relevance, challenges and possibilities," *Biosensors (Basel)*, vol. 11, no. 12, p. 476, November 25 2021.

[31] P. Catalão Moura, M. Raposo, and V. Vassilenko, "Breath volatile organic compounds (VOCs) as biomarkers for the diagnosis of pathological conditions: A review," *Biomedical Journal*, vol. 46, p. 100623, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2319417023000604

[32] J. Hernández Torres *et al.*, "Nanoestructuras en el diagnóstico no invasivo de la detección de compuestos orgánicos volátiles," *Elementos*, vol. 126, pp. 71–74, 2022.

[33] M. A. Araya Quijada, "Evaluación de compuestos orgánicos volátiles antropogénicos en especies arbóreas urbanas en Santiago," 2024, proyecto de grado, Universidad de Chile. Universidad de Chile, Facultad de Ciencias Forestales y de la Conservación de la Naturaleza, Programa Interfacultades Magíster en Gestión y Planificación Ambiental.

[34] G. Mahecha, G. Conto Echeverria, J. Barrera, and G. Pino, "Estudio cinético e identificación de productos de la reacción del radical OH con nerol: Implicancias atmosféricas," *INFIQC (CONICET-UNC). Dpto. de Fisicoquímica, Fac. de Ciencias Químicas. Centro Láser de Ciencias Moleculares. Universidad Nacional de Córdoba*, 2021, ciudad Universitaria Pabellón Argentina, X5000HUA Córdoba, Argentina. Correspondencia: gpino@unc.edu.ar.

[35] J. Villalobos-Morales, J. Mora-Barrantes, J. Sibaja-Brenes, G. Zárate-Montero, R. Alfaro-Solís, H. Borbón-Alpízar, and A. Retana-Díaz, "Evaluación de la exposición laboral a compuestos orgánicos volátiles (COV). Caso de estudio en una empresa química," *Tecnología en Marcha*, vol. 37, no. 3, pp. 141–155, Julio-Septiembre 2024, laboratorio de Química de la Atmósfera, Escuela de Química, Universidad Nacional (LATAQ-UNA). Costa Rica. Correspondencia: jose.sibaja.brenes@una.ac.cr, henry.borbon.alpizar@una.ac.cr.

[36] N. Drabińska, C. Flynn, N. Ratcliffe, I. Belluomo, A. Myridakis, O. Gould, M. Fois, A. Smart, T. Devine, and B. De Lacy Costello, "A literature survey of all volatiles from healthy human breath and bodily fluids: the human volatilome," *Journal of Breath Research*, vol. 15, no. 3, 2021, published 21 April 2021.

[37] A. Korpi, J. Järnberg, and A. L. Pasanen, "Microbial Volatile Organic Compounds," *Critical Reviews in Toxicology*, vol. 39, no. 2, pp. 139–193, 2009.

[38] A. Amann, B. de Lacy Costello, W. Miekisch, J. Schubert, B. Buszewski, J. Pleil, N. Ratcliffe, and T. Risby, "The human volatilome: volatile organic compounds (VOCs) in exhaled

breath, skin emanations, urine, feces and saliva," *Journal of Breath Research*, vol. 8, no. 3, 2014, published 19 June 2014.

[39] J. D. Pleil and A. Williams, "Centralized resource for chemicals from the human volatilome in an interactive open-sourced database," *Journal of Breath Research*, vol. 13, no. 4, 2019, published 15 August 2019. Not subject to copyright in the USA. Contribution of US Environmental Protection Agency.

[40] S. Sethia, R. Nanda, and T. Chakraborty, "Clinical Application of Volatile Organic Compound Analysis for Detecting Infectious Diseases," *Clinical Microbiology Reviews*, vol. 26, no. 3, 2013, july 2013, Volume 26, Issue 3.

[41] J. Dummer, M. Storer, M. Swanney, M. McEwan, A. Scott-Thomas, S. Bhandari, S. Chambers, R. Dweik, and M. Epton, "Analysis of biogenic volatile organic compounds in human health and disease," *TrAC Trends in Analytical Chemistry*, vol. 30, no. 7, pp. 960–967, 2011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165993611001270

[42] A. P. Siegel, A. Daneshkhah, D. S. Hardin, S. Shrestha, K. Varahramyan, and M. Agarwal, "Analyzing breath samples of hypoglycemic events in type 1 diabetes patients: towards developing an alternative to diabetes alert dogs," *Journal of Breath Research*, vol. 11, no. 1, 2017. [Online]. Available: https://doi.org/10.1088/1752-7163/aa6f5d

[43] M. A. G. Wallace and J. D. Pleil, "Evolution of clinical and environmental health applications of exhaled breath research: Review of methods and instrumentation for gas-phase, condensate, and aerosols," *Analytica Chimica Acta*, vol. 1024, pp. 18–38, 2018. [Online]. Available: https://doi.org/10.1016/j.aca.2018.01.069

[44] S. Dragonieri, G. Scioscia, V. N. Quaranta, P. Carratu, M. P. Venuti, M. Falcone, G. E. Carpagnano, M. P. Foschino Barbaro, O. Resta, and D. Lacedonia, "Exhaled volatile organic compounds analysis by e-nose can detect idiopathic pulmonary fibrosis," *Journal of Breath Research*, vol. 14, no. 4, 2020. [Online]. Available: https://doi.org/10.1088/1752-7163/ab98f6

[45] J. Xu, K. Liu, and C. Zhang, "Electronic nose for volatile organic compounds analysis in rice aging," *Trends in Food Science & Technology*, vol. 109, pp. 83–93, 2021. [Online]. Available: https://doi.org/10.1016/j.tifs.2021.01.027

[46] M. A. B. Khan, M. J. Hashim, J. K. King, R. D. Govender, H. Mustafa, and J. A. Kaabi, "Epidemiology of type 2 diabetes – global burden of disease and forecasted trends," *Journal of Epidemiology and Global Health*, vol. 10, pp. 107–111, 2019. [Online]. Available: https://doi.org/10.2991/jegh.k.191028.001

[47] J. A. Sanchez Delgado and N. E. Sanchez Lara, "Epidemiología de la diabetes mellitus tipo 2 y sus complicaciones," *Revista Finlay*, vol. 12, no. 2, pp. 168–176, June 30 2022, accessed: November 19, 2024. [Online]. Available: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2221-24342022000200168&lng=es&nrm=iso

[48] N. G. Forouhi and N. J. Wareham, "Epidemiology of diabetes," *Medicine*, vol. 47, no. 1, pp. 22–27, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1357303918302640

[49] A. S. Oyerinde, V. Selvaraju, J. R. Babu, and T. Geetha, "Potential role of oxidative stress in the production of volatile organic compounds in obesity," *Antioxidants*, vol. 12, no. 1, p. 129, 2023.

[50] T. Xu, J. Wang, J. Tan, T. Huang, G. Han, Y. Li, H. Yu, J. Zhou, and M. Xu, "Gas chromatography-mass spectrometry pilot study to identify volatile organic compound biomarkers of childhood obesity with dyslipidemia in exhaled breath," *Journal of Translational Internal Medicine*, vol. 11, no. 1, pp. 81–89, 2023.

[51] E. Gaude, M. K. Nakhleh, S. Patassini, J. Boschmans, M. Allsworth, B. Boyle, and M. P. van der Schee, "Targeted breath analysis: exogenous volatile organic compounds (evoc) as metabolic pathway-specific probes," *Journal of Breath Research*, vol. 13, no. 3, p. 032001, may 2019. [Online]. Available: https://dx.doi.org/10.1088/1752-7163/ab1789

[52] M. Laakso, "Biomarkers for type 2 diabetes," *Molecular Metabolism*, vol. 27, pp. S139–S146, 2019, biomarkers of Beta-Cell Health and Dysfunction: Towards Personalised Diabetes Care. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212877819305770

[53] T. S. Ahluwalia, T. O. Kilpeläinen, S. Singh, and P. Rossing, "Editorial: Novel biomarkers for type 2 diabetes," *Frontiers in Endocrinology*, vol. 10, p. 649, 2019.

[54] N. Teonata, V. Wijaya, V. Vitaloka, M. Attamimi, and M. Kartikawati, "An introduction to different types of gas chromatography," *Jurnal Sains dan Terapan Kimia*, vol. 15, p. 8, 01 2021.

[55] R. A. Hites, "Gas chromatography mass spectrometry," in *Handbook of Instrumental Techniques for Analytical Chemistry*, F. A. Settle, Ed. Upper Saddle River, NJ: Prentice Hall PTR, 1997, ch. 31, indiana University, School of Public and Environmental Affairs and Department of Chemistry.

[56] M. Zoccali, P. Q. Tranchida, and L. Mondello, "Fast gas chromatography-mass spectrometry: A review of the last decade," *TrAC Trends in Analytical Chemistry*, vol. 118, pp. 444–452, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165993619302456

[57] A. Roszkowska, N. Miękus, and T. Bączek, "Application of solid-phase microextraction in current biomedical research," *Journal of Separation Science*, October 5 2018.

[58] W. Wang, W. Zhou, S. Wang *et al.*, "Accuracy of breath test for diabetes mellitus diagnosis: a systematic review and meta-analysis," *BMJ Open Diabetes Research and Care*, vol. 9, p. e002174, 2021.

[59] I. Horváth, J. Hunt, P. J. Barnes, K. Alving, A. Antczak, E. Baraldi, G. Becher, W. J. van Beurden, M. Corradi, R. Dekhuijzen, R. A. Dweik, T. Dwyer, R. Effros, S. Erzurum, B. Gaston, C. Gessner, A. Greening, L. P. Ho, J. Hohlfeld, Q. Jöbsis, D. Laskowski, S. Loukides, D. Marlin, P. Montuschi, A. C. Olin, A. E. Redington, P. Reinhold, E. L. van Rensen, I. Rubinstein, P. Silkoff, K. Toren, G. Vass, C. Vogelberg, and H. Wirtz, "Exhaled breath condensate: methodological recommendations and unresolved questions," *European Respiratory Journal*, vol. 26, no. 3, pp. 523–548, Sep. 2005.

[60] M. J. Davies, V. R. Aroda, B. S. Collins, R. A. Gabbay, J. Green, N. M. Maruthur, S. E. Rosas, S. Del Prato, C. Mathieu, G. Mingrone, P. Rossing, T. Tankova, A. Tsapas, and J. B. Buse, "Management of Hyperglycemia in Type 2 Diabetes, 2022. A Consensus Report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD)," *Diabetes Care*, vol. 45, no. 11, pp. 2753–2786, Nov. 2022.

[61] Y.-Y. Zhou, J.-P. Chen, L. Gan, W. Xu, Y. Liu, Y.-G. Zhao, and Y. Zhu, "A non-invasive method for the detection of glucose in human exhaled breath by condensation collection coupled with ion chromatography," *Journal of Chromatography A*, vol. 1685, p. 463564, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0021967322007555

[62] H.-H. Chiu and C.-H. Kuo, "Gas chromatography-mass spectrometry-based analytical strategies for fatty acid analysis in biological samples," *Journal of Food and Drug Analysis*, vol. 28, pp. 60–73, 2020.

[63] J. Lindström and J. Tuomilehto, "The diabetes risk score: A practical tool to predict type 2 diabetes risk," *Diabetes Care*, vol. 26, no. 3, pp. 725–731, 2003. [Online]. Available: https://doi.org/10.2337/diacare.26.3.725

[64] M. Bajo-Fernández, A. Souza-Silva, C. Barbas, M. F. Rey-Stolle, and A. García, "Gc-ms-based metabolomics of volatile organic compounds in exhaled breath: applications in health and disease. a review," *Frontiers in Molecular Biosciences*, vol. Volume 10 - 2023, 2024. [Online]. Available: https://www.frontiersin.org/journals/molecular-biosciences/articles/10.3389/fmolb.2023.1295955

[65] M. B. Greiter, L. Keck, T. Siegmund, C. Hoeschen, U. Oeh, and H. G. Paretzke, "Differences in exhaled gas profiles between patients with type 2 diabetes and healthy controls," *Diabetes Technology & Therapeutics*, vol. 12, no. 6, pp. 455–463, Jun 2010.

[66] A. K. Pathak, K. Swargiary, N. Kongsawang, P. Jitpratak, N. Ajchareeyasoontorn, J. Udomkittivorakul, and C. Viphavakit, "Recent advances in sensing materials targeting clinical volatile organic compound (voc) biomarkers: A review," *Biosensors*, vol. 13, no. 1, 2023. [Online]. Available: https://www.mdpi.com/2079-6374/13/1/114

[67] National Center for Biotechnology Information, "PubChem Compound Summary for CID 22386, 2-Pentanol," https://pubchem.ncbi.nlm.nih.gov/compound/2-Pentanol, 2025, retrieved May 20, 2025.

[68] G. Ferrandino, F. Ricciardi, A. Murgia, M. Manhota, D. Fonseca, L. Nicholson-Scott, C. Fitzpatrick, O. Gandelman, M. Allsworth, B. Boyle, A. Smolinska, H. Davies, A. G. Frings, J. Contreras, V. Gabrielli, C. Asenjo-Lobos, F. Bascur, V. Barrientos, N. Clavo, M. Jerez, and L. Méndez, "S2053 targeted breath biopsy profiling of induced biomarkers unveils a metabolic adaptation in cirrhosis toward alcohol production," *The American Journal of Gastroenterology*, vol. 119, no. 10S, pp. S1467–S1468, October 2024.

[69] R. Fernández del Río, M. O'Hara, A. Holt, P. Pemberton, T. Shah, T. Whitehouse, and C. Mayhew, "Volatile biomarkers in breath associated with liver cirrhosis — comparisons of pre- and post-liver transplant breath samples," *eBioMedicine*, vol. 2, no. 9, pp. 1243–1250, 2015. [Online]. Available: https://doi.org/10.1016/j.ebiom.2015.07.027

[70] F. Bakkali, S. Averbeck, D. Averbeck, and M. Idaomar, "Biological effects of essential oils – a review," *Food and Chemical Toxicology*, vol. 46, no. 2, pp. 446–475, 2008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0278691507004541

[71] Human Metabolome Database, "2-phenoxyethanol (hmdb0041607)," https://www.hmdb.ca/metabolites/HMDB0041607, 2025, consultado el 20 de mayo de 2025.

[72] T. Li, L. Huo, C. Pulley, and A. Liu, "Decarboxylation mechanisms in biological system," *Bioorganic Chemistry*, vol. 43, pp. 2–14, 2012, mechanistic Enzymology - Part II. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045206812000107

[73] S. Xu, Q. Zhang, G. Dong, Z. Liu, J. Fu, and S. Shi, "Advances in the microbial biosynthesis of methyl ketones," *Renewable and Sustainable Energy Reviews*, vol. 208, p. 115038, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032124007640

[74] A. C. Bishop, M. Libardoni, A. Choudary, B. Misra, K. Lange, J. Bernal, M. Nijland, C. Li, M. Olivier, and P. W. Nathanielsz, "Nonhuman primate breath volatile organic compounds associate with developmental programming and cardio-metabolic status," *Journal of Breath Research*, vol. 12, no. 3, p. 036016, 2018. [Online]. Available: https://doi.org/10.1088/1752-7163/aaba84

[75] A. Pysanenko, T. Wang, P. Španěl, and D. Smith, "Acetone, butanone, pentanone, hexanone and heptanone in the headspace of aqueous solution and urine studied by selected ion flow tube mass spectrometry," *Rapid Communications in Mass Spectrometry*, vol. 23, no. 5, pp. 676–682, 2009. [Online]. Available: https://doi.org/10.1002/rcm.3963

[76] P. Mochalski, J. King, M. Klieber, K. Unterkofler, H. Hinterhuber, M. Baumann, and A. Amann, "Blood and breath levels of selected volatile organic compounds in healthy volunteers," *Analyst*, vol. 138, no. 7, pp. 2134–2145, 2013, epub 2013 Feb 25. [Online]. Available: https://doi.org/10.1039/c3an36756h

[77] S. W. Harshman, B. A. Geier, M. Fan, S. Rinehardt, B. S. Watts, L. A. Drummond, G. Preti, J. B. Phillips, D. K. Ott, and C. C. Grigsby, "The identification of hypoxia biomarkers from

exhaled breath under normobaric conditions," *Journal of Breath Research*, vol. 9, no. 4, p. 047103, 2015. [Online]. Available: https://doi.org/10.1088/1752-7155/9/4/047103

[78] R. Cozzolino, B. De Giulio, A. Martignetti, M. Di Stasio, G. Mazzone, E. Bugianesi, L. Perullo, C. Astarita, G. D'Argenio, V. Fogliano, and A. Chambery, "Urinary volatile organic compounds in non-alcoholic fatty liver disease (nafld), type two diabetes mellitus (t2dm) and nafld-t2dm coexistence," *Metabolomics*, vol. 18, no. 10, p. 98, 2022. [Online]. Available: https://doi.org/10.1007/s11306-022-01960-1

[79] Human Metabolome Database, "Dimethyl succinate (hmdb0033837)," https://www.hmdb.ca/metabolites/HMDB0033837, 2023, accedido el 20 de mayo de 2025.

[80] National Center for Biotechnology Information, "PubChem Compound Summary for CID 11565, 3,5-Dimethylpyridine," https://pubchem.ncbi.nlm.nih.gov/compound/3_5-Dimethylpyridine, 2025, retrieved May 20, 2025.

[81] Human Metabolome Database, "2-phenylpropanal (hmdb0031626)," https://hmdb.ca/metabolites/HMDB0031626, 2023, accedido el 20 de mayo de 2025.

[82] B. de Lacy Costello, A. Amann, H. Al-Kateb, C. Flynn, W. Filipiak, T. Khalid, D. Osborne, and N. M. Ratcliffe, "A review of the volatiles from the healthy human body," *Journal of Breath Research*, vol. 8, no. 1, p. 014001, 2014, epub 2014 Jan 13. [Online]. Available: https://doi.org/10.1088/1752-7155/8/1/014001

[83] P. Fuchs, C. Loeseken, J. K. Schubert, and W. Miekisch, "Breath gas aldehydes as biomarkers of lung cancer," *International Journal of Cancer*, March 2010. [Online]. Available: https://doi.org/10.1002/ijc.24970

[84] S. R. Sutaria, S. S. Gori, J. D. Morris, Z. Xie, X. A. Fu, and M. H. Nantz, "Lipid peroxidation produces a diverse mixture of saturated and unsaturated aldehydes in exhaled breath that can serve as biomarkers of lung cancer—a review," *Metabolites*, vol. 12, no. 6, p. 561, 2022. [Online]. Available: https://doi.org/10.3390/metabo12060561

[85] A. Cork and K. C. Park, "Identification of electrophysiologically-active compounds for the malaria mosquito, anopheles gambiae, in human sweat extracts," *Medical and Veterinary Entomology*, vol. 10, no. 3, pp. 269–276, July 1996. [Online]. Available: https://doi.org/10.1111/j.1365-2915.1996.tb00742.x

[86] E. Kononova, L. Mežmale, I. Poļaka, V. Veliks, L. Anarkulova, I. Vilkoite, I. Tolmanis, A. M. Ļeščinska, I. Stonāns, A. Pčolkins *et al.*, "Breath fingerprint of colorectal cancer patients based on the gas chromatography–mass spectrometry analysis," *International Journal of Molecular Sciences*, vol. 25, no. 3, p. 1632, 2024. [Online]. Available: https://doi.org/10.3390/ijms25031632