



Universidad Autónoma de Querétaro  
Facultad de Química

Identificación de perfiles de metilación global de ADN en espectros de infrarrojo  
de muestras séricas de pacientes con cáncer de pulmón

Tesis

Que como parte de los requisitos para  
obtener el Grado de  
Maestra en Química Clínica Diagnóstica

Presenta

Lic. en Biol. María Arellano Sosa

Dirigido por

Dra. Karla Isabel Lira de León

Querétaro, Qro. a 11 de agosto de 2025

La presente obra está bajo la licencia:  
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>



CC BY-NC-ND 4.0 DEED

Atribución-NoComercial-SinDerivadas 4.0 Internacional

### Usted es libre de:

**Compartir** — copiar y redistribuir el material en cualquier medio o formato

La licenciante no puede revocar estas libertades en tanto usted siga los términos de la licencia

### Bajo los siguientes términos:



**Atribución** — Usted debe dar [crédito de manera adecuada](#), brindar un enlace a la licencia, e [indicar si se han realizado cambios](#). Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que usted o su uso tienen el apoyo de la licenciante.



**NoComercial** — Usted no puede hacer uso del material con [propósitos comerciales](#).



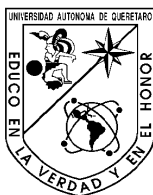
**SinDerivadas** — Si [remezcla, transforma o crea a partir](#) del material, no podrá distribuir el material modificado.

**No hay restricciones adicionales** — No puede aplicar términos legales ni [medidas tecnológicas](#) que restrinjan legalmente a otras a hacer cualquier uso permitido por la licencia.

### Avisos:

No tiene que cumplir con la licencia para elementos del material en el dominio público o cuando su uso esté permitido por una [excepción o limitación](#) aplicable.

No se dan garantías. La licencia podría no darle todos los permisos que necesita para el uso que tenga previsto. Por ejemplo, otros derechos como [publicidad, privacidad, o derechos morales](#) pueden limitar la forma en que utilice el material.



Universidad Autónoma de Querétaro  
Facultad de Química  
Maestría en Química Clínica Diagnóstica

“Identificación de perfiles de metilación global de ADN en espectros de infrarrojo de muestras séricas de pacientes con cáncer de pulmón”

**Tesis**

Que como parte de los requisitos para obtener el grado de Maestra en  
Química Clínica Diagnóstica

**Presenta**

Lic. en Biol. María Arellano Sosa

**Dirigido por**

Dra. Karla Isabel Lira de León

Dra. Karla Isabel Lira de León

**Presidente**

Dr. Raúl Salas Coronado

**Secretario**

M.I.M. David Gustavo García Gutiérrez

**Vocal**

Dr. Gerardo Manuel Nava Morales

**Suplente**

Dr. Juan Campos Guillén

**Suplente**

Centro Universitario, Querétaro, Qro.  
Agosto 2025  
México

## Agradecimientos

A **David**, por ser mi principal apoyo desde el primer día de la maestría, por escuchar con paciencia en cada momento de frustración y de duda. Gracias por aceptar todas mis facetas a lo largo de más de una década de amistad y seguir recordándome quién soy y de lo que soy capaz.

A **mi madre**, porque a la distancia siempre fue mi motor durante estos dos años. Gracias por sembrar en mí la semilla de hacer las cosas con amor y con dedicación.

A la **Dra. Karla Isabel Lira de León**, por ser siempre atenta conmigo, por facilitarme todo lo necesario para mi formación y para lograr terminar el proyecto de investigación. Ha sido una gran bendición tenerla como asesora.

A **mis sinodales, maestros y tutora**, gracias por su compromiso, por sus atentos consejos, sus observaciones, opiniones y recomendaciones, cada una de sus palabras ha resonado en mí y me ha llevado a tomar mejores decisiones académicas y profesionales.

A la **Maestría en Química Clínica Diagnóstica**, por demostrar desde el principio, durante los momentos más complicados, que están comprometidos con sus estudiantes. Gracias por apoyarnos para continuar en el programa.

A la **Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI)**, por el apoyo económico brindado durante estos 2 años, que me permitió adquirir nuevos conocimientos y vivir experiencias académicas invaluable.

A **cada uno de los participantes del estudio**, a quienes amablemente decidieron aportar muestras, reactivos, equipos, asesorías. Jamás hubiese podido completar este proyecto sin una larga lista de personas que contribuyó a su realización.

En especial, gracias a los **pacientes de cáncer de pulmón**, por permitir el estudio de su enfermedad, porque a la mayoría los conocí en un momento vulnerable, gracias por participar en esas condiciones y por confiarme sus vivencias, ojalá pudiera hacer más por ellos.

## Índice general

Página

<b>Índice de cuadros</b>	<b>iv</b>
<b>Índice de figuras</b>	<b>v</b>
<b>Abreviaturas</b>	<b>viii</b>
<b>Resumen</b>	<b>1</b>
<b>Summary</b>	<b>2</b>
<b>Declaración de responsabilidad del estudiante</b>	<b>3</b>
<b>1 Introducción</b>	<b>4</b>
1.1 Planteamiento del problema	5
1.2 Justificación	6
<b>2 Antecedentes</b>	<b>8</b>
2.1 Cáncer	8
2.1.1 Definición y epidemiología	8
2.1.2 Bases moleculares del cáncer	8
2.2 Epigenética	10
2.2.1 Metilación del ADN	10
2.2.2 Los patrones de metilación durante el cáncer	12
2.2.3 La metilación global del ADN como marcador de cáncer	13
2.2.4 Técnicas de detección del ADN metilado global	15
2.3 Cáncer de pulmón	18
2.3.1 Factores de riesgo para el cáncer de pulmón	18
2.3.2 Clasificación	19
2.3.3 Estadificación	20
2.4 Métodos diagnósticos	20
2.4.1 Biopsia líquida	22
2.5 La metilación del ADN en sangre periférica	23
2.6 Espectroscopía infrarroja	24
2.6.1 Espectroscopía infrarroja por transformada de Fourier (FTIR)	26
2.6.2 Aplicaciones clínicas de la FTIR	27
2.6.3 Regiones y bandas de interés en el estudio del cáncer de pulmón	29
2.6.4 Las conformaciones del ADN en la espectroscopía FTIR	34
<b>3 Hipótesis</b>	<b>38</b>

<b>4</b>	<b>Objetivos</b>	<b>39</b>
4.1	General	39
4.2	Particulares	39
4.3	Profesionalizantes	39
<b>5</b>	<b>Metodología</b>	<b>40</b>
5.1	Diseño	40
5.1.1	Definición del Universo	40
5.1.2	Tamaño de muestra	40
5.1.3	Definición de las unidades de observación	41
5.1.4	Grupo control	41
5.1.5	Criterios de inclusión	42
5.1.6	Criterios de exclusión	42
5.1.7	Criterios de eliminación	42
5.1.8	Definición de variables y unidades de medida	42
5.1.9	Fuentes, métodos, técnicas y procedimientos de recolección de la información.	43
5.2	Procedimientos	45
5.2.1	Obtención de muestra sanguínea	45
5.2.2	Obtención de suero	46
5.2.3	Determinación del porcentaje de metilación global del ADN	46
5.2.4	Identificación de los perfiles infrarrojos séricos	51
5.2.5	Procesamiento de los espectros FTIR-ATR	55
5.2.6	Plan de procesamiento y presentación de la información	55
5.2.7.	Consideraciones éticas	56
5.2.8.	Manejo de residuos peligrosos biológico-infecciosos (RPBI)	57
<b>6</b>	<b>Resultados y discusión</b>	<b>58</b>
6.1	Características de la población	58
6.2	Antecedentes familiares de cáncer como factor de riesgo en la población muestreada	63
6.3	Análisis de la distribución de las variables	63
6.4	Porcentaje de metilación global	65
6.4.1	Porcentaje de metilación global del ADN entre grupos	66
6.4.2	Evaluación de la asociación entre variables	67
6.5	Correlación de las variables demográficas con el porcentaje de metilación del ADN	70
6.6	Análisis espectral	72

6.6.1	Pre-procesamiento de los espectros de ADN	72
6.6.2	Análisis multivariado	82
6.6.3	Análisis por regiones espectrales	93
6.6.4	Pre-procesamiento de los espectros de suero	98
6.6.5	Análisis multivariado	100
6.6.6	Análisis por regiones espectrales	108
<b>7</b>	<b>Conclusiones</b>	<b>113</b>
<b>8</b>	<b>Referencias</b>	<b>114</b>
	<b>Anexo 1: Consentimiento informado</b>	<b>123</b>
	<b>Anexo 2: Carta de revocación del consentimiento</b>	<b>129</b>
	<b>Anexo 3. Cuestionario para participantes</b>	<b>130</b>
	<b>Anexo 4: Consentimiento informado (ampliado)</b>	<b>133</b>
	<b>Anexo 5. Análisis de plagio</b>	<b>140</b>

## Índice de cuadros

	Página
Cuadro 1. Comparación de métodos para el análisis de metilación global del ADN.....	17
Cuadro 2. Comparación de métodos diagnósticos de cáncer de pulmón.....	21
Cuadro 3. Asignaciones preliminares de los espectros IR medios del suero de pacientes con CP. ....	30
Cuadro 4. Asignación a las bandas en suero de pacientes con CP. ....	31
Cuadro 5. Asignación de bandas. ....	32
Cuadro 6. Definición de unidades de observación.....	41
Cuadro 7. Características de las variables.....	43
Cuadro 8. Características demográficas y clínicas de los participantes. ....	58
Cuadro 9. Criterios del Cuestionario Internacional de Actividad Física. ....	60
Cuadro 10. ANOVA entre variables.....	68
Cuadro 11. Asignación de bandas de ADN.....	86



## Índice de figuras

	Página
Figura 1. Efectos de mutaciones en genes .....	9
Figura 2. Alteración de los patrones de metilación en cáncer.....	13
Figura 3. Correlación de los niveles de metilación obtenidos mediante pirosecuenciación de LINE-1 y ELISA .....	15
Figura 4. Fundamento de biomarcadores estudiados por biopsia líquida.....	23
Figura 5. Aplicaciones biomédicas de la espectroscopía infrarroja. ....	29
Figura 6. Espectros FTIR de plasma.....	33
Figura 7. Espectroscopía FTIR para monitorear las señales de ADN .....	36
Figura 8. Espectros IR medios del suero de pacientes con cáncer de pulmón y de personas sanas. ....	37
Figura 9. Corrimiento electroforético del ADN extraído.....	48
Figura 10. Curva estándar de kit de metilación. ....	51
Figura 11. Liofilización de muestras de suero. ....	52
Figura 12. Liofilización de muestras de ADN. ....	54
Figura 13. Distribución sexo - edad de los participantes.....	59
Figura 14. Nivel de actividad física por grupos.. ....	61
Figura 15. Índice de Masa Corporal por grupos.....	62
Figura 16. Distribución de la variable edad. ....	64
Figura 17. Distribución de la variable IMC.....	65
Figura 18. Distribución de la variable metilación del ADN. ....	66
Figura 19. Comparación de medias del porcentaje de metilación global.....	67
Figura 20. Comparación del porcentaje de metilación del ADN por sexo.....	69
Figura 21. Comparación del porcentaje de metilación del ADN por nivel de actividad física. ....	70
Figura 22. Correlación de las variables demográficas con el porcentaje de metilación del ADN.. ....	71
Figura 23. Pre-procesamiento de espectros a las 4 horas de liofilización. ....	74

Figura 24. Pre-procesamiento de espectros de ADN a las 8 horas de liofilización...	74
Figura 25. Comparación de espectros de ADN a 4 y 8 horas de liofilización. ....	76
Figura 26. Efectos del agua en la conformación y el espectro infrarrojo del ADN ....	78
Figura 27. A: Espectros FTIR superpuestos de muestras de ADN y control de metilación. ....	79
Figura 28. Bandas reportadas para enlaces presentes en el ADN. ....	80
Figura 29. Bandas reportadas para enlaces de la citosina. ....	81
Figura 30. Bandas reportadas para la citosina metilada. ....	82
Figura 31. Mapa de varianza explicada por componentes principales en la región de 4000 a 3500 $\text{cm}^{-1}$ . ....	83
Figura 32. Mapa de varianza explicada por componentes principales en la región de 1500 a 1200 $\text{cm}^{-1}$ . ....	84
Figura 33. Mapa de varianza explicada por componentes principales en la región de 1200 a 850 $\text{cm}^{-1}$ . ....	84
Figura 34. Porcentaje de varianza explicado por cada componente principal para la región de 1200 a 850 $\text{cm}^{-1}$ . ....	85
Figura 35. Acercamiento de la región 1200 - 850 $\text{cm}^{-1}$ del espectro de ADN. ....	87
Figura 36. Acercamiento de la región 1000 - 750 $\text{cm}^{-1}$ del espectro de ADN. ....	88
Figura 37. Biplot de PCA de espectros de ADN región 1200 - 850 $\text{cm}^{-1}$ . ....	89
Figura 38. Biplot de PCA empleando el algoritmo de agrupamiento DBSCAN. ....	91
Figura 39. Espectro de muestra A110 resaltado para observar diferencias con el grupo. ....	92
Figura 40. Dendograma de las muestras de ADN. ....	93
Figura 41. Comparación de absorbancias por regiones espectrales. ....	94
Figura 42. Comparación de absorbancias por regiones espectrales. ....	96
Figura 43. Correlación entre absorbancias por ELISA y FTIR en región Amida II de muestras de ADN aislado. ....	97
Figura 44. Correlación de Pearson entre absorbancias por ELISA y FTIR. ....	98
Figura 45. Pre-procesamientos de los espectros de suero a las 16 horas de liofilización. ....	99

Figura 46. Espectros promedio del suero.....	100
Figura 47. Mapa de varianza explicada por componentes principales en la región de 4000 a 500 cm <sup>-1</sup> .....	101
Figura 48. Porcentaje de varianza explicado por cada componente principal para todo el espectro del suero. ....	102
Figura 49. Biplot de PCA de espectros de suero. ....	103
Figura 50. Biplot de PCA empleando el algoritmo de agrupamiento DBSCAN. ....	104
Figura 51. Dendograma de las muestras de suero. ....	105
Figura 52. PCA con muestras indicadas por características demográficas. ....	106
Figura 53. PCA de muestras de suero eliminando regiones espectrales ....	107
Figura 54. Área de banda espectral calculada por regiones.....	111
Figura 55. Correlación de Pearson entre absorbancias por ELISA y FTIR en región Metilación de muestras de suero.....	112

## Abreviaturas

<i>5-mC</i>	5-metilcitosina
<i>ACP</i>	Análisis de componentes principales
<i>ADN</i>	Ácido desoxirribonucleico
<i>ADNlc</i>	Ácido desoxirribonucleico libre circulante
<i>ADNtc</i>	Ácido desoxirribonucleico tumoral circulante
<i>ANOVA</i>	Análisis de varianza
<i>ARN</i>	Ácido ribonucleico
<i>C2</i>	Carbono 2
<i>C3</i>	Carbono 3
<i>C5</i>	Carbono 5
<i>CP</i>	Cáncer pulmonar/cáncer de pulmón
<i>CPCNP</i>	Cáncer de pulmón de células no pequeñas
<i>CPCP</i>	Cáncer de pulmón de células pequeñas
<i>CTC</i>	Células tumorales circulantes
<i>DNMT</i>	ADN metiltransferasas
<i>DO</i>	Densidad óptica
<i>ELISA</i>	Ensayo por inmunoabsorción ligado a enzimas
<i>EPOC</i>	Enfermedad pulmonar obstructiva crónica
<i>FTIR</i>	Espectroscopía Infrarroja por Transformada de Fourier
<i>HPLC</i>	Cromatografía líquida de alta eficiencia
<i>IR</i>	Infrarrojo
<i>FTIR</i>	Espectroscopía infrarroja con transformada de Fourier
<i>MRI</i>	Imagen por resonancia magnética
<i>PET</i>	Tomografía por emisión de positrones
<i>ATR</i>	Accesorio de Reflectancia Total Atenuada
<i>TC</i>	Tomografía computarizada
<i>VE</i>	Vesículas extracelulares

## Resumen

Las alteraciones de los mecanismos epigenéticos, como la metilación del ADN, se han propuesto como potenciales marcadores para el cáncer de pulmón. Sin embargo, las metodologías convencionales de detección de metilación global de ADN son costosas y laboriosas. La espectroscopía infrarroja por transformada de Fourier (FTIR), presenta ventajas como la rapidez en la obtención de los datos a partir de una variedad de muestras con mínima preparación. Bandas espectrales de la región 1800 a 800  $\text{cm}^{-1}$  indican información sobre la estructura de los ácidos nucleicos, con la posibilidad de detectar cambios como los ocasionados por la metilación del ADN. Por lo tanto, el objetivo del presente estudio fue explorar la asociación entre datos de infrarrojo de muestras séricas y datos espectrofotométricos del grado de metilación del ADN en pacientes con cáncer de pulmón. Se reclutaron 20 participantes del grupo control y 6 pacientes con cáncer de pulmón. Los perfiles infrarrojos se obtuvieron mediante el análisis de suero y ADN en un espectrómetro FTIR-ATR y el grado de metilación global del ADN mediante un ensayo ELISA. Se observó una disminución estadísticamente significativa del 13.6 % en la metilación global del ADN en el grupo de cáncer de pulmón en comparación con el grupo control ( $0.985 \pm 0.143$  vs.  $1.140 \pm 0.131$ ; prueba de Welch,  $p = 0.048$ ), confirmando un proceso de hipometilación durante el cáncer. En el análisis espectral del suero, la región de 1470 – 1250  $\text{cm}^{-1}$  fue la que tuvo mayor contribución a la identificación de la metilación, mostrando una absorbancia del área de banda disminuida para el grupo cáncer (prueba t de Student,  $p = 0.0421$ ). El análisis de correlación entre el nivel de metilación detectado con ELISA y con FTIR en las regiones espectrales identificadas fue débil para las muestras con suero (correlación de Pearson,  $r = 0.247$ ,  $p = 0.223$ ), sin embargo, para las muestras de ADN la correlación fue moderada (correlación de Pearson,  $r = 0.501$ ,  $p = 0.0127$ ). Aunque el análisis espectral del ADN puede aportar información relevante sobre su estructura, la detección de la metilación es una metodología que requiere ajustes que reducirían su practicidad en el área clínica.

**Palabras clave:** cáncer, metilación, FTIR.

## Summary

Alterations in epigenetic mechanisms, such as DNA methylation, have been suggested as potential lung cancer markers. However, conventional methodologies for detecting global DNA methylation are expensive and laborious. Fourier transform infrared spectroscopy (FTIR) offers advantages like rapid data acquisition from various samples with minimal preparation. Spectral bands in the 1800–800  $\text{cm}^{-1}$  region provide information about nucleic acid structure and can detect changes caused by DNA methylation. Thus, this study aimed to explore the association between infrared data from serum samples and spectrophotometric data on the degree of DNA methylation in lung cancer patients. Twenty participants were recruited from the control group and six patients with lung cancer. Infrared profiles were obtained by analyzing serum and DNA on an FTIR-ATR spectrometer, and the degree of global DNA methylation was determined using an ELISA assay. A statistically significant decrease of 13.6% in overall DNA methylation was observed in the lung cancer group compared to the control group ( $0.985 \pm 0.143$  vs.  $1.140 \pm 0.131$ ; Welch's test,  $p = 0.048$ ), which confirms the hypomethylation process that occurs during cancer development. Serum spectral analysis revealed that the 1470–1250  $\text{cm}^{-1}$  region contributed most to identifying methylation. This region showed decreased band area absorbance in the cancer group (Student's t-test,  $p = 0.0421$ ). Correlation analysis revealed a weak correlation between the level of methylation detected by ELISA and FTIR in the identified spectral regions for serum samples (Pearson's correlation,  $r = 0.247$ ,  $p = 0.223$ ). However, a moderate correlation was observed for DNA samples (Pearson's correlation,  $r = 0.501$ ,  $p = 0.0127$ ). While spectral analysis of DNA can provide relevant information about its structure, detecting methylation requires adjustments that reduce its practicality in clinical settings.

**Keywords:** cancer, methylation, FTIR.

### **Declaración de responsabilidad del estudiante**

Declaro que los datos propios obtenidos en esta investigación fueron generados durante el desarrollo de mi trabajo de tesis de forma ética y que reporto detalles necesarios para que los resultados de esta tesis sean reproducibles en eventuales investigaciones futuras. Finalmente, este manuscrito de tesis es un trabajo original en cual se declaró y dio reconocimiento a cualquier colaboración o cita textual presentadas en el documento.

## **1 Introducción**

El cáncer de pulmón (CP) es la neoplasia maligna con la mayor tasa de mortalidad. En la fase inicial del CP, los pacientes normalmente no presentan ningún síntoma aparente, siendo una de las razones por las que acuden al servicio médico cuando la enfermedad se manifiesta con síntomas que impiden la realización de actividades cotidianas. Por otro lado, el acceso limitado a métodos diagnósticos, así como la falta de éstos en centros de salud, hacen que el diagnóstico de esta enfermedad, en la mayoría de los casos, sea tardío (Hana & Faizah, 2023).

Actualmente se considera que tanto la aparición como la progresión del cáncer son resultado de las interacciones entre diferentes cambios genéticos, modificaciones epigenéticas y una diversidad de factores ambientales. El descubrimiento de la existencia de una huella de metilación del ADN alterada en células cancerosas ofrece ventajas potenciales: a) puede manifestarse como una hipometilación global o una hipermetilación específica en genes supresores de tumores, característico de una etapa temprana de desarrollo tumoral, funcionando como un marcador de diagnóstico precoz y b) la metilación en ADN puede detectarse mediante muestras obtenidas por técnicas mínimamente invasivas o incluso no invasivas (Hong & Kim, 2021; Ortiz Zamora et al., 2023). Las muestras de plasma o suero pueden considerarse óptimas en comparación con las de tejido, ya que además de ser de fácil acceso, no están sujetas a sesgos de biopsia y pueden extraerse repetidamente del mismo paciente (Souza et al., 2023).

En el contexto de las tecnologías accesibles y sustentables que pueden analizar muestras de plasma o suero, el uso de la espectroscopía infrarroja por transformada de Fourier (FTIR) tanto para el diagnóstico funcional como para el seguimiento de tratamientos en cáncer ha ganado atención en la última década, estableciendo la oportunidad de seguir investigando y aterrizando su aplicación en el ámbito clínico (Vitorino et al., 2023).



## 1.1 Planteamiento del problema

El cáncer de pulmón representa un preocupante problema de salud mundial, ya que su elevada incidencia genera elevados costos de atención para el sector de salud pública (GLOBOCAN, 2025). En México, el 75 % de los pacientes que sufren cáncer pulmonar y que acuden al servicio de salud, al momento del diagnóstico ya padecen la enfermedad en etapa avanzada. Cuando el cáncer de pulmón logra detectarse en etapa temprana hay una mayor probabilidad de que el paciente responda al tratamiento, aumentando sus probabilidades de supervivencia (Heredia et al., 2023).

Un diagnóstico oportuno y la determinación del pronóstico del paciente son de suma importancia para abordar esta problemática. En la actualidad, solo se ha visto una mejora en el diagnóstico de cáncer de pulmón mediante el tamizaje con tomografía axial computarizada (TAC) de tórax de baja dosis (National Cancer Institute, 2023), desafortunadamente, este método no es accesible para toda la población y su aplicación rutinaria representa una gran inversión para los servicios de salud pública. Por lo tanto, es necesario explorar nuevas herramientas que mejoren el rendimiento diagnóstico. A diferencia de la TAC de tórax de baja dosis, la mayoría de tecnologías emergentes que permiten detectar biomarcadores moleculares no implican la exposición a la radiación, son relativamente más baratas y de alto rendimiento, por lo que satisfacen la necesidad de un proceso de tamizaje rentable y rápido (Guo et al., 2023).

Los cambios en los patrones de la metilación del ADN son candidatos prometedores para desarrollar biomarcadores del cáncer de pulmón, no obstante, se necesitan más investigaciones que puedan respaldar sus aplicaciones clínicas (Hong & Kim, 2021). Por otro lado, las técnicas convencionales de detección de la metilación global del ADN son laboriosas, costosas y casi exclusivas de laboratorios de investigación que han estandarizado su metodología. En México, son pocos los estudios que han analizado el estado de la metilación global del ADN en pacientes

con cáncer de pulmón (Alcázar-Ramos et al., 2019), lo que ha dificultado el avance en el desarrollo del conocimiento y aplicación de las ventajas que ofrece este mecanismo epigenético como potencial biomarcador de cáncer de pulmón.

## **1.2 Justificación**

Se ha encontrado que en pacientes con diferentes tipos de cáncer existen alteraciones en los patrones de metilación global de ADN, además, algunas pruebas sugieren que cierta metilación aberrante del ADN en sangre podría reflejar cambios patológicos en tejidos diana que no podrían obtenerse de forma fácil o segura (Hong & Kim, 2021). Además, cualquier alteración en los sistemas biológicos inducida por una condición patológica produce cambios estructurales y funcionales significativos que se reflejan directamente en los espectros vibracionales (Balan et al., 2019).

Al ser más sencilla, rápida, precisa, barata, no destructiva y apta para la automatización que los actuales métodos de detección, diagnóstico, gestión y seguimiento, la espectroscopía FTIR puede mejorar potencialmente el enfoque de las intervenciones clínicas y el tratamiento de los pacientes al detectar cambios bioquímicos en las muestras de pacientes con cáncer de pulmón (Su & Lee, 2020).

La capacidad de la espectroscopía FTIR para mostrar bandas correspondientes a la estructura de los ácidos nucleicos plantea la posibilidad de utilizarla como técnica de detección de la metilación global del ADN, ya que se han reportado bandas que puedan estar indicando la presencia de los grupos metilo en el ADN. Si bien existen estudios que analicen la metilación del ADN, estos son enfocados a genes específicos, procedimiento que conlleva un mayor tiempo de realización y costo, a diferencia de evaluar la metilación global del ADN. Por otro lado, a la fecha, son pocas las investigaciones de espectroscopía FTIR en cáncer de pulmón y ninguna ha sido realizada en México. Además, son escasos los estudios que han buscado explorar si la metilación del ADN puede ser detectada en muestras biológicas de pacientes mediante espectroscopía FTIR.

Por lo tanto, el objetivo general de este estudio es explorar la correlación de datos de infrarrojo de muestras séricas con datos espectrofotométricos del grado de metilación del ADN en pacientes con y sin cáncer de pulmón.

Comprobar una asociación entre el porcentaje de metilación global del ADN y regiones espectrales determinadas de FTIR podría permitir en un futuro que la detección de la metilación global se realizara mediante esta técnica, simplificando el proceso y ahorrando costos de técnicas como la espectrometría de masas o la secuenciación.

## **2 Antecedentes**

### **2.1 Cáncer**

#### **2.1.1 Definición y epidemiología**

Bajo el término genérico cáncer, neoplasia o tumor maligno, se engloba a un conjunto de enfermedades que pueden originarse en casi cualquier órgano o tejido del cuerpo cuando se da un crecimiento descontrolado de las células anormales, y éstas posteriormente se desplazan fuera de sus límites normales para invadir otros órganos en un proceso llamado metástasis (Organización Mundial de la Salud, 2023). El tipo de célula de origen, las causas, los mecanismos subyacentes, así como el grado de malignidad, son característicos de cada tipo de cáncer (Herráez, 2012). De acuerdo con la Organización Mundial de la Salud (OMS), en la actualidad se considera al cáncer como la principal causa de muerte en el mundo, ocasionando casi 10 millones de defunciones en el todo el mundo durante el 2022, siendo el cáncer de pulmón el que causó el mayor número de fallecimientos (1.8 millones), seguido del cáncer colorrectal y el hepático (Organización Mundial de la Salud, 2023). En México, durante el 2022 hubo un registro de 841,318 defunciones, de las cuales 96, 210 fueron a causa de cáncer y de estas 7,808 por cáncer de pulmón (GLOBOCAN, 2025).

#### **2.1.2 Bases moleculares del cáncer**

##### **2.1.2.1 Equilibrio entre proliferación y muerte celular**

El cáncer es considerado una enfermedad multifactorial que implica tanto una falta de control en el crecimiento y la proliferación celular como una evolución maligna, es decir, la pérdida de los rasgos y la fisiología celular normal. Cada órgano presenta un equilibrio específico entre los procesos de proliferación y muerte celular, que se regula dependiendo la función de sus tejidos. La alteración de ese equilibrio con predominio de la proliferación celular ya sea por mecanismos genéticos o epigenéticos de genes que codifican para proteínas reguladoras es el principio que da origen a las neoplasias (Salazar Montes et al., 2016).

### 2.1.2.2 Genes responsables del cáncer

Las mutaciones que ocurren en las células cancerígenas perturban directamente a dos tipos de genes cuya alteración está íntimamente relacionada con la carcinogénesis: oncogenes y genes supresores de tumores (también llamados oncosupresores), así como a los productos de éstos (Herráez, 2012). Como se observa en la Figura 1, existen dos formas de manifestación de la proliferación celular descontrolada: a través de la activación de protooncogenes por el incremento de factores proliferativos, o por la inactivación de genes supresores de tumores por la disminución de los factores inhibidores (Iwasa & Marshall, 2019).

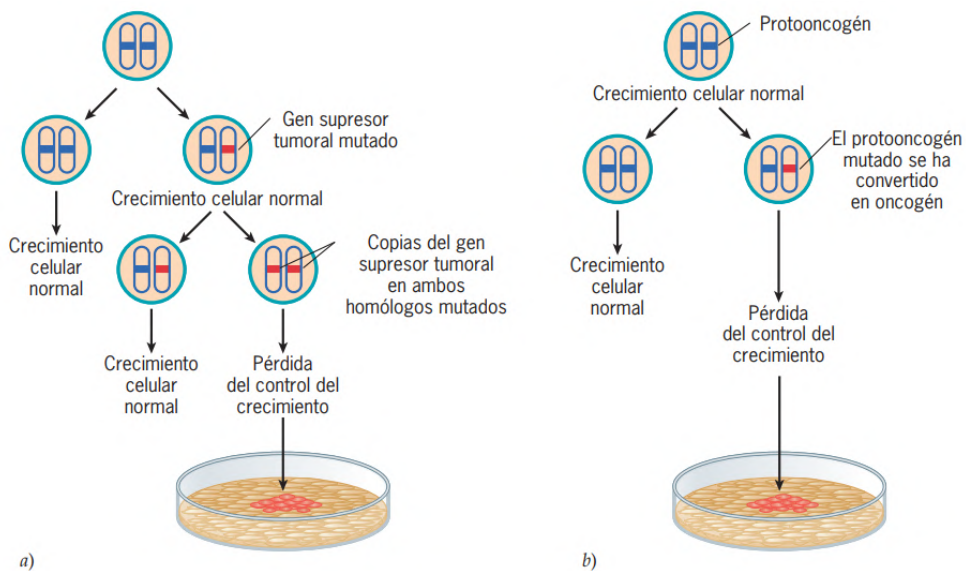


Figura 1. Efectos de mutaciones en genes a) supresores de tumores y b) oncogenes (Iwasa y Marshall, 2019).

Los oncogenes son versiones mutadas de los protooncogenes, que generalmente codifican para proteínas que cuando se activan incorrectamente por mutaciones pueden promover la proliferación descontrolada de las células. Los genes supresores de tumores, por el contrario, codifican proteínas cuya función principal es bloquear/controlar el desarrollo del cáncer (Fior & Zilhão, 2019).

## 2.2 Epigenética

El término epigenética hace referencia al estudio del conjunto de modificaciones en la expresión génica que acontecen sin producir alteraciones estructurales en la secuencia del ADN. Anteriormente se creía que la heredabilidad de la información epigenética se limitaba a las divisiones celulares mitóticas, sin embargo, ahora se sabe que los procesos epigenéticos pueden transferirse a través de la meiosis de una generación a otra en los organismos (Tollefsbol, 2023). Actualmente se conocen cuatro mecanismos epigenéticos que actúan coordinadamente: la metilación del ADN, las modificaciones covalentes de las histonas (acetilación, metilación, fosforilación, ubiquitinación), la regulación por ARN no codificante o dependiente de micro ARNs y la remodelación de la cromatina (Ortiz Zamora et al., 2023).

### 2.2.1 Metilación del ADN

El proceso epigenético más estudiado es la metilación del ADN. Esta se refiere a la transferencia mediada por enzima de un grupo metilo de la S-adenosilmetionina (SAM) al carbono 5 de las citosinas que se encuentran en determinados dinucleótidos CpG. En el genoma humano se ha demostrado la presencia de regiones de secuencias que poseen un elevado contenido de citosina y guanina unidos por fosfatos, denominadas islas CpG. Esta importante reacción de transferencia está catalizada por las ADN metiltransferasas (DNMT), las cuales se clasifican en: las ADN-metiltransferasas de mantenimiento (DNMT1) y las ADN-metiltransferasas *de novo* (DNMT3A, DNMT3B y DNMT3L). Se denomina metilación de mantenimiento cuando ésta se produce durante cada replicación celular a medida que el ADN se duplica, mientras que la metilación *de novo* conduce a la introducción de 5-metilcitosinas (5-mC) en el genoma en sitios que no estaban previamente metilados (Rojas Moreno et al., 2016; Tollefsbol, 2023). En células sanas, mientras que las secuencias genómicas repetitivas se encuentran metiladas, la mayoría de las islas CpG no lo están, permitiendo que los genes se expresen al presentarse los activadores transcripcionales correspondientes. Sin embargo, las

regiones promotoras de los genes de células normales se metilan durante ciertos procesos de desarrollo como la impronta de genes, genes del cromosoma X en las mujeres y genes específicos de tejidos y de la línea germinal. Aunque la metilación no es necesaria durante la regulación constitutiva de los procesos de expresión génica, a veces puede cumplir ese propósito, como en el caso de los genes específicos de tejidos, algunos de ellos contienen islas CpG, mientras que otros contienen sólo unos pocos dinucleótidos CpG dispersos en su región reguladora 5' (Tollefsbol, 2023).

El epigenoma de un individuo puede verse influenciado durante tres etapas cronológicas: la etapa ancestral (que se refiere a las marcas epigenéticas de la línea parental que permanecieron en la línea germinal y se heredaron), la etapa prenatal (alteraciones epigenéticas ocasionadas por la exposición de la madre a diversos factores) y la etapa postnatal (la influencia de factores ambientales sobre el individuo desde que nace hasta el final de su vida) (Aguilera et al., 2010; Esteller, 2019; Tollefsbol, 2023).

Existe una amplia diversidad de factores que se sospecha pueden tener una baja o alta influencia en alguno o varios de los mecanismos epigenéticos, que pueden ir desde factores intrínsecos como la calidad de vida, los cuidados maternos, el aprendizaje, el estrés, el envejecimiento, hasta agentes químicos y físicos, como contaminantes ambientales, metales pesados, medicamentos y drogas (Esteller, 2019; Tollefsbol, 2023). La investigación referente a los impactos ambientales que conducen a los cambios transcripcionales mediante los diferentes mecanismos epigenéticos sigue siendo insuficiente hasta la fecha, principalmente por la falta de la estandarización de una técnica para cuantificar los cambios en las marcas epigenéticas y a la dificultad para realizar estudios longitudinales en cohortes humanas que permitan comparar los cambios. En este contexto, se sabe que la edad, el sexo y la actividad física regular son los principales factores biológicos que pueden afectar los valores globales de metilación del ADN (Lam et al., 2012a; Szigeti et al., 2022).

### 2.2.2 Los patrones de metilación durante el cáncer

Cuando las células normales se transforman en cancerígenas los patrones de la metilación del ADN cambian notablemente como resultado de dos fenómenos principales que se ilustran en la Figura 2, primero, el genoma tumoral se vuelve globalmente hipometilado, a diferencia de las células normales, debido principalmente a la desmetilación generalizada en los CpGs dispersos por toda la secuencia génica. En segundo lugar, puede ocurrir una intensa hipermetilación de las regiones promotoras de genes supresores de tumores (Esteller, 2019).

Durante el cáncer, las células pierden del 20 al 40 % de su contenido normal de 5-metilcitosinas, principalmente debido a la hipometilación de las regiones codificantes y los intrones de los genes, y también debido a la desmetilación de las secuencias repetitivas que conforman entre el 20 y 30 % del genoma humano. La pérdida global de metilación del ADN da lugar a deleciones, translocaciones y reordenamientos cromosómicos en los tumores. La hipometilación del genoma es un acontecimiento que ocurre a inicios del desarrollo del cáncer y se acumula a lo largo de todas las etapas, desde la proliferación benigna hasta el cáncer invasivo. Por otro lado, la hipermetilación está asociada a las islas CpG de los promotores de los genes supresores de tumores, que en tejidos normales no están metilados, pero que durante el desarrollo del cáncer se metilan y conducen al silenciamiento de esos genes. La presencia de la hipermetilación en los promotores de las islas CpG afecta la regulación de los genes de casi todas las funciones celulares, entre las que destacan el ciclo celular, la adherencia celular, apoptosis, metabolismo carcinogénico, respuesta hormonal, señalización vía Ras y microRNAs. La desregulación de los genes implicados en estos mecanismos celulares está íntimamente implicada con el cáncer (Ortiz Zamora et al., 2023).



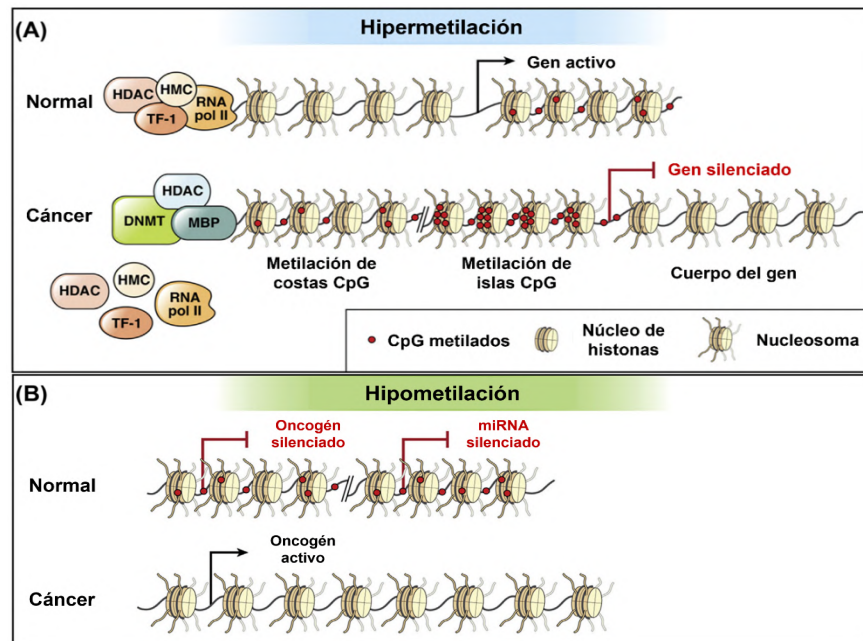


Figura 2. Alteración de los patrones de metilación en cáncer. TF-1, factor de transcripción 1; RNA pol II, ARN polimerasa II; HDAC, complejo histona deacetilasa; DNMT, ADN metiltransferasa; MBP, proteína de unión a metilo; miRNA, micro ARN. A) Hipermetilación aberrante del ADN. B) Hipometilación aberrante del ADN. Modificado de (Tollefsbol, 2023).

### 2.2.3 La metilación global del ADN como marcador de cáncer

El nivel de metilación global del ADN puede funcionar como una medida integrada e intermedia del desarrollo del cáncer, ya que es un reflejo de la acumulación de múltiples factores tanto reversibles como irreversibles, incluyendo edad, sexo, exposiciones ambientales y polimorfismos genéticos (Kuchiba et al., 2014; Z.-Z. Zhu et al., 2012).

En un estudio de casos y controles que evaluó el grado de metilación global del ADN en sangre periférica como factor de riesgo para el cáncer de mama mediante ensayo de metilación luminométrica (LUMA), los niveles más bajos de metilación global se asociaron con un riesgo significativamente mayor de cáncer de mama (tendencia OR = 1.68; IC 95 % IC = 1.35 – 2.09) (Kuchiba et al., 2014).

En otro estudio para evaluar si la metilación de ADN libre circulante detectada en suero y plasma era capaz de discernir entre cáncer de pulmón, fibrosis quística y enfermedad pulmonar obstructiva crónica (EPOC) se utilizó la digestión con enzimas de restricción sensibles a la metilación y el enriquecimiento del ADN metilado mediante amplificación dirigida. Se identificaron 63 loci candidatos diferencialmente metilados que se utilizaron para construir un modelo predictivo y estimar el valor diagnóstico: el análisis ROC reveló un área bajo la curva (AUC) de 0.91 (IC 95 %: 0.84 - 0.96) para el cáncer de pulmón, 0.815 (IC 95%: 0.73 - 0.88) para la fibrosis quística, 0.73 (IC 95 %: 0.62 – 0.83) para la EPOC y 0.828 (IC 95 %: 0.76 – 0.89) para todas las enfermedades frente a los controles sanos (Wielscher et al., 2015).

Por otro lado, se han comparado métodos de detección de metilación global evaluando directamente su potencial como marcador de tamizaje para cáncer. Un grupo de investigación cuantificó la metilación global de muestras de tejido de pacientes con cáncer cervicouterino, de pacientes con lesiones intraepiteliales benignas de cérvix, líneas celulares de cáncer cervicouterino y leucocitos de personas sanas como control mediante la pirosecuenciación de elementos nucleares largos intercalados (LINE-1) y la comparó con los valores obtenidos mediante un ELISA. Se reportaron resultados de hipometilación consistentes con ambos métodos (Figura 3), y se encontraron diferencias estadísticamente significativas al comparar los niveles de metilación entre muestras de cáncer y de lesiones precancerosas mediante ELISA (3.03 % - 3.85 %, respectivamente,  $p < 0.05$ ) y mediante la pirosecuenciación de LINE-1 (48.8 % - 56.9 %, respectivamente,  $p < 0.05$ ), proponiendo que la metilación global puede ser utilizada como test de tamizaje para el cáncer cervicouterino (Thumbovorn et al., 2022).

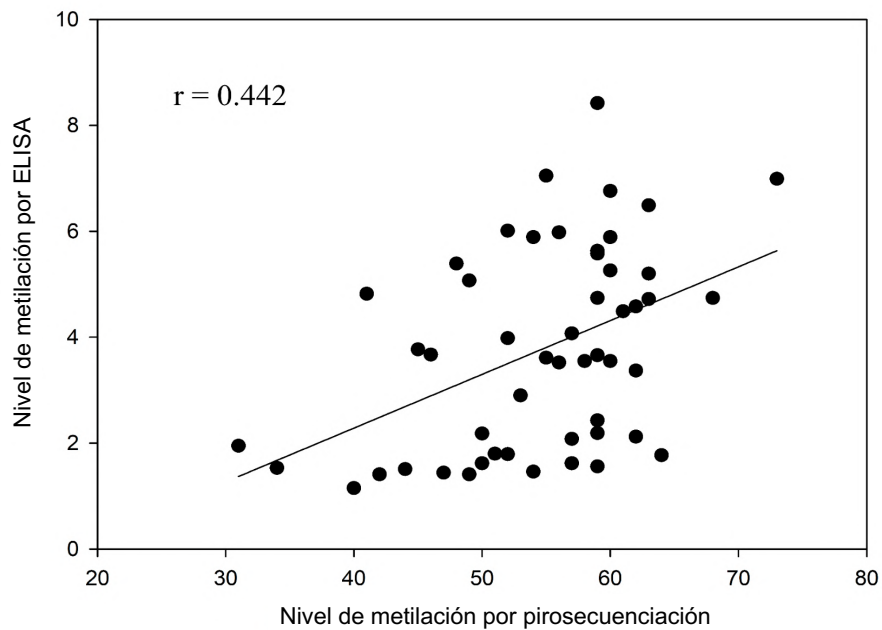


Figura 3. Correlación de los niveles de metilación obtenidos mediante pirosecuenciación de LINE-1 y ELISA (Modificado de Thumbovorn et al. 2022).

#### 2.2.4 Técnicas de detección del ADN metilado global

Actualmente existe una variedad de técnicas de detección de la metilación global del ADN, la mayoría enfocadas solamente al estudio de regiones específicas.

##### Ensayo de inmunoabsorción ligado a enzima (ELISA)

Es un método simple y rentable adecuado para determinar el nivel general de metilación del ADN. El contenido de citosina de 5-mC de las muestras desconocidas es determinado a través de su comparación con una curva estándar de 5-mC predeterminada. El ADN metilado se detecta mediante anticuerpos específicos y luego se cuantifica leyendo la absorbancia en un espectrofotómetro de microplacas. Por lo tanto, el porcentaje de ADN metilado es proporcional a la densidad óptica medida (Martisova et al., 2021).

## Espectrometría de masas

La cuantificación más precisa del ADN se basa en la determinación de la abundancia absoluta de las diferentes bases del ADN y sus modificaciones mediante el uso de monitoreo de reacciones múltiples en un espectrómetro de masas (MS-MRM) acoplado a nano-ultra Cromatografía Líquida de Alta Eficiencia (HPLC). Esta técnica ofrece muchas ventajas, ya que la detección es muy precisa y sensible, además, el análisis abarca todo el genoma, independientemente del sitio o la secuencia. Sin embargo, también tiene una gran desventaja, la complejidad técnica para su realización, además de los grandes costos asociados (Fernandez et al., 2018).

## Ensayo de metilación luminométrica (LUMA)

Esta detección está basada en la digestión del ADN con enzimas de restricción sensibles a la metilación (*HpaII*) o insensibles (*MspI*), seguida de un ensayo de bioluminiscencia extendido con polimerasa para cuantificar el grado de escisión de restricción. Al no requerir una conversión previa del ADN genómico con bisulfito, esta técnica es relativamente rápida, sin embargo, tiene las siguientes desventajas: los resultados obtenidos se ven afectados por la técnica de aislamiento de ADN y la escisión enzimática se limita a los sitios de restricción 5'-CCGG-3', que representan solo alrededor del 8 % de todos los CpG en el genoma (Martisova et al., 2021).

## LINE 1 y pirosecuenciación

Los Elementos largos intercalados 1 (LINE 1) son una familia de retrotransposones de ADN repetitivo que constituye el 17 % del genoma humano, por lo tanto, su grado de metilación se considera un marcador sustituto del ADN global, para su detección tiene que hacerse conversión con bisulfito y amplificación por PCR para su posterior pirosecuenciación (Kurdyukov & Bullock, 2016).

En el Cuadro 1, modificado de (Kurdyukov & Bullock, 2016) se enlistan los tres métodos más utilizados en investigación clínica para el análisis de metilación global del ADN.

Cuadro 1. Comparación de métodos para el análisis de metilación global del ADN.

Método	Cobertura del genoma	Nivel mínimo de detección	Cantidad de ADN inicial	Costo por muestra
LC-MS/MS	100 %	0.05 % - 10 %	100 ng - 1ug	\$1,374.90
LINE-1 y pirosecuenciación	17 %	0.5 %	50 ng	\$376.50
ELISA	100 %	0.05 %	5-50 ng	\$116.01

Debido a la diversidad de métodos de detección de la metilación global del ADN, la comparación entre los resultados obtenidos por cada investigación puede complicarse, ya que cada equipo o método arroja resultados en distintas unidades. Sin embargo, (Alcázar-Ramos, 2019) en un proyecto de investigación de tesis de maestría, reportó un promedio de metilación global del 3.31 % para un grupo de no fumadores utilizado como control, mientras que se obtuvo un promedio de metilación global de 1.8 % para el grupo de pacientes con cáncer de pulmón. Un segundo estudio de la misma línea de investigación (Beltran, 2022) reportó un promedio de 0.32 % para el grupo de no fumadores y 1.58 % para el grupo con cáncer de pulmón. La importancia de estos trabajos es el empleo de la misma técnica de detección (ELISA), por lo que los resultados obtenidos en el presente proyecto podrán ser discutidos conforme a lo reportado previamente.

## **2.3 Cáncer de pulmón**

Según el Observatorio Global del Cáncer (GLOBOCAN) en 2022 el cáncer de pulmón se consideró el tercer tipo de cáncer con más frecuencia en el mundo (2, 480 675 casos) y constituye el cáncer con mayor tasa de mortalidad. En los hombres, éste representa el tumor maligno más común, mientras que en mujeres es el segundo. El cáncer de pulmón se manifiesta clínicamente de distintas formas, y a menudo se presenta con pocos o ningún síntoma en una fase temprana. Dado que los primeros síntomas del cáncer de pulmón son mínimos e inespecíficos, a muchos pacientes se les diagnostica en una fase avanzada (Hong & Kim, 2021).

### **2.3.1 Factores de riesgo para el cáncer de pulmón**

El consumo de tabaco se considera el mayor factor de riesgo para el cáncer pulmonar, los fumadores tienen entre 15 y 30 veces más probabilidades de sufrirlo o morir a causa de este cáncer que las personas no fumadoras, y representa la causa del 80 % de casos y defunciones (Division of Cancer Prevention and Control, Centers for Disease Control and Prevention, 2023). La exposición involuntaria a humo de tabaco es otro factor de gran importancia, ya que el humo de segunda mano contiene numerosas sustancias químicas tóxicas y al menos 50 carcinógenos, además, se asocia con el mayor riesgo de cáncer de pulmón de células pequeñas, por consiguiente, cualquier nivel de exposición a este puede representar un riesgo (Brody et al., 2021).

Se ha establecido que la exposición al gas radón representa la segunda causa de cáncer de pulmón, con el mayor riesgo para los fumadores. También se consideran riesgosas las exposiciones ocupacionales a diferentes compuestos carcinógenos (amianto, radiación, arsénico, cromatos, níquel, éteres de clorometil, hidrocarburos aromáticos policíclicos, gas mostaza, emisiones de horno de coque, humo de cocina artesanal y calefacción en chozas o cabañas), siendo las personas que trabajan en la producción de aluminio, la pintura y la fundición las que tienen un riesgo elevado de desarrollar cáncer de pulmón (American Cancer Society, 2023).

La inhalación de partículas, una mezcla microscópica de contaminantes sólidos y líquidos presentes en la compleja variedad de componentes de la contaminación atmosférica está relacionada con un aumento de la susceptibilidad de padecer cáncer pulmonar de entre el 8 y el 9 % (International Agency for Research on Cancer, 2016).

La inflamación crónica, como la que sucede durante diferentes padecimientos incluidos el asma, la bronquitis crónica, neumonía, enfisema, la enfermedad pulmonar obstructiva crónica (EPOC), la deficiencia de alfa-1 antitripsina y la fibrosis pulmonar, aumentan la vulnerabilidad de padecer cáncer pulmonar (aumento de riesgo entre el 16 % en las personas con antecedentes de asma y 2.5 veces entre las que tienen antecedentes de EPOC) (American Cancer Society, 2023). De igual forma, antecedentes de infecciones que ocasionaron cicatrices en los pulmones, como la tuberculosis, aumentan la probabilidad de desarrollar la enfermedad (Heredia et al., 2023).

La predisposición genética también se ha considerado un factor de riesgo en las últimas décadas, el Consorcio Internacional de Cáncer de Pulmón estima que las personas que tengan un familiar de primer grado con cáncer de pulmón presentan un riesgo un 50 % mayor de padecer la enfermedad, aumentando al 82 % cuando se trata de un hermano diagnosticado con la enfermedad (S. Li et al., 2022).

### 2.3.2 Clasificación

El cáncer de pulmón engloba una variedad de tumores biológicamente distintos. Los diferentes tipos de cáncer de pulmón se organizan en dos grupos: el cáncer de pulmón no microcítico (CPNM), mayormente conocido como cáncer de pulmón de células no pequeñas (CPCNP), que representa el 81 % de los casos, y el cáncer de pulmón de células pequeñas (CPCP), que representa el 14 % de los casos (American Cancer Society, 2023). El CPCNP se clasifica a su vez en adenocarcinoma, que es más frecuente en las mujeres, seguido del carcinoma de células escamosas y carcinoma de células grandes. El adenocarcinoma se origina

en las glándulas que segregan moco, y por lo general, es más tratable que los otros subtipos debido a la alta probabilidad de su localización en la periferia pulmonar y a que presente mutaciones que puedan ser objeto de tratamientos. El carcinoma de células escamosas se disemina más rápido que el adenocarcinoma y tiene su origen en las células que se encuentran recubriendo las vías respiratorias pulmonares. El carcinoma de células grandes es el más maligno de los tres CPCNP, ya que puede originarse en cualquier parte del pulmón (Thai et al., 2021). El subtipo de CPCP más frecuente es el carcinoma de células pequeñas, seguido del carcinoma combinado de células pequeñas. El CPCP es generalmente más agresivo que el CPCNP, por lo que es más probable que al momento del diagnóstico de pacientes con CPCP la enfermedad ya se haya extendido más allá de los pulmones, dificultando su tratamiento (Miller et al., 2022).

### 2.3.3 Estadificación

La determinación del estadio del cáncer de pulmón permite al médico conocer si el cáncer se ha extendido y a qué distancia, ayudando a establecer el mejor tratamiento. Para el CPCNP por lo general se utiliza el sistema de clasificación TNM del American Joint Committee on Cancer (AJCC), que se basa en tres características: 1) el tamaño y la extensión del tumor principal (T), 2) la propagación a los ganglios (nódulos) linfáticos adyacentes (N) y 3) la propagación a sitios distantes, conocida como metástasis (M). Después de determinar las categorías T, N y M, estos datos se agrupan para determinar una etapa general, de acuerdo con la clasificación más reciente del sistema, que se puede consultar en las guías médicas (Detterbeck, 2018).

## 2.4 Métodos diagnósticos

La radiografía de tórax es el estudio que normalmente se realiza primero en el abordaje del diagnóstico del cáncer de pulmón, y puede ir acompañada o seguida de una tomografía computarizada. El diagnóstico definitivo se realiza mediante una biopsia de tejido (Arroyo-Hernández et al., 2022). Sin embargo, dependiendo del



caso y del enfoque del médico, se puede optar por una variedad de métodos diagnósticos que pueden ser útiles en distintas etapas de la enfermedad.

En el Cuadro 2 se resumen algunos de estos métodos y se comparan de acuerdo con las principales características y limitaciones.

Cuadro 2. Comparación de métodos diagnósticos de cáncer de pulmón.

<b>Tipo de método</b>	<b>Método</b>	<b>Estadios detectados</b>	<b>Limitaciones</b>
Imagenología	Radiografía de tórax	Inicial	Requiere confirmación.
	Tomografía axial computarizada (TAC)	Tamizaje, I-II	Sensible pero poco específico, requiere complementarse.
	Imágenes por resonancia magnética (MRI)	Avanzados	Para determinar metástasis a médula espinal o cerebro.
	Tomografía por Emisión de Positrones (PET)	Iniciales y avanzados	Inyección de análogo radiactivo de la glucosa.
Citología	Con esputo, líquido pleural.	Iniciales y detección temprana en grupos de alto riesgo	Para cáncer originado en laringe, tráquea y bronquios. Baja sensibilidad.
Histología	Biopsia con aguja	En avanzados	Tejido colectado insuficiente o poco representativo
	Broncoscopía	Avanzados	Altamente invasivo. Sensibilidad depende del tamaño del tumor
Quirúrgicos	Mediastinoscopia y mediastinotomía	En avanzados	Incisión de mayor tamaño.
	Biopsia pulmonar quirúrgica	Avanzados	Requiere toracotomía. En presencia de tumor en pared torácica.

### 2.4.1 Biopsia líquida

La biopsia líquida se refiere al procedimiento en el que se analizan derivados de un tumor, como células tumorales circulantes (CTC), ADN libre circulante (ADNlc) y vesículas extracelulares (VE), que pueden ser obtenidas a través de varios fluidos corporales, como la sangre, la orina y líquido de derrame pleural (Souza et al., 2023). Las proteínas y el ADNlc son los biomarcadores mejor estudiados para la biopsia líquida, ilustrados en la Figura 4. El ADNlc está conformado por fragmentos de ADN libre que circulan en la sangre periférica y que tienen una longitud de entre 80 y 200 pares de bases (pb). Los fragmentos de ADN generados por células tumorales conforman el ADN tumoral circulante (ADNtc). La llegada del ADNtc a la circulación periférica puede deberse a la secreción espontánea de fragmentos de ADN a la sangre por parte de los tumores, o como resultado de la apoptosis y necrosis de las células tumorales (Li et al., 2019).

Las ventajas que ofrece este procedimiento son diversas, principalmente debido a su mínima invasión y su potencial para mejorar significativamente el manejo de pacientes con cáncer de pulmón avanzado, ya que permite el monitoreo no invasivo y en tiempo real del progreso de la enfermedad, la respuesta al tratamiento y la recurrencia del cáncer. Debido a que es un procedimiento emergente, actualmente los resultados de la biopsia líquida pueden ser utilizados solamente para apoyar el diagnóstico de cáncer de pulmón obtenido por otros métodos. Sin embargo, la aplicación de la biopsia líquida está siendo estudiada como un ensayo de tamizaje, ya que podría ser una herramienta importante para mejorar la detección temprana de cáncer de pulmón entre las poblaciones de riesgo (W. Zhu et al., 2023).

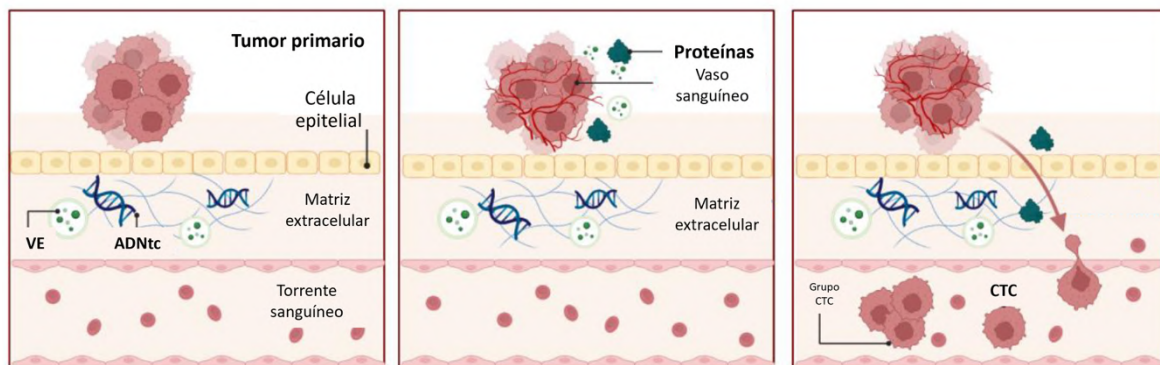


Figura 4. Fundamento de biomarcadores estudiados por biopsia líquida. VE: Vesículas extracelulares. ADNtc: ADN tumoral circulante. CTC: Células tumorales circulantes.

## 2.5 La metilación del ADN en sangre periférica

Los cambios en la metilación global del ADN en células de sangre periférica se asocian con el riesgo de padecer diversos tipos de cáncer, como los de mama, pulmón, próstata, ovario, endometrio, colon y recto (Pei et al., 2020).

Aunque las alteraciones epigenéticas en las células de sangre periférica pueden no representar exactamente los cambios epigenéticos en los tumores primarios, las alteraciones epigenéticas globales pueden reflejar la inestabilidad genómica individual y servir como una firma de exposición integrada de múltiples factores cancerígenos conocidos y desconocidos, reflejando la acumulación de múltiples factores reversibles e irreversibles que pueden incluir la edad, el género, el estilo de vida y la exposición a distintos factores ambientales, así como los polimorfismos genéticos (Kuchiba et al., 2014; Z.-Z. Zhu et al., 2012). Por este motivo, los biomarcadores epigenéticos en células de sangre periférica se están convirtiendo en una nueva herramienta para predecir el riesgo y el pronóstico de diversos tipos de cáncer.

Si bien las células mononucleares de sangre periférica no están incluidas dentro de lo que se consideran las fuentes de biopsia líquida clásica, la detección de su metilación tanto en el diagnóstico temprano como en el pronóstico tumoral ha sido objeto de estudio en años recientes, siendo reportados hasta ahora, cambios en la metilación del ADN de este tipo de células en muestras de cáncer de mama, cáncer de ovario, carcinoma escamoso de cabeza y cuello, relacionándolos de manera directa con la aparición y progresión del cáncer (Y. Li et al., 2023). El fundamento fisiológico detrás de estos cambios en la metilación de células mononucleares de sangre periférica radica en la respuesta que tienen los órganos diana a la estimulación de diferentes factores ambientales, factores fisiológicos y lesiones patológicas, regulando el nivel de expresión de los genes diana, además, los genes relacionados en las células sanguíneas en la circulación periférica también pueden exhibir cambios similares (L. Li et al., 2017).

Los tumores pueden promover cambios genéticos o epigenéticos en leucocitos del sistema circulatorio, igualmente, la metilación del ADN en células de sangre periférica contiene información de metilación relacionada con el tumor. Una ventaja de detectar la metilación del ADN en células mononucleares de sangre periférica es que estas se encuentran en mayor concentración, además, el proceso de extracción es más simple y el ADN es estable durante un tiempo prolongado para su almacenamiento (L. Li et al., 2017; Shirahata & Hibi, 2014).

## **2.6 Espectroscopía infrarroja**

La espectroscopía estudia la interacción de la radiación electromagnética con una sustancia y está basada en el fundamento que explica que la radiación que incide en la materia no es igual a la radiación que sale. Dicha diferencia energética es denominada energía de interacción o energía de absorción. Esta interacción tiene como consecuencia cambios que pueden darse a nivel de átomos o de moléculas. Las formas en las que se pueden dar estas interacciones dependen directamente de las propiedades tanto físicas como químicas de la muestra, así

como de la intensidad de la radiación empleada, ésta última a su vez, determinará el tipo de cambio que se producirá a nivel atómico o molecular (El-Azazy, 2018).

La región infrarroja (IR) de la radiación electromagnética suele dividirse en tres subregiones: infrarrojo cercano (NIR), infrarrojo medio (MIR) e infrarrojo lejano (FIR). La NIR abarca de  $12,500 - 4000 \text{ cm}^{-1}$  (800-2500 nm). Las bandas NIR surgen de transiciones vibracionales de baja probabilidad y por tanto, son muy débiles. Sus coeficientes de absorción suelen ser dos órdenes de magnitud inferiores a los de las bandas fundamentales. La región MIR corresponde a los números de onda  $4000-400 \text{ cm}^{-1}$  (longitudes de onda 2500 - 25.000 nm) y se asocia principalmente con las transiciones vibracionales fundamentales. Esto implica que el coeficiente de absorción alcance niveles elevados en la región MIR. Las bandas de la región del infrarrojo lejano (FIR;  $400-10 \text{ cm}^{-1}$ ; 25.000 nm a 1 mm de longitud de onda) proceden principalmente de las deformaciones esqueléticas y los movimientos de torsión de las moléculas, así como de los modos intermoleculares. Estas bandas fundamentales suelen ser difíciles de analizar, ya que son más débiles y anchas que la mayoría de las bandas MIR (Beć et al., 2022).

Un requerimiento importante para que se lleve a cabo la absorción de la radiación en el espectro infrarrojo por parte de determinados enlaces atómicos de una molécula es la existencia de un momento dipolar de enlace activo. Se llama momento dipolar al producto entre el valor de las cargas ( $q$ ) y la longitud del enlace (Balan et al., 2019). La resonancia es otro de los fenómenos requeridos para la absorción de la radiación infrarroja, es decir, la existencia de una concurrencia entre la frecuencia de vibración basal en el enlace y la frecuencia de la radiación infrarroja incidente que es absorbida. Cuando esto ocurre, el enlace atómico incrementa su nivel de vibración en dicha frecuencia, lo que da como resultado la manifestación de un espectro con la formación de un pico de absorción, dichos picos o secuencia de bandas representan el espectro de absorción en un intervalo de frecuencias en la región del infrarrojo. Es decir, cada pico en el espectro corresponde a la forma de vibración que sucedió en un enlace atómico al interaccionar el haz de infrarrojo

incidente con esa muestra. Dichas vibraciones proceden de enlaces atómicos específicos que conforman las moléculas que contiene el analito. Por otro lado, cuando las frecuencias basal y de absorción no coinciden, en lugar de obtener los picos de absorción antes mencionados, se obtiene una interferencia destructiva y no hay aumento en la vibración del enlace basal en la molécula (Mondragón-Cortez, 2020).

El tipo de vibración de las moléculas es un aspecto complejo, pero está caracterizado en función de la combinación de varios movimientos vibratorios simples que ya están definidos y se denominan modos normales de vibración. Estas vibraciones pueden referirse a cambios longitudinales en los enlaces, lo que sería el modo de estiramiento, o pueden involucrar el cambio del ángulo de los enlaces, es decir, el modo de flexión. Por su parte, el modo de estiramiento puede darse de dos maneras: en fase, llamado estiramiento simétrico, o fuera de fase, entonces denominado estiramiento asimétrico (Mondragón-Cortez, 2017).

Dentro del espectro obtenido, se pueden observar picos de absorción diferenciales, habrá unos con mayor absorbancia, que naturalmente corresponden a los enlaces de las moléculas de la muestra que vibraron a una mayor intensidad al absorber la energía en infrarrojo. La intensidad de los picos en un espectro depende de la cantidad de enlaces que absorban la energía infrarroja a una intensidad determinada, ya que al incidir la energía de infrarrojo sobre la muestra solo los enlaces que entren en resonancia con esta energía podrán absorberla, lo que se asocia con un cambio en el momento dipolar de la molécula, que resulta en una vibración molecular en conjunto. Por lo tanto, diferentes moléculas darán lugar a espectros específicos que permiten su distinción de otras (El-Azazy, 2018).

#### 2.6.1 Espectroscopía infrarroja por transformada de Fourier (FTIR)

Los espectrómetros de infrarrojo por transformada de Fourier (FTIR por sus siglas del inglés *Fourier Transform Infrared Spectroscopy*), constituyen la tercera generación de espectrómetros e incorporan un principio distinto. En estas máquinas

la muestra es irradiada por la fuente de banda ancha y la respuesta específica de la longitud de onda de la muestra se consigue mediante la transformada de Fourier de un interferograma, que acumula la señal de la muestra en toda la región de longitud de onda (de banda ancha) mediante una única acción de barrido; es la denominada medición multiplexada (Zapata et al., 2021).

La transformada de Fourier es en realidad un algoritmo matemático que se integró en la espectrometría infrarroja cuando el monocromador que se utilizaba anteriormente en los equipos se reemplazó por el interferómetro, que permitía registrar simultáneamente varias frecuencias, lo que daba como resultado múltiples ondas sinusoidales superpuestas, el algoritmo de Fourier convierte estas señales en un espectro convencional de infrarrojo, poniendo solución a este inconveniente (Beć et al., 2022).

En el espectrómetro de FTIR se llama accesorio de interacción al dispositivo donde se produce la interacción entre la radiación infrarroja y la muestra a ser analizada. Una de las configuraciones más utilizadas actualmente es con el accesorio de reflectancia total atenuada (de sus siglas en inglés ATR, por *Attenuated Total Reflection*), debido principalmente, a la ventaja que representa al ser compatible con cualquier tipo de presentación física de la muestra. Además, cuando se utiliza el modo de reflectancia total atenuada (ATR), no es necesario tratar la muestra y, por lo tanto, ésta permanece intacta tras el análisis espectroscópico, de forma que puede conservarse para análisis posteriores (Zapata et al., 2021).

#### 2.6.2 Aplicaciones clínicas de la FTIR

En la última década el creciente desarrollo tanto de la instrumentación como de los métodos quimiométricos situó a la técnica analítica cualitativa y cuantitativa por infrarrojo como uno de los métodos más extendidos para determinar las estructuras químicas de las moléculas. Las muestras biológicas son esencialmente complejas, compuestas por una fusión de diferentes proteínas, lípidos, ácidos nucleicos y carbohidratos, por lo tanto, el espectro infrarrojo es el resultado de las bandas de

absorción características de estas moléculas. La FTIR ha ganado un lugar en las herramientas que se utilizan actualmente para investigar material biológico, Balan et al. (2019) resumen sus ventajas:

- Accesibilidad experimental a una serie de transiciones activas infrarrojas derivadas de moléculas específicas en regiones espacialmente localizadas dentro de las biomoléculas.
- Método no invasivo que no requiere marcadores ni sondas fluorescentes.
- Sin límite de peso molecular de la muestra.
- Permite obtener imágenes instantáneas de todas las conformaciones moleculares.
- Mínima preparación de la muestra.
- Sencillez, rapidez y bajo costo.
- Alta sensibilidad molecular unida a una resolución espacial de hasta unos pocos micrómetros.

Recientemente se ha puesto atención en la importancia de la técnica de infrarrojo en la evaluación sensible y favorable de todas las “ómicas” de una muestra biológica, que permite la detección de cambios moleculares que pueden reflejar un diagnóstico precoz y un pronóstico eficaz de la enfermedad. La espectroscopía infrarroja, por lo tanto, se ha incorporado en medicina para el diagnóstico de diversos cánceres, trastornos neurológicos, diabetes, arterosclerosis, glóbulos rojos infectados por la malaria y seguimiento de la artritis reumatoide en modelos celulares, animales o clínicos (Balan et al., 2019; Ghosal et al., 2008).

El análisis de suero sanguíneo y en especial, el análisis de las proteínas mediante métodos espectroscópicos vibracionales (espectroscopía de Raman y FTIR), pueden proporcionar información esencial sobre el estado del paciente e indicar la existencia de diferentes patologías (Figura 5). Los espectros del suero sanguíneo permiten identificar las especies orgánicas más abundantes como proteínas totales, glucosa, triglicéridos, urea, albúmina y colesterol. El suero sanguíneo incluye cerca



de 20,000 proteínas distintivas con una concentración total de 1mM. El peptidoma, conocido como la fracción de bajo peso molecular del suero, está presente en concentraciones pequeñas, sin embargo, tiene un potencial de datos diagnósticos específicos del cáncer debido principalmente a la capacidad de huella dactilar de los eventos moleculares de diferentes órganos o tejidos asociados a la presencia de cáncer (Balan et al., 2019; Yang et al., 2021a).

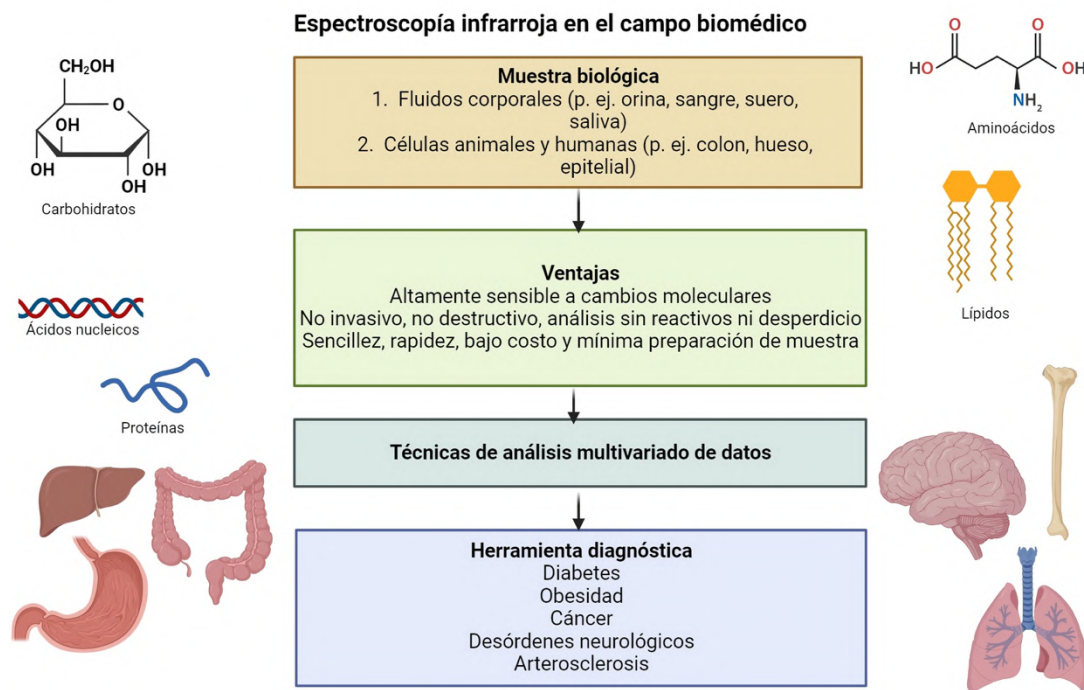


Figura 5. Aplicaciones biomédicas de la espectroscopía infrarroja.

### 2.6.3 Regiones y bandas de interés en el estudio del cáncer de pulmón

La región conocida como la “huella digital” (*bio-fingerprint region*) del espectro del infrarrojo medio comprende de  $1800$  a  $900\text{ cm}^{-1}$  y contiene los modos vibracionales característicos de los enlaces químicos que pueden ser explorados para comprender mecanismos intracelulares, por lo tanto, la espectroscopía infrarroja proporciona una “huella dactilar bioquímica celular” de la muestra que se está estudiando, con una asociación directa entre picos y enlaces químicos (Yang et al., 2021a). Esto representa un enfoque no destructivo y de tamizaje para el diagnóstico que puede ser realizado de forma rápida.

En la obtención de espectros FTIR de sueros de pacientes con cáncer de pulmón y de un grupo de personas sanas, Wang y colaboradores caracterizaron las bandas preliminares que se resumen en el Cuadro 3 y que posteriormente se asignan en el Cuadro 4, reportando que la razón A1080/A1170 puede discriminar entre los sueros de ambos grupos, siendo 1080  $\text{cm}^{-1}$  correspondiente al estiramiento simétrico de PO2 de ácidos nucleicos y 1170  $\text{cm}^{-1}$  a los grupos C-O (H) de los residuos de treonina, tirosina y serina de proteínas, y que esto se relaciona con que el suero de pacientes con cáncer de pulmón contiene mayor cantidad de ácidos nucleicos (Wang et al., 2014).

Cuadro 3. Asignaciones preliminares de los espectros IR medios del suero de pacientes con CP (Wang et al., 2014).

Frecuencia ( $\text{cm}^{-1}$ )		Asignación de banda
Pacientes con cáncer de pulmón	Personas sanas	
3293	3295	Banda amida A
2958 y 2872	2959 y 2872	Estiramiento simétrico y asimétrico CH3
2931	2931	Estiramiento simétrico y asimétrico CH2
1743	1741	Estiramiento C=2
1650 y 1546	1651 y 1545	Banda de la amida I y II
1451	1451	Deformación CH2
1399	1399	Doblamiento CH3
1313	1314	Banda de la amida III
1243 y 1079	1243 y 1080	Estiramiento simétrico y asimétrico PO2
1169	1170	Estiramiento C-O(H) en proteínas

Cuadro 4. Asignación de bandas por (Wang et al., 2014) en suero de pacientes con CP.

<b>Banda/región (cm<sup>-1</sup>)</b>	<b>Asignación</b>	<b>Relación cáncer de pulmón</b>
2997-2887	Estiramiento C-H (lípidos).	El contenido de lípidos incrementa en relación con el contenido de proteínas en pacientes con CP.
1593-1480	Banda amida II (proteínas).	
1650	Banda amida I (proteínas).	Sugiere que el contenido de flexión N-H y estiramiento C-N aumentó en relación con el contenido de estiramiento carbonilo en las proteínas del suero de pacientes con CP.
1080	Estiramiento simétrico PO <sub>2</sub> de ácidos nucleicos.	Sugiere que el contenido de ADN está incrementado en CP, relacionado probablemente con necrosis y apoptosis de células de CP y/o liberación de ADN por éstas.
1243	Estiramiento asimétrico de ácidos nucleicos.	Puede indicar cambios estructurales en ácidos nucleicos.
1170	Grupos C-O (H) de residuos de treonina, tirosina y serina en proteínas	Puede describir el incremento relativo de ácidos nucleicos en CP.

En el análisis FTIR de esputo de pacientes con cáncer de pulmón y sanos se reportaron las siguientes bandas resumidas en el Cuadro 5, con una absorbancia incrementada en el grupo de cáncer de pulmón: 964 cm<sup>-1</sup>, resultante de los modos vibracionales de ácidos nucleicos y del estiramiento simétrico de monoésteres de

fosfato y proteínas fosforiladas, así como picos altamente diferenciados en  $1024\text{ cm}^{-1}$  y  $1049\text{ cm}^{-1}$ , correspondientes al estiramiento C-O y la flexión C-O típica del glucógeno; posiblemente relacionado con mayores niveles de glucógeno característicos de los tumores pulmonares. También se obtuvo un pico en  $1411\text{ cm}^{-1}$ , asociada con el estiramiento COO- y la flexión C-H. Las bandas de  $1577\text{ cm}^{-1}$  y  $1656\text{ cm}^{-1}$ , correspondientes a las regiones amida II y amida I, respectivamente, también presentaron absorbancias mayores en el grupo de cáncer de pulmón (Lewis et al., 2010).

Cuadro 5. Asignación de bandas en esputo por (Lewis et al., 2010).

Posición del pico obtenido por la segunda derivada ( $\text{cm}^{-1}$ )		Modo vibracional propuesto	Fuente primaria propuesta
Normal	Cáncer de pulmón		
964	966	Estiramiento PO <sub>4</sub> , Estiramiento C-C	Proteínas y ácidos nucleicos
1024	1024	Estiramiento C-O, flexión C-O	Glucógeno
1049	1051	Estiramiento C-O, flexión C-O	Glucógeno
1411	1417	Estiramiento COO-, flexión C-H	Proteína
1577	1577	Amida II, flexión NH, estiramiento C-N, estiramiento C=N del anillo de imidazol	Proteínas y ácido nucleico
1656	1654	Amida I, estiramiento C=O	Proteína

El proceso de diseminación de las células cancerosas representa una etapa fundamental y que ocurre como resultado del éxito de la intravasación y la supervivencia en el torrente sanguíneo de estas células. La metástasis, por lo tanto, requiere un incremento en la producción de energía, para lo cual se llevan a cabo una serie de cambios metabólicos que involucran la regulación de la glucosa, los lípidos y proteínas (Weber, 2016). La espectroscopía FTIR permite analizar estos cambios a través de la longitud de las cadenas lipídicas, la conformación de las proteínas y ácidos nucleicos, así como la relación lípido/proteína. Por ejemplo, la alteración en el metabolismo de los lípidos puede notarse en las intensidades de la banda  $2926\text{ cm}^{-1}$  correspondiente al estiramiento asimétrico  $\text{CH}_2$ , y  $2852\text{ cm}^{-1}$  del estiramiento simétrico  $\text{CH}_2$ , que se asocian a fosfolípidos y ácidos grasos de cadena larga, y que puede relacionarse con la actividad de la fosfolipasa D en la hidrólisis de los fosfolípidos que participan en la invasión de las células cancerosas (Blat et al., 2019; Staniszevska-Slezak et al., 2018).

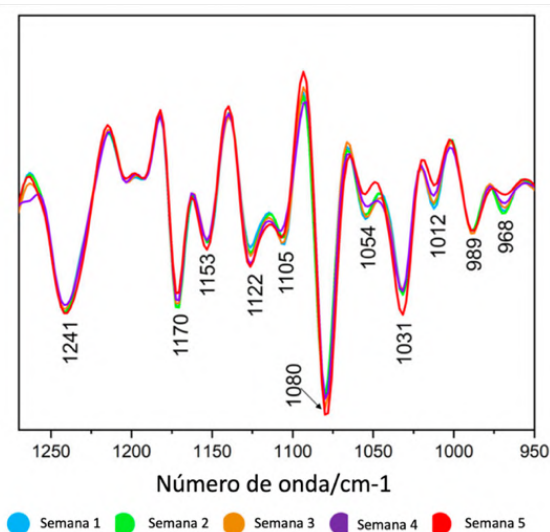


Figura 6. Espectros FTIR de plasma de las semanas 1-5 de progresión del cáncer de mama. Modificado de (Blat et al. 2019).

En un análisis de plasma de pacientes con cáncer de mama (Figura 6) y metástasis a pulmón se encontraron bandas en el rango espectral  $1300 - 950\text{ cm}^{-1}$ , asignados a las vibraciones C-O de glucosa, glicerol y ARN, asociadas probablemente a

diferentes vías metabólicas de azúcares y ácidos nucleicos que se incrementan como parte de los requerimientos metabólicos durante las distintas fases del desarrollo del cáncer (Blat et al., 2019).

#### 2.6.4 Las conformaciones del ADN en la espectroscopía FTIR

La facilidad con que la espectroscopía FTIR proporciona una instantánea de la estructura y conformación del ADN en forma de picos tanto intensos como discernibles, ha permitido utilizar esta herramienta ampliamente para estudiar sus conformaciones, por consiguiente, se ha logrado recopilar los diversos estudios y de esta manera integrar una biblioteca de picos de ADN en la región  $1800 - 800 \text{ cm}^{-1}$ , que es la región de interés cuando se estudian ácidos nucleicos (Gurbanov et al., 2019).

El ADN bicatenario puede presentarse en tres conformaciones: la forma B, la forma A y la forma Z, siendo la forma B la que se encuentra con mayor frecuencia en la célula, mientras que la forma A es la segunda más frecuente. Existe un proceso transicional de ADN-B a ADN-A que se lleva a cabo en el contexto de la replicación del ADN o del reconocimiento de secuencias específicas por proteínas o la interacción con ligandos, que se basan en interacciones ADN-proteína (Gurbanov et al., 2019; Whelan et al., 2011). La conformación ADN-Z es una hélice hacia la izquierda, que presenta un surco menor más estrecho que el ADN-B y ningún surco mayor evidente, su estabilización es posible gracias al superenrollamiento negativo generado por la transcripción, como un cambio conformacional temporal (Zhang et al., 2016). En el espectro FTIR, la banda del estiramiento asimétrico  $\text{PO}_2$  proporciona información sobre la conformación del esqueleto de los ácidos nucleicos, y está reportada alrededor de  $1245 - 1235 \text{ cm}^{-1}$  para la forma A,  $1225 - 1220 \text{ cm}^{-1}$  para la forma B y  $1216 - 1213 \text{ cm}^{-1}$  para la forma Z del ADN (Banyay et al., 2003; Blat et al., 2019; Gurbanov et al., 2019; Zhang et al., 2016).

Información adicional sobre la conformación del ADN puede obtenerse con las bandas que se asocian a los modos de plegamiento de los azúcares. El plegamiento

del azúcar se refiere a la torsión que sufren las moléculas de pentosa en el ADN, los anillos de azúcar constituyen el enlace flexible entre la base nitrogenada y el esqueleto de fosfato, por lo que al adoptar diferentes conformaciones de plegamiento alteran la orientación relativa del esqueleto y la base, resultando en variaciones en la estructura global de los ácidos nucleicos, así como en su función (Huang et al., 2014). La región  $900 - 800 \text{ cm}^{-1}$  puede mostrar información relacionada, ya que se asocia a las vibraciones del azúcar/azúcar-fosfato, en específico, las bandas  $882 - 877$ ,  $865 - 860$  y  $815-802 \text{ cm}^{-1}$  corresponden al modo de plegamiento de azúcares de tipo N (plegamiento C3'-endo), y las bandas del rango  $842 - 820 \text{ cm}^{-1}$  se asocian a los azúcares de tipo S (plegamiento C2'-endo) (Gurbanov et al., 2019).

#### Conformación Z del ADN

El ADN-Z es una doble hélice levógira, estrecha y alargada, con un diámetro de  $44,6 \text{ Å}$ , que incorpora 12 bases por vuelta en oposición a las 10 del ADN-B. La cadena de azúcar-fosfato forma un zigzag causado por orientaciones alternas de los azúcares ribosa en los nucleótidos de guanosina y citidina. En la conformación Z, los residuos de guanosina y citidina tienen el oxígeno O4' de furanosa "apuntando" en direcciones opuestas a lo largo de la cadena principal de fosfato (Krall et al., 2023).

Las modificaciones covalentes de las bases nitrogenadas del ARN y del ADN afectan la estabilidad de la conformación Z desestabilizando las conformaciones A y B o estabilizando la conformación Z. Se conoce que cambios epigenéticos como la metilación del ADN facilitan la transición de formas B-Z (Figura 7), y aunque el mecanismo no está completamente elucidado, se cree que estos cambios conformacionales inducidos por la metilación están relacionados con menos diferencias energéticas que implica la transición del ADN-B metilado al ADN-Z, así como con su función biológica (Temiz et al., 2012). Esto sugiere una barrera energética más baja, que puede ayudar a explicar una mayor facilidad del cambio de ADN-B metilado a ADN-Z que de ADN-B no metilado (Temiz et al., 2012).

Además, se ha caracterizado que para el ADN-Z, la metilación en C5 de citidina desestabiliza la conformación B al hacer que el surco mayor sea más hidrófobo mientras que solo desestabiliza parcialmente la conformación Z, ya que el grupo metilo llena parcialmente un parche hidrófobo en la superficie del surco mayor, resultando en la promoción de la conformación Z (Krall et al., 2023).

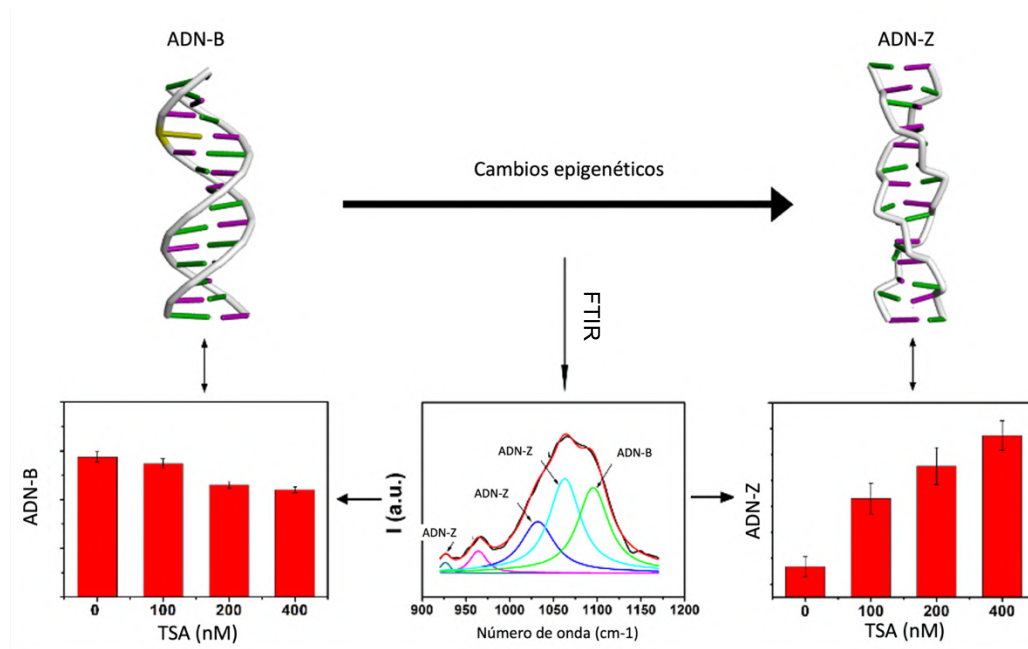


Figura 7. Espectroscopía FTIR para monitorear las señales de ADN en células con acetilación de histonas regulada por tricostatina A (TSA). Modificado de (Zhang et al. 2016).

También se ha reportado que secuencias metiladas analizadas con espectroscopía FTIR han presentado incrementos en las bandas correspondientes a los azúcares de tipo S (plegamiento C2'-endo), como lo demostrado por Gurbanov et al. en 2019, al comparar datos IR con un análisis de metilación de secuencias de ADN para evaluar la capacidad de aclimatación a metales pesados por mecanismos epigenéticos de *Gordonia sp.*

Por otro lado, la adición de un grupo metilo en el C5' de la citosina ocasionará vibraciones CH<sub>3</sub>, y ya que un cambio en el entorno químico puede dar lugar a desplazamientos y/o intensidad de los picos, también ocurrirán cambios en los picos



asociados con el grupo NH<sub>2</sub> (del C4 adyacente a la posición metilada de la citosina), así como cambios en los estiramientos C-C. En una comparación entre la espectroscopía FTIR y la espectroscopía de Raman para analizar la sensibilidad de ambas técnicas para detectar cambios en oligonucleótidos diferencialmente metilados y no metilados, se reportaron las siguientes bandas: 1548 cm<sup>-1</sup> de la flexión N-H, 1612 cm<sup>-1</sup> del estiramiento C4-NH<sub>2</sub>, 1658 cm<sup>-1</sup> del estiramiento C=O, 818 cm<sup>-1</sup> del esqueleto de fosfato, 822 cm<sup>-1</sup> de la estructura de la citosina y 1608 cm<sup>-1</sup> de la vibración C4-NH<sub>2</sub> de la citosina (Kelly et al., 2011).

De tal manera, la variabilidad en estas regiones previamente mencionadas puede caracterizar muestras de dos poblaciones o grupos distintos, y su representación gráfica espectral correspondería a picos definidos, como se muestra en el ejemplo de la Figura 8.

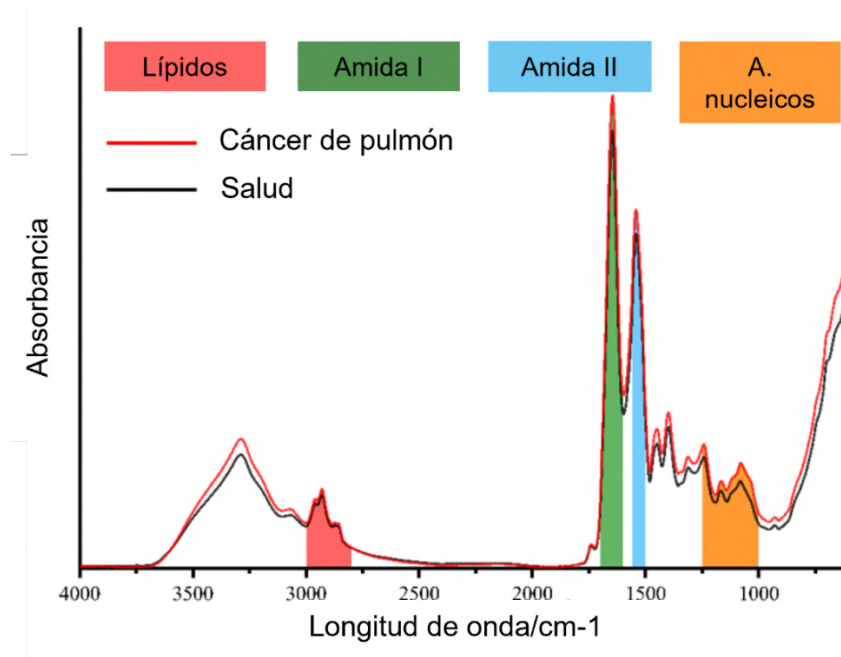


Figura 8. Espectros IR medios del suero de pacientes con cáncer de pulmón y de personas sanas. Modificado de (Yang et al., 2021).

### **3 Hipótesis**

Los datos obtenidos por espectroscopía FTIR de las muestras séricas correlacionarán con el grado de metilación global obtenido por ELISA en pacientes con cáncer de pulmón.

## **4 Objetivos**

### **4.1 General**

Explorar la correlación de datos de infrarrojo de muestras séricas con datos espectrofotométricos del grado de metilación del ADN en pacientes con y sin cáncer de pulmón.

### **4.2 Particulares**

1. Determinar el grado de metilación global del ADN en pacientes con y sin cáncer de pulmón mediante ELISA.
2. Establecer las variables de los espectros de absorbancia de infrarrojo de las muestras estudiadas que se someterán al análisis de correlación con el porcentaje de metilación global del ADN.
3. Evaluar la posible aplicación de los datos de infrarrojo para la determinación del porcentaje de metilación global del ADN en pacientes con cáncer de pulmón empleando técnicas de correlación y medidas de similitud.

### **4.3 Profesionalizantes**

- Adquirir habilidades y conocimientos en la aplicación de técnicas analíticas avanzadas para el diagnóstico de enfermedades, como la espectroscopía infrarroja por transformada de Fourier.
- Desarrollar competencias en el diseño y la implementación de protocolos de análisis clínico para la detección y el seguimiento de enfermedades como el cáncer de pulmón.

## 5 Metodología

### 5.1 Diseño

El diseño de este proyecto de investigación es de tipo observacional transversal y comparativo.

#### 5.1.1 Definición del Universo

Individuos de 40 años en adelante diagnosticados con cáncer de pulmón en instituciones de salud (ISSSTE, IMSS) de México.

Individuos de 40 años en adelante que residan en México y que no tengan historial de ningún tipo de cáncer, sin familiares consanguíneos de primer grado diagnosticados con cáncer y sin antecedentes de enfermedades pulmonares.

#### 5.1.2 Tamaño de muestra

El tamaño muestral se obtuvo por la fórmula de medias de dos muestras independientes:

$$n = \frac{(Z_{\alpha} + Z_{\beta})^2 (Var1 + Var2)}{(\mu1 - \mu2)^2} \quad (1)$$

Donde:

$Z_{\alpha}$  es el nivel de confianza (95 % = 1.64)

$Z_{\beta}$  es el poder de la prueba (80 % = 0.84)

Var 1 y Var 2 son las varianzas de cada muestra, respectivamente.

$\mu1$  y  $\mu2$  son las medias de cada muestra, respectivamente.

Dado que no existen antecedentes relacionados a la metilación global del ADN que se ajusten a las características de este estudio se consideró solo la variable de absorbancia obtenida por espectroscopía FTIR para la estimación del tamaño de muestra.

De acuerdo con lo reportado por (Wang et al., 2014), en muestras de suero de pacientes con cáncer de pulmón, para las bandas espectrales A2959/A1545 se

encontró una media de absorción de  $0.2737 \pm 0.0461$ , mientras que la media para el grupo de personas sanas fue de  $0.2451 \pm 0.0224$ . Se tomó como referencia esa razón de absorción ya que ha sido la mejor variable de respuesta reportada en la literatura que represente la mínima diferencia estadísticamente significativa entre ambos grupos.

Sustituyendo tales valores en la fórmula (1) y considerando las posibles pérdidas, se conforman los grupos de la siguiente forma:

Grupo de pacientes diagnosticados con cáncer: 20

Grupo control: 20

### 5.1.3 Definición de las unidades de observación

De acuerdo con el Cuadro 6, se definen los grupos a evaluar.

Cuadro 6. Definición de unidades de observación.

Unidades de observación	Definición	Unidad de medida
<b>Pacientes con cáncer de pulmón</b>	Personas que hayan sido diagnosticadas con cáncer de pulmón	Porcentaje de metilación global del ADN Unidades de absorbancia
<b>Grupo control</b>	Individuos sin cáncer, que hayan residido en México al menos el 51 % de su vida, sin familiares consanguíneos de primer grado con cáncer, sin antecedentes de enfermedades pulmonares.	Porcentaje de metilación global del ADN Unidades de absorbancia

### 5.1.4 Grupo control

El grupo control se conformará por personas mayores de 40 años sin discriminación de sexo.

#### 5.1.5 Criterios de inclusión

- a) Para ambos grupos los integrantes debían tener de 40 años en adelante, que hubiesen residido en el país la mayor parte de su vida (mínimo el 51 % de su vida).
- b) Para el grupo de pacientes con cáncer de pulmón: Que contaran con diagnóstico de cualquier estadio de cáncer de pulmón por primera vez. Tumor primario originado en pulmón.

#### 5.1.6 Criterios de exclusión

- a) Ningún participante del grupo control debió tener exposición pasiva por más de una hora semanal a humo de tabaco, ni exposición mayor a 100 horas/año a humo de leña, humo de fábricas y chimeneas.
- b) Participantes que previamente hubiesen padecido cualquier tipo de cáncer.
- c) Participantes con familiares consanguíneos de primer grado que hubiesen sido diagnosticados con cáncer de pulmón.
- d) Participantes diagnosticados con asma, fibrosis pulmonar, o cualquier otra enfermedad pulmonar crónica.
- e) Para el grupo de pacientes con cáncer de pulmón: personas cuyo cáncer de pulmón fuese producto de la metástasis de otro tipo de cáncer.

#### 5.1.7 Criterios de eliminación

- a) Participantes que ya no desearan continuar en el estudio.
- b) Participantes cuya muestra haya sido insuficiente o cuyos datos estén incompletos y no se pueda repetir la toma de muestra o información.

#### 5.1.8 Definición de variables y unidades de medida

De acuerdo con el Cuadro 7, se mencionan tanto la definición, las unidades de medida y el tipo de variable.

Cuadro 7. Características de las variables.

<b>Variable</b>	<b>Definición</b>	<b>Unidad de medida</b>	<b>Tipo de variable</b>
Perfiles infrarrojos	Espectro de absorción que a través de bandas o picos representa las frecuencias en las que la muestra absorbe la radiación	Unidades de absorbancia Número de onda en cm-1	Cuantitativa continua
Metilación global del ADN	Porcentaje de ADN con 5-metilcitosina	%	Cuantitativa continua

#### 5.1.9 Fuentes, métodos, técnicas y procedimientos de recolección de la información.

1. La búsqueda de los pacientes con cáncer de pulmón se realizó mediante una colaboración con el Cancer Center del Hospital Tec 100 Querétaro. También se promovió la participación mediante un aviso de reclutamiento para ambos grupos de estudio que se colocó en instituciones de salud con previa autorización, así como en diversos puntos de la ciudad y de la Universidad Autónoma de Querétaro.
2. A los participantes que cumplían con los criterios de inclusión se les pidió leer la carta de consentimiento informado, se les hizo saber que cualquier duda que tuvieran podían consultarla con las investigadoras responsables. Una vez que terminaron de leer, se les preguntó si habían comprendido en su totalidad la carta de consentimiento informado o si tenían alguna otra duda, si todo había quedado comprendido se procedió a solicitar su firma (Anexo 1, Anexo 4). El Anexo 4, correspondiente al consentimiento informado amplio, se incluye con el propósito de almacenar las muestras de suero y sangre total a -20 °C por un plazo máximo de 2 años en un espacio reservado del laboratorio de la Unidad

de Servicios Clínicos y Diagnóstico Molecular ubicado en el Edificio 9 de la Facultad de Química, Centro Universitario. Esto con el fin de obtener la autorización de los participantes para utilizar las muestras y los datos obtenidos de ellas en proyectos de investigación sobre cáncer de pulmón que se realicen posteriormente en el Laboratorio de Enfermedades Crónico-Degenerativas como parte del mismo programa académico y que sean dirigidas por la Dra. Karla Isabel Lira de León.

3. Se comunicó a los participantes la existencia del formato de revocación del consentimiento (Anexo 2). Si los participantes deciden retirarse del estudio se les preguntará si desean la destrucción de la muestra recolectada y la eliminación de sus datos o si permiten el uso de la muestra y los datos ya recabados. Ambas opciones se presentan en el formato. Se le preguntará al participante si tiene alguna duda respecto a su retiro del estudio, si todo quedó comprendido se solicitará que seleccione cómo proceder con sus datos y muestra sanguínea y que firme la revocación del consentimiento. Se respetará la elección del participante en relación con la conservación o destrucción de su muestra y/o datos recolectados.
4. Se registraron los datos personales de cada participante, los cuales fueron tratados de acuerdo con los principios éticos correspondientes, y se les proporcionó un cuestionario para su llenado (Anexo 3).
5. La toma de muestra sanguínea se llevó a cabo en el laboratorio de análisis clínico del Cancer center del Hospital Tec 100 de Querétaro para el grupo de cáncer de pulmón. La toma de muestra sanguínea para el grupo control fue realizada por la Q.F.B. Alitzel Arcos de la Cruz, con cédula profesional 13357730, y la aplicación de los formularios fue realizada por la Lic. María Arellano Sosa con cédula profesional 13057599, responsable del proyecto. Ante cualquier incidente con el participante y en caso de que fuese necesario, estaba disponible la atención médica por parte de los médicos del consultorio clínico universitario de FARMA-UAQ: Dra. Irhe Sue Acevedo Baltazar y Dra. Karen Itzel Suárez Torrez.



## **5.2 Procedimientos**

### **5.2.1 Obtención de muestra sanguínea**

Primero se realizó la extracción de sangre periférica mediante punción venosa utilizando el sistema Vacutainer™. Se obtuvieron 5 ml de sangre con tubos de recolección lila con EDTA y 4 ml con tubo dorado sin anticoagulante (9 ml de sangre en total) en participantes con previo ayuno de 8 horas mínimo. El procedimiento fue ambulatorio, siguiendo todas las medidas de higiene y seguridad recomendadas para la extracción sanguínea, implicando riesgo de formación de hematoma, dolor y/o malestar en el área de punción, reacción alérgica, mareo o desmayo e infección del área de punción si se expone a microorganismos.

La extracción de sangre periférica se realizó de la siguiente manera: una vez que la persona estuvo sentada, se despejó el área antecubital del brazo, se colocó el torniquete con la mano del participante cerrada para palpar la vena, se exploró primero la vena cubital, que es la más larga, gruesa y se prefiere por bordear la musculatura del brazo, posteriormente se buscó la vena cefálica, por ser menos gruesa, y por última opción, de ser necesario se buscó la vena basilica, ya que es más pequeña que las anteriores y se encuentra en un área más sensible (González de Buitrago, 2010). Una vez seleccionada la vena, se descontaminó el área de punción con una torunda de alcohol al 70 % y se introdujo la aguja con el sistema Vacutainer en un ángulo de 45°, se acoplaron los tubos en orden, primero el tubo sin anticoagulante y posteriormente el tubo con anticoagulante. Los tubos se invirtieron inmediatamente después de quitarlos del sistema. Se aseguró que la persona tuviese abierta la mano, se liberó el torniquete y se colocó sin presionar una torunda sin alcohol sobre la aguja para retirarla. Una vez retirada se presionó el área puncionada con la torunda y se colocó una banda adhesiva para evitar contaminación. Las muestras de sangre fueron procesadas en un lapso no mayor a 3 horas, se obtuvo primero el suero y las alícuotas de sangre total para su posterior almacenamiento (descrito en sección 5.2.2).

### 5.2.2 Obtención de suero

Para obtener el suero, los tubos con gel separador fueron centrifugados a 3,200 rpm durante 15 minutos y se recuperó el sobrenadante. Se hicieron alícuotas de 1 ml y se les añadió 0.6 g de trehalosa de la marca Sigma (T9631-10G) para preservar las proteínas. Las muestras fueron almacenadas en congelación a -4 °C en el Laboratorio de Microbiología Molecular del Edificio 9 de la Facultad de Química hasta su análisis. Las muestras de sangre total fueron almacenadas en refrigeración en el tubo primario en el mismo laboratorio.

### 5.2.3 Determinación del porcentaje de metilación global del ADN

Para cumplir el primer objetivo particular propuesto, se realizarán los siguientes pasos.

#### 5.2.3.1 Extracción de ADN

La extracción del ADN se realizó a partir de la sangre total, utilizando el estuche *Quick-DNA™ Miniprep Plus Kit*, número de catálogo D4068 de la casa comercial *Zymo Research* (Irvine, California, Estados Unidos), que se basa en el principio de unión del ADN a una membrana de gel de sílice, mientras que los contaminantes la atraviesan. Para comenzar los pasos del procedimiento indicado por el proveedor se colocó una cantidad de 200 µL de muestra de sangre total en tubos estériles de 1.5 mL para microcentrifuga. Posteriormente se emplearon los buffers correspondientes para inducir la lisis de las células presentes en las muestras de sangre, la unión del ADN genómico a la membrana de una columna de centrifugación, el lavado de la membrana y finalmente la elución del ADN genómico de la membrana, de acuerdo con las instrucciones y las modificaciones necesarias indicadas por el proveedor.

#### 5.2.3.2 Cuantificación de la concentración y pureza del ADN

Se obtuvo la concentración del ADN mediante espectrofotometría, mediante el valor de absorbancia obtenido a una longitud de onda de 260 nm, mientras que la pureza

del ADN se evaluó con la lectura de las absorbancias en la relación A260/280. Para estas cuantificaciones se utilizó un espectrómetro *Thermo™ Scientific Multiskan GO Microplate Spectrophotometer*. La absorbancia del ADN a 280 nm es característica de las proteínas, especialmente las que contienen aminoácidos aromáticos (triptófano, tirosina y fenilalanina). Una muestra de ADN con relación de absorbancia 260/280 cercano a 1.8 se considera pura, mientras que <1.7 indica posible contaminación por proteínas o fenoles, lo que significa que hay una mayor absorbancia a 280 nm. La relación 260/230 se utiliza para evaluar la pureza respecto a contaminantes orgánicos como fenoles, sales, azúcares o reactivos utilizados en la extracción del ADN. A 230 nm absorben sales (EDTA) y algunos carbohidratos. Una muestra de ADN con una relación 260/230 entre 2.0 – 2.2 se considera de alta pureza, mientras que valores < 2.0 indican posible contaminación por compuestos orgánicos, sales o reactivos que absorben a 230 nm (Desjardins & Conklin, 2010). Los valores obtenidos fueron en general de una pureza aceptable.

#### 5.2.3.3 Evaluación de la integridad del ADN

Para conocer la integridad del ADN extraído se realizó una electroforesis en gel de agarosa al 0.85 %, para lo cual se pesaron 0.25 g de agarosa y se disolvieron en un matraz con 30 ml de buffer TAE (TRIS, ácido acético, EDTA) 1X como buffer de corrida, el matraz se calentó en el horno de microondas durante 40 segundos, observando que la solución quedara completamente transparente, lo que indica que la agarosa se ha disuelto. Se añadieron 3 µl de bromuro de etidio para visualizar las bandas, se dejó enfriar y se vertió en la bandeja de la cámara de electroforesis con precaución para evitar la formación de burbujas, se insertó el peine y se dejó gelificar durante aproximadamente 20 minutos. Se utilizaron 2 µl de 6X buffer de carga mezclado con 5 µl de muestra. En el primer pozo se cargaron 5 µl de marcador de peso molecular 1 Kb Plus DNA Ladder de *Invitrogen™* y posteriormente cada una de las muestras. Las condiciones de corrida fueron: voltaje constante de 100 volts durante 30 minutos. Las bandas fueron analizadas utilizando un sistema de foto documentación UV. En la Figura 9 se presenta el corrimiento electroforético del ADN de 9 muestras escogidas al azar.

Se considera que el ADN de una muestra está íntegro cuando su perfil en una electroforesis en gel de agarosa se corresponde a una banda discreta. El nivel de degradación de una muestra está determinado por la pérdida de definición de la banda predominante y el acompañamiento de una estela a lo largo del gel. De acuerdo con el *Joint Genome Institute (JGI)*, cuando se observa una presencia simultánea de la banda en la parte superior del gel y una ligera estela, como en la Figura 9, se considera un ADN de integridad adecuada (Departamento de control de calidad del Banco Nacional de ADN, 2024).

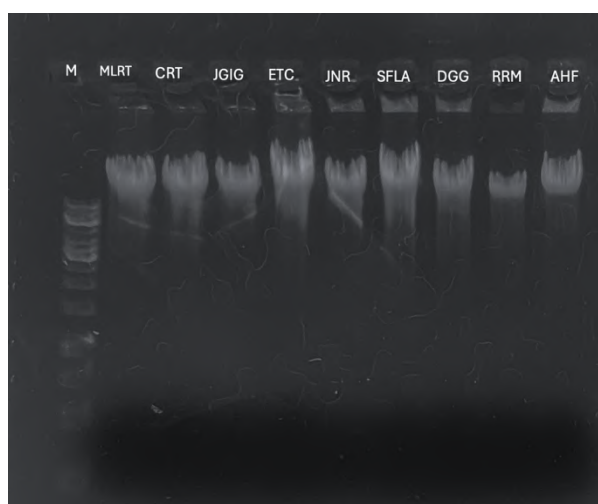


Figura 9. Corrimiento electroforético del ADN genómico extraído. Las iniciales representan los nombres de las muestras cargadas. M: Marcador de peso molecular 1 Kb.

#### 5.2.3.4 Ensayo por inmunoabsorción ligado a enzima (ELISA)

Para cuantificar el porcentaje global de metilación del ADN se utilizó el estuche *MethylFlash™ Methylated DNA 5-mC Quantification Kit (Colorimetric)* de la casa comercial *EpiGentek* (Farmingdale, Nueva York, Estados Unidos), que se basa en el principio de captura por anticuerpos del ADN metilado. Los pasos para llevar a cabo el procedimiento se siguieron de acuerdo con las indicaciones del proveedor (EpigenTek, 2022). Por recomendación del proveedor, se utilizó una cantidad de 100 ng de ADN, por lo que el ADN fue diluido dependiendo la concentración de cada

muestra, añadiéndose un volumen de 4 µl por cada muestra. La curva estándar se realizó mediante 6 diluciones del control positivo que incluía el estuche, también se incluyó el control negativo. Tanto los controles, las diluciones para la curva y las muestras, fueron cargadas por duplicado en los pocillos. El procedimiento realizado fue el siguiente:

1. Antes de comenzar la metodología, se verificó la disponibilidad de todo el material requerido, así como la preparación de las soluciones stock que estaban concentradas de acuerdo con las instrucciones del estuche.
2. Preparación de la curva estándar. Se diluyó 1 µl del control positivo (PC) con 9 µl del control negativo (NC) para obtener PC diluido en un tubo de 1.5 ml estéril y previamente rotulado. Posteriormente se prepararon los 6 puntos de concentraciones para el control combinado, PC diluido y NC de acuerdo con las proporciones indicadas por el estuche.
3. Unión del ADN. Se seleccionaron las tiras a utilizar y se dibujó el esquema de la microplaca con el orden elegido de las muestras. Para los pozos control negativo (NC) se añadió 100 µl de *Binding solution* (BS) y 2 µl del control negativo (NC). Para pozos de control positivo (PC) se añadió 100 µl de BS y 2 µl de PC a diferentes concentraciones (0.1 % - 5 %) para generar la curva estándar. Para pozos de muestra se añadió 100 µl de BS y 100 ng de muestra de ADN. Por sugerencia del estuche, el volumen de ADN a añadir debía estar en un rango de 2 – 4 µl, por lo que se eligió un volumen de 4 µl ya que con esta cantidad se facilitaba el pipeteo y se reducía la probabilidad de cometer errores al cargar un volumen tan bajo. El ADN se diluyó en solución salina fisiológica (NaCl 0.9 %) de la marca PISA. Para calcular el volumen de ADN a utilizar por muestra para las diluciones, se realizó la sustitución de la concentración de cada muestra en la fórmula  $\text{Volumen } (\mu\text{l}) = \text{cantidad deseada (ng)} / \text{concentración (ng}/\mu\text{l})$ .
4. Se mezcló la solución inclinando la placa suavemente, se cubrió con papel Parafilm™ y se incubó a temperatura ambiente en un agitador de placas durante

60 minutos. Durante los últimos 10 minutos de incubación, se preparó la solución “Complejo de detección”.

5. Se removió el BS de cada pocillo después de 60 min de incubación mediante un movimiento de inversión de la placa. Se lavó cada pozo con 150  $\mu$ l de *Washing buffer (WB)* cada vez durante 3 repeticiones, empleando una micropipeta multicanal.
6. Detección del ADN metilado y medición de la señal. Se añadieron 50  $\mu$ l de la solución Complejo de detección a cada pozo. La placa se cubrió con papel Parafilm™ y se incubó a temperatura ambiente en el agitador de placas durante 50 minutos. Se removió la solución Complejo de detección de cada pozo mediante inversión. Se lavó cada pozo con 150  $\mu$ l de WB cada vez por 5 repeticiones.
7. Se añadieron 100  $\mu$ l de *Detection solution (DS)* a cada pozo en orden de columnas, no por filas, de manera simultánea en vertical con la micropipeta multicanal. Se agitó la placa suavemente contra la mesa de trabajo durante 10 segundos y se incubó a temperatura ambiente durante 4 minutos. Se supervisó el desarrollo del color en los pozos de las muestras y de los controles. Después de unos minutos el DS se volvió azul en presencia de ADN metilado.
8. A los 5 minutos, el color de los pozos de PC al 5 % se tornó azul oscuro, por lo que se detuvo la reacción enzimática añadiendo 100  $\mu$ l de *Stop signal (SS)* a cada pozo de manera simultánea en forma de columnas, empleando la micropipeta multicanal. Se mezcló la solución agitando suavemente la placa y se esperó 2 minutos para permitir que la reacción de color se detuviera completamente. El color azul cambió a amarillo después de añadir SS.

La cuantificación se llevó a cabo leyendo la placa de ELISA en el espectrómetro *Thermo™ Scientific Multiskan GO Microplate Spectrophotometer* a 450 nm, siendo la cantidad de ADN metilado proporcional a la intensidad de DO medida, que puede calcularse con las fórmulas incluidas en el kit (2), para ello se realizó una curva estándar trazando los valores de DO frente a la cantidad del control positivo en cada

punto de concentración en Excel (EpigenTek, 2022). Se observó que la ecuación con mejor valor de  $R^2$  fue una de tipo logarítmica de segundo orden, por lo tanto, se utilizó la ecuación que arrojó la gráfica, se despejó y se sustituyeron los valores correspondientes.

$$5mC(ng) = \frac{DO \text{ de la muestra} - DO \text{ del control negativo}}{\text{Pendiente de la curva} \times 2} \quad (2)$$

$$5mC \% = \frac{\text{cantidad de } 5mC(ng)}{\text{Cantidad de muestra inicial de ADN}(ng)} \times 100 \%$$

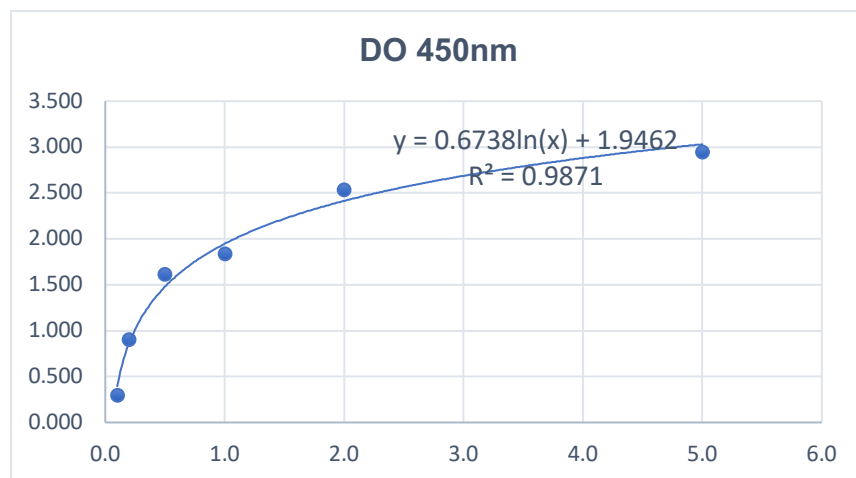


Figura 10. Curva estándar del estuche de metilación.

#### 5.2.4 Identificación de los perfiles infrarrojos séricos

Para cumplir con el segundo objetivo particular planteado, se llevaron a cabo los siguientes procedimientos para la obtención de los perfiles infrarrojos FTIR.

##### 5.2.4.1 Preparación de las muestras

El primer paso consiste en la liofilización del suero, utilizando la liofilizadora *FreeZone* de 2.5 litros Benchtop (Labconco, Kansas City, Misuri, Estados Unidos). Antes de realizar la liofilización, se aseguró que las muestras estuvieran completamente congeladas, ya que es un requerimiento imprescindible para que el proceso pueda llevarse a cabo correctamente. Debido a que esta etapa metodológica se llevó a cabo en la Universidad Tecnológica de la Mixteca, después del transporte en hielo seco, las muestras se dejaron en el congelador horizontal

del Laboratorio de Análisis Instrumental del Instituto de Agroindustrias por 48 horas para asegurar el congelamiento a una temperatura constante. Una vez que se confirmó el congelamiento, se procedió a prepararlas para la liofilización. Para realizar este procedimiento, a cada tubo de microcentrífuga que contuviera la muestra, previamente etiquetados con el número de identificación, se le cortó la tapa y en su lugar se taparon con papel filtro, sellado con papel Parafilm™. Esta manipulación de las muestras ocasionó cambios en la temperatura, por lo que una vez concluida la preparación, se dejaron congelar por 2 horas más. Posteriormente se transportaron en una nevera con hielo hasta el cubículo donde se encontraba la liofilizadora. Se utilizaron 3 vasos de vidrio de Labconco para colocar las muestras. Se verificó el estado del equipo, así como de las válvulas, tubos y la colocación correcta de los vasos. La liofilización se programó con una duración de 16 horas a  $-58\text{ }^{\circ}\text{C}$  y con una presión de 0.014 a 0.021 mBar. Al finalizar la liofilización, mediante un proceso inverso para eliminar el vacío, se dejó que el aire entrara lentamente, después a las muestras se les colocaron sus tapas y fueron preservadas en un desecador hasta el momento en el que se realizara la obtención de los espectros FTIR (Figura 11).



Figura 11. Liofilización de muestras de suero.



Este procedimiento de liofilización también se realizó con el ADN extraído previamente con el fin de obtener espectros del ADN. Para tal fin, primero se hicieron alícuotas de 10  $\mu$ l de ADN. Se planteó utilizar bromuro de potasio (KBr) para obtener una muestra sólida que fuera más estable para su análisis espectral y no tener la presencia de bandas de agua, dadas las características del KBr: es transparente al FTIR ya que no absorbe en la región del infrarrojo medio y no reacciona químicamente con la estructura del ADN, además de que es ampliamente utilizado en el análisis espectral (Han et al., 2018). El KBr se dejó durante toda una noche en una estufa a 105 °C, para eliminar cualquier humedad. A cada muestra de ADN se le agregaron 50  $\mu$ l de agua ultrapura esterilizada, y 20 mg de KBr, se mezcló e inmediatamente se regresaron al congelador por 30 minutos. Posteriormente se les cortó la tapa y se colocó el filtro siguiendo los mismos cuidados que con las muestras de suero.

La liofilización se llevó a cabo a 0.014 mBar a -58 °C durante 4 horas. Una vez concluido el tiempo, se dejaron en el desecador. Después de analizar en el espectro estas muestras y al observar la presencia de bandas de agua, se decidió volver a liofilizar todas las muestras por 4 horas más, para lo que se repitió el mismo proceso de congelamiento y preparación. Se utilizó un control de metilación positiva del estuche *MethylFlash™ Methylated DNA 5-mC Quantification Kit (Colorimetric)* de la casa comercial EpigenTek (Farmingdale, Nueva York, Estados Unidos) que contenía 50 % de 5mC en una concentración de 20  $\mu$ g/ml, del cual se tomaron 20  $\mu$ l, se agregaron 50  $\mu$ l de agua ultrapura esterilizada y 20 mg de KBr para posteriormente llevar a cabo el proceso de liofilización siguiendo el mismo protocolo utilizado para las muestras de ADN.

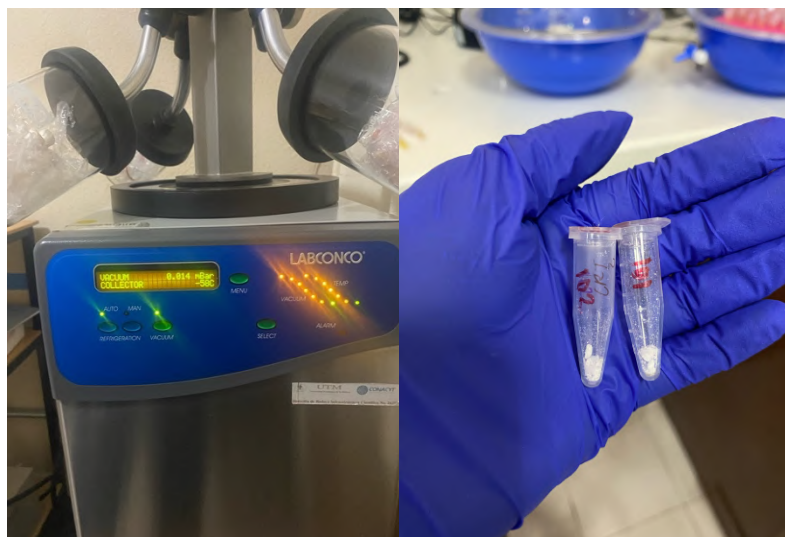


Figura 12. Liofilización de muestras de ADN.

#### 5.2.4.2 Obtención de los espectros FTIR-ATR

Para analizar las muestras y obtener los espectros se utilizó un equipo Bruker de transformada de Fourier modelo ALPHA P (Bruker Optik GmbH) acoplado con un accesorio de reflectancia total atenuada (ATR), y con un detector de sulfato de triglicina deuterada (DTGS) y un porta muestra de diamante con una geometría de paralelogramo de 45°.

Las muestras mantenidas en el desecador se seleccionaron en orden para su lectura espectral. Primero se analizaron las muestras de ADN. Antes de utilizar el equipo de FTIR se realizó la verificación de desviación mínima y la prueba automático. Después de hacer una prueba para evaluar la calidad de los espectros a 64 o 128 escaneos, se decidió utilizar 128, por lo que el equipo se configuró para leer 128 escaneos en modo de absorbancia en la región del infrarrojo medio (4000 a 400  $\text{cm}^{-1}$ ). Se estableció como fondo o *background* el KBr, para esto, primero se leyó el KBr en polvo. Posteriormente se leyeron cada una de las muestras de ADN por triplicado, colocando con una espátula una cantidad aproximada de 10 mg de muestra y bajando la prensa con la misma fuerza en cada colocación, se limpió el diamante y la placa del accesorio con agua ultrapura y etanol después de cada lectura.

La selección de la región espectral se llevó a cabo tomando primero varios intervalos del espectro infrarrojo medio en los que se concentran las bandas de absorción más intensas, posteriormente se fue delimitando la región comprendida hasta encontrar la que proporcionara los mejores resultados en el análisis multivariado, es decir, que en el análisis de componentes principales presentara un mayor porcentaje de variabilidad.

#### 5.2.5 Procesamiento de los espectros FTIR-ATR

Con el fin de eliminar información que no pudiera ser de utilidad para el análisis posterior de los datos espectrales, éstos se sometieron a métodos de preprocesamiento:

1. Corrección de la línea base.
2. Obtención de la media de las tres réplicas de cada muestra.
3. Normalización.
4. Filtro de Savitzky-Golay.
5. Escalado.

El tratamiento de los espectros, la selección de la región espectral, el preprocesamiento y análisis de los datos espectrales se llevó a cabo en Rstudio.

#### 5.2.6 Plan de procesamiento y presentación de la información

La distribución y homocedasticidad de todos los datos obtenidos fueron determinados mediante la prueba de Shapiro Wilk y las pruebas de Levene y Barlett, respectivamente. Los datos demográficos recolectados mediante el cuestionario fueron evaluados mediante un análisis de correlación de Pearson y Spearman para observar si presentan cambios en función de la metilación global del ADN.

##### 5.2.6.1 Análisis estadístico para el porcentaje de metilación global del ADN

Primero se evaluó la distribución de los datos mediante la prueba de Shapiro Wilk. Dado que los datos presentaban una distribución normal, se llevó a cabo una prueba de t de Student para buscar diferencias estadísticamente significativas en el porcentaje de metilación global del ADN entre ambos grupos (pacientes con cáncer de pulmón y control), planteando las siguientes hipótesis estadísticas:

$H_0$ : No hay diferencia significativa en el porcentaje de metilación global del ADN entre el grupo de pacientes con cáncer de pulmón y el grupo control.

$H_A$ : Existe diferencia significativa entre el porcentaje de metilación global del ADN del grupo de pacientes con cáncer de pulmón y el grupo control.

#### 5.2.6.2 Análisis estadístico para los espectros infrarrojos

Una vez obtenidos los espectros infrarrojos y posterior al procesamiento, se identificaron las bandas de interés relacionadas con la metilación del ADN de acuerdo con lo reportado en la literatura.

Adicionalmente, después de obtener la absorción del área de las bandas en forma de media y error medio estándar, estas se compararon mediante una prueba de t Student con el fin de buscar diferencias estadísticamente significativas entre los grupos de estudio. Estos valores de absorbancia se utilizaron posteriormente en los análisis de correlación.

#### 5.2.6.3 Análisis de correlación

Con el fin de cumplir el tercer objetivo particular propuesto, se analizó la posible correlación entre los datos obtenidos mediante la lectura de la placa del ELISA en el espectrofotómetro y los datos obtenidos mediante FTIR. Se llevó a cabo un análisis de correlación de Pearson utilizando R Studio.

#### 5.2.7. Consideraciones éticas

Este estudio, considerado en la categoría B, con riesgo mínimo, se realizó de acuerdo con lo manifestado en la Ley Federal de Salud en materia de Investigación para la salud en los capítulos I y V, en referencia a los principios de la no maleficencia, autonomía, beneficencia y justicia para con los participantes, así como con la responsabilidad de protección de la integridad, dignidad, confidencialidad y autodeterminación de su información personal, según lo que establecen las Pautas Éticas internacionales para la Investigación Biomédica en Seres Humanos y la Declaración de Helsinki. El protocolo de este estudio fue aprobado por el Comité de Bioética de la Facultad de Química de la UAQ con el No. de oficio CBQ24/003b.

#### 5.2.8. Manejo de residuos peligrosos biológico-infecciosos (RPBI)

Una vez utilizados, los residuos de las muestras y los desechos biológicos que incluya la metodología fueron manejadas de acuerdo con el protocolo de manejo de Residuos Peligrosos Biológico-Infecciosos (RPBI) que marca la NOM-087-ECO-SSA1-2002, envasando la sangre líquida y sus derivados obtenidos en recipientes herméticos, resistentes a fracturas y pérdidas de contenido, destructibles por métodos físicos, marcados con la leyenda “Residuos Peligrosos Líquidos Biológico-Infecciosos” y con el símbolo universal de riesgo biológico. Las agujas y objetos punzocortantes utilizados fueron colocados en recipientes para tales fines, de polipropileno color rojo y con la leyenda “Residuos Peligrosos Punzocortantes Biológico-Infecciosos” y con el símbolo universal de riesgo biológico correspondiente, debidamente cerrados y hasta el 80 % de su capacidad. Estos residuos fueron llevados a la zona destinada de almacenaje de la Unidad de Servicios Clínicos y Diagnóstico molecular del Edificio 9 de la Facultad de Química de la Universidad Autónoma de Querétaro, Centro Universitario, donde se tiene un calendario de recolección de cada 15 días por parte de la empresa PRIBESA, para su disposición final.

## 6 Resultados y discusión

### 6.1 Características de la población

Se recolectaron datos de edad, sexo, Índice de Masa Corporal (IMC) y el nivel de actividad física, ya que son variables reportadas previamente como factores que pueden influenciar cambios en el grado de metilación (Lam et al., 2012).

Mediante el cuestionario aplicado a los participantes, se obtuvieron las características demográficas que se presentan en el Cuadro 8.

Cuadro 8. Características demográficas y clínicas de los participantes. La edad y el IMC se muestran en promedio y desviación estándar. GC: Grupo control, CP: Cáncer de pulmón.

	<b>GC</b>	<b>CP</b>
Hombres	8	4
Mujeres	12	2
Edad	52.8 ± 11.8	73.5 ± 16.1
IMC	28.7 ± 5	23.9 ± 2.47
Nivel de actividad física	Alto	Bajo
Tipo de cáncer	No aplica	Adenocarcinoma (100%)
Estadio del cáncer	No aplica	IV (100%)
Exfumador	No aplica	6
Familiares directos con cáncer	No aplica	6

La participación de los pacientes con cáncer de pulmón estuvo limitada por la baja incidencia de casos diagnosticados en el estado de Querétaro, de acuerdo con el último conteo del Sistema Nacional de Vigilancia Epidemiológica, hasta junio de 2024 solo existía un registro de 6 pacientes con cáncer de pulmón y un acumulado de 3 durante la totalidad del 2023, sin embargo, no hay datos de la sobrevivencia de estos pacientes ni de su procedencia. La edad de los participantes, de acuerdo con los criterios de inclusión considerados, fue de un rango de los 40 a los 90 años. El 46.15 % de los participantes fueron mujeres de un rango de edad de los 40 a los 55 años (Figura 13). Tanto el IMC como el nivel de actividad física fueron más altos

en el grupo control, situación que puede estar relacionada con las manifestaciones clínico-patológicas de la enfermedad y que se discute en los apartados posteriores. La totalidad de los pacientes con cáncer fueron diagnosticados con adenocarcinoma por sus respectivos servicios médicos, como se esperaba, ya que es el subtipo de cáncer de pulmón con mayor prevalencia en México y el mundo (Heredia et al., 2023). El adenocarcinoma también es el que tiene mayor probabilidad de ocurrir en jóvenes y es el subtipo de cáncer con mayor incidencia en personas fumadoras activas o exfumadoras (American Cancer Society, 2023). Esta característica coincide con el historial de tabaquismo de todos los pacientes con cáncer de pulmón que participaron en el estudio. También puede observarse que todos los pacientes con cáncer de pulmón reportan haber tenido un familiar sanguíneo con diagnóstico de cáncer, lo que junto con el hábito tabáquico, representa dos de los principales factores de riesgo para el desarrollo de esta enfermedad.

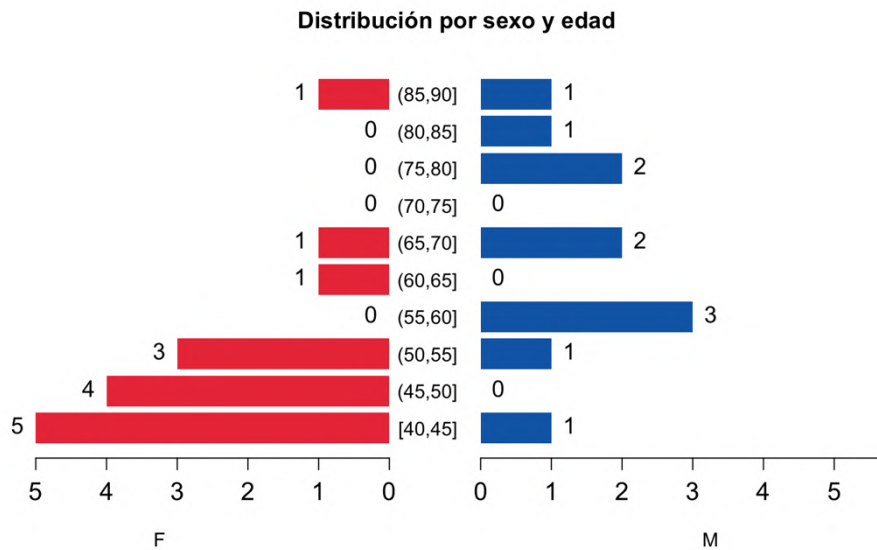


Figura 13. Distribución sexo - edad de los participantes. F: Femenino, M: Masculino. El rango de edad se muestra entre paréntesis al centro de la imagen y las barras representan el número de mujeres (rojo) y hombres (azul).

El Cuestionario Internacional de Actividad Física (IPAQ) es una herramienta que forma parte de una serie de estándares que se han utilizado para establecer los criterios a evaluar cuando se requiere valorar las actividades físicas que se realizan en cualquier país (Medina et al., 2022). Contiene 7 preguntas que evalúan la frecuencia, duración y la intensidad de la actividad realizada (moderada e intensa), a cada nivel de actividad se le calcula un valor en función de la periodicidad con la que se realiza y se multiplica por un coeficiente asignado en unidades de medición de la prueba (MET). Los criterios para el puntaje de la prueba se presentan en el Cuadro 9.

Cuadro 9. Criterios del Cuestionario Internacional de Actividad Física.

<b>Valor de la prueba</b>	<b>Criterios de clasificación</b>	<b>Nivel de actividad</b>
Caminatas: 3'3 MET x minutos de caminata x días por semana	<ol style="list-style-type: none"> <li>1. 3 o más días de actividad física intensa por lo menos 20 minutos por día.</li> <li>2. 5 o más días de actividad física moderada y/o caminata al menos 30 minutos por día.</li> <li>3. 5 o más días de cualquiera de las combinaciones de la caminata, actividad física moderada o intensa logrando como mínimo un total de 600 MET.</li> </ol>	Moderada
Actividad física moderada: 4 MET x minutos x días por semana	<ol style="list-style-type: none"> <li>1. Actividad física intensa por lo menos 3 días por semana logrando un total de al menos 1500 MET.</li> </ol>	Alta



	2. 7 días de cualquier combinación de caminata, con actividad física moderada y/o actividad física vigorosa, logrando un total de al menos 3000 MET.	
Actividad física intensa: 8 MET x minutos x días por semana	Cuando el nivel de actividad física del sujeto no esté incluido en las categorías alta o moderada.	Baja

El nivel de actividad física predominante en el grupo control fue alta, mientras que para el grupo con cáncer de pulmón fue baja, debido principalmente a que el 50 % de estos participantes se encontraban en cama o tenían movilidad limitada (Figura 14).

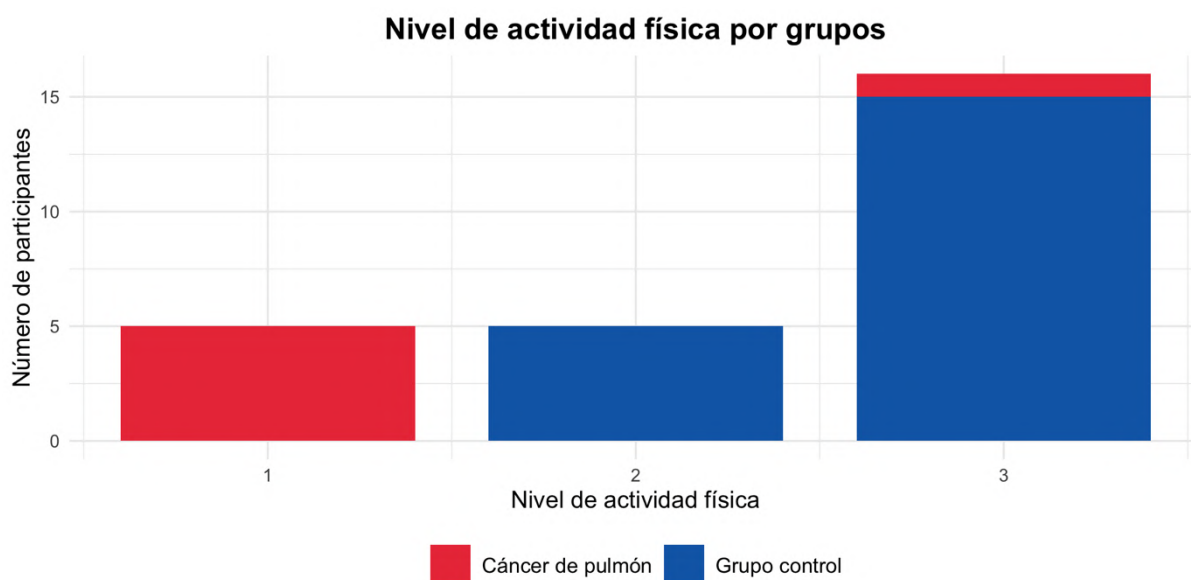


Figura 14. Nivel de actividad física por grupos. CP: Cáncer de pulmón, GC: Grupo control

El IMC se registró cuando los participantes lo conocían, de lo contrario, se calculaba con el peso y la estatura. En todos los casos de participantes con cáncer de pulmón, los valores de peso y estatura fueron comprobados con su expediente clínico, sin embargo, algunos de los participantes del grupo control dieron un aproximado de su peso. De acuerdo con el Instituto Mexicano del Seguro Social (IMSS) un IMC menor a 18.5 se considera peso insuficiente, entre 18.5 y 24.9 se clasifica dentro del rango de peso normal, entre 25.0 y 29.9 se encuentra dentro del rango de sobrepeso y si el IMC es igual o superior a 3.0, se encuentra dentro del rango de obesidad. El grupo control presentó un IMC mayor, contrario al grupo de cáncer de pulmón, en el que la totalidad de los pacientes estaban sometidos a tratamientos y alimentación especiales, reflejándose en valores de IMC más bajos (Figura 15).

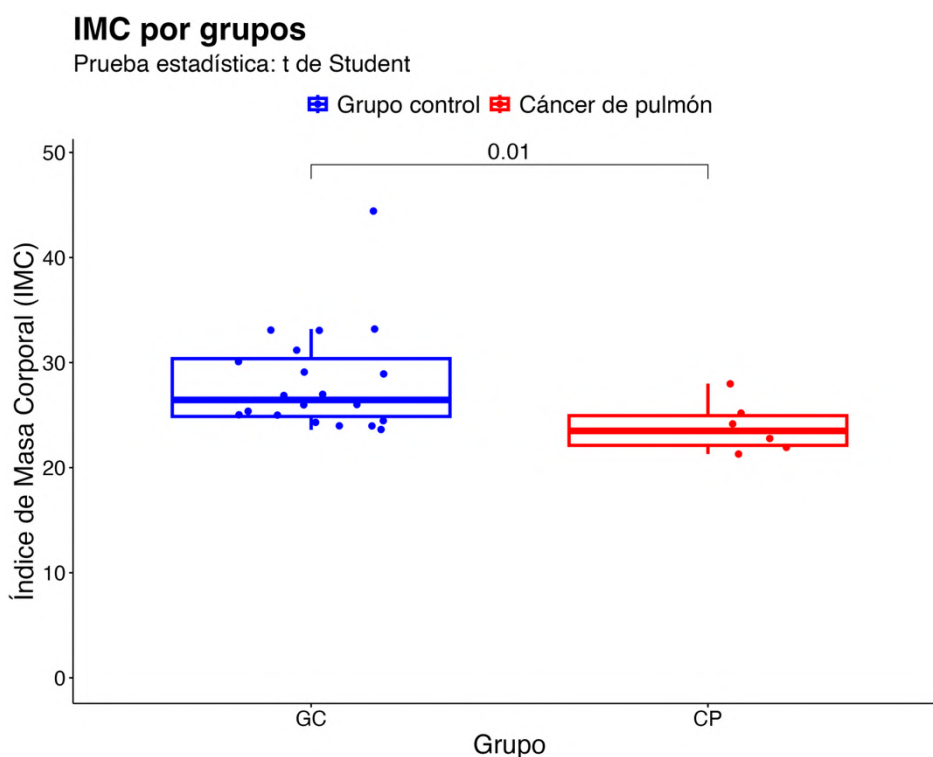


Figura 15. Comparación de promedios del Índice de Masa Corporal por grupos. CP: Cáncer de pulmón, GC: Grupo control. Prueba t de Student,  $p < 0.05$ .

## 6.2 Antecedentes familiares de cáncer como factor de riesgo en la población muestreada

La totalidad de participantes de cáncer de pulmón fueron identificados como adenocarcinoma etapa IV, de acuerdo con el historial clínico proporcionado por su oncólogo. Además, todos los pacientes con cáncer de pulmón refirieron haber tenido familiares directos con cáncer (ninguno de pulmón). Una variedad de estudios ha reportado un vínculo entre antecedentes de cáncer de pulmón en la familia directa y un mayor riesgo de desarrollar cáncer de pulmón. En una investigación realizada con la cohorte del Estudio de Genoma y Epidemiología de Corea (KoGES), se observó que entre los nunca fumadores, el riesgo de padecer cáncer de pulmón se elevó significativamente para aquellas personas con antecedentes de familiares con cáncer de pulmón ( $ORa = 3.77$ , IC del 95 % 1.59 - 8.93) (Kim et al., 2024). Además, se ha observado que los antecedentes familiares de otros tipos de cáncer en los familiares directos también aumentan las probabilidades de desarrollar cáncer de pulmón. En un estudio de casos y controles (Pineda-Lancheros et al., 2022), la presencia de cualquier tipo de cáncer en el historial familiar se asoció a un 49.5 % de los casos de cáncer de pulmón, 13.24 % de estos eran casos de nunca fumadores, 39.71 % eran ex fumadores y 47.96 % eran fumadores activos, se estableció que este historial familiar de cáncer se asociaba con el diagnóstico de CPCNP con un OR de 15.2 (IC del 95 %: 9.55 - 25.2).

## 6.3 Análisis de la distribución de las variables

Para analizar la distribución de las variables continuas y poder establecer el tipo de prueba estadística a realizar posteriormente, se construyeron histogramas y gráficas QQ (*Quantile-quantile*). El histograma muestra cómo se distribuyen los valores en rangos específicos y una curva de distribución normal (línea azul) representa una distribución normal ajustada a los datos. Las gráficas QQ por su parte, también permiten evaluar la normalidad de los datos. Estas gráficas contienen una línea diagonal de referencia que representa la distribución teórica que se espera de datos que siguen una distribución normal, por lo tanto, si una variable presenta

una distribución normal, los puntos deberían estar alineados a lo largo de esta diagonal. En el caso de la edad, puede observarse que la curva no se ajusta bien al histograma (Figura 16A), lo que sugiere que los datos de esta variable no presentan una distribución normal, la mayoría de los datos se concentran en el rango de 40 a 60 años, es decir, la distribución es asimétrica y está sesgada a la izquierda. La gráfica QQ muestra puntos alejados significativamente de la línea diagonal, indicando una vez más, una distribución sesgada (Figura 16B). En el extremo superior se pueden ver varios puntos acumulados, indicando que la distribución de los datos presenta una cola positiva, como lo que se pudo observar en el histograma. Al aplicar la prueba Shapiro Wilk se obtuvo un valor  $p = 0.00428$ , siendo menor a 0.05, con esto se confirma que la edad no sigue una distribución normal. Los datos correspondientes al IMC presentaron una distribución sesgada de acuerdo con lo observado en el histograma y el gráfico QQ (Figura 17). La mayoría de los participantes se encontraron en el rango de 25 a 35, lo que corresponde a las categorías de sobrepeso y obesidad grado I, respectivamente. Sin embargo, la prueba Shapiro Wilk arrojó un valor de  $p = 0.0006659$ , y al ser menor que 0.05 permite rechazar la hipótesis nula que plantea que los datos siguen una distribución normal.

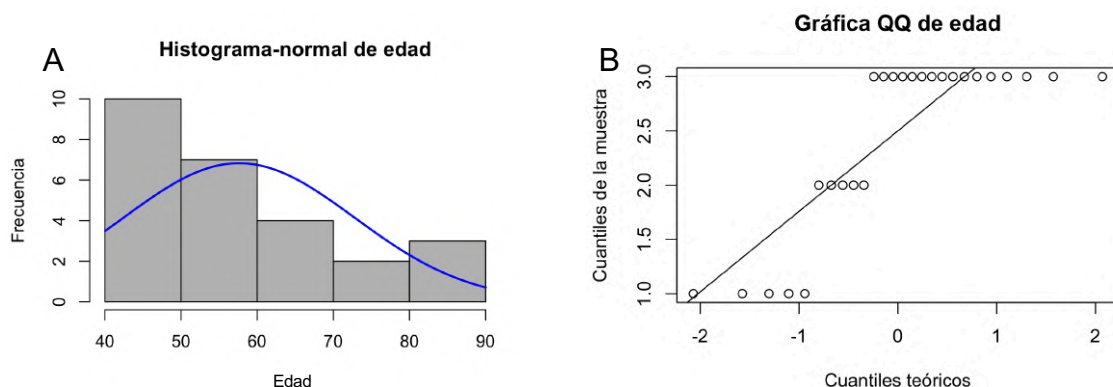


Figura 16. Distribución de la variable edad. A: Histograma con curva de densidad normal teórica (línea azul) B: Gráfico QQ que compara los cuantiles observados con los esperados bajo normalidad.

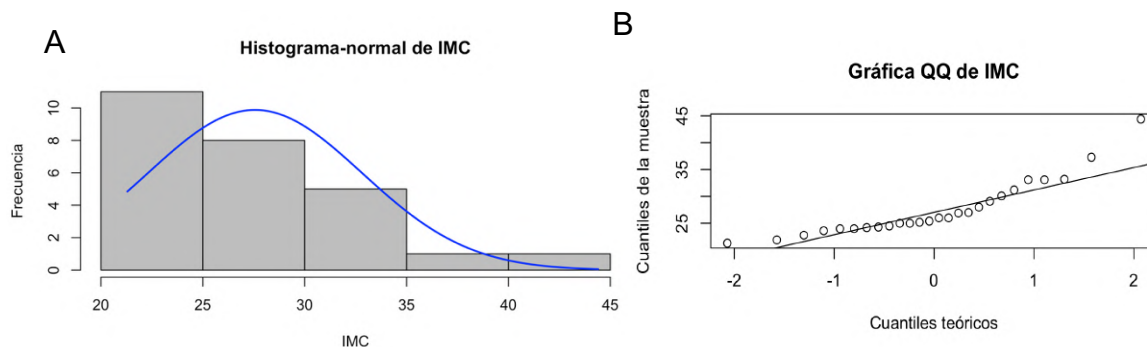


Figura 17. Distribución de la variable IMC. A: Histograma con curva de densidad normal teórica (línea azul) B: Gráfico QQ que compara los cuantiles observados con los esperados bajo normalidad.

#### 6.4 Porcentaje de metilación global

En el histograma realizado con los valores obtenidos del porcentaje de metilación global del ADN (Figura 18A) se muestra una distribución que se asemeja a una normal, aunque presenta valores que se salen de la curva ajustada, especialmente en los extremos. Esta distribución aproximadamente normal y los valores que se alejan también pueden verse en la gráfica QQ (Figura 18B). Esta normalidad se comprobó mediante la prueba Shapiro Wilk ( $W = 0.097885$ ,  $p = 0.8487$ ), al obtenerse un valor de  $p > 0.05$  no hay evidencia suficiente que permita rechazar la hipótesis nula que establece que esta variable sigue una distribución normal.

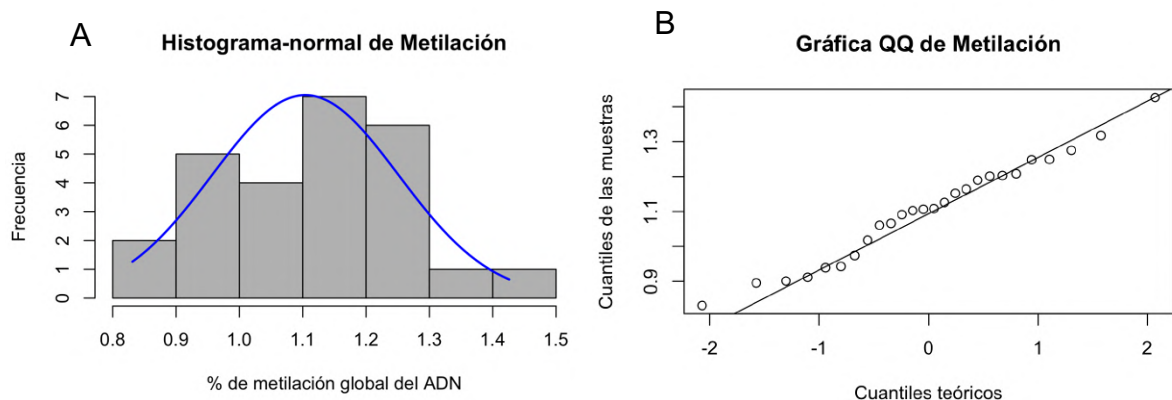


Figura 18. Distribución de la variable metilación del ADN. A: Histograma con curva de densidad normal teórica (línea azul) B: Gráfico QQ que compara los cuantiles observados con los esperados bajo normalidad.

Las pruebas de Levene y Bartlett permiten evaluar la homocedasticidad de los datos, y después de aplicarlas a la variable metilación, se obtuvieron valores de  $p = 0.925$ , y  $p = 0.803$ , respectivamente, por lo tanto, siendo  $> 0.05$ , no se rechaza la hipótesis nula de que las varianzas entre los grupos son iguales, es decir, las varianzas son homogéneas, por lo que pueden realizarse análisis paramétricos posteriores con esta variable.

#### 6.4.1 Porcentaje de metilación global del ADN entre grupos

Después de corroborar la distribución normal y la homocedasticidad de los datos, se aplicó la prueba  $t$  de Student para buscar diferencias estadísticamente significativas entre las medias de ambos grupos de estudio. El test de Welch es una variante de la prueba  $t$  de Student que no asume varianzas iguales, y que además permite comparar grupos con números de muestras muy diferentes, como en este caso (20 muestras de grupo control vs. 6 muestras de grupo cáncer de pulmón), por esta razón, se eligió implementar esta prueba (Lu & Yuan, 2010). Se obtuvo un valor de  $p = 0.048$ , siendo menor a  $p = 0.05$  se puede afirmar que existen diferencias estadísticamente significativas entre los porcentajes de metilación de ambos grupos.

con un 95 % de confianza, y que además, hay un nivel más bajo en el grupo de cáncer de pulmón. El valor de p obtenido con la prueba convencional de t de Student fue de  $p = 0.02065$ , es decir, sigue confirmando la existencia de una diferencia estadísticamente significativa entre las medias de los grupos (Figura 19).

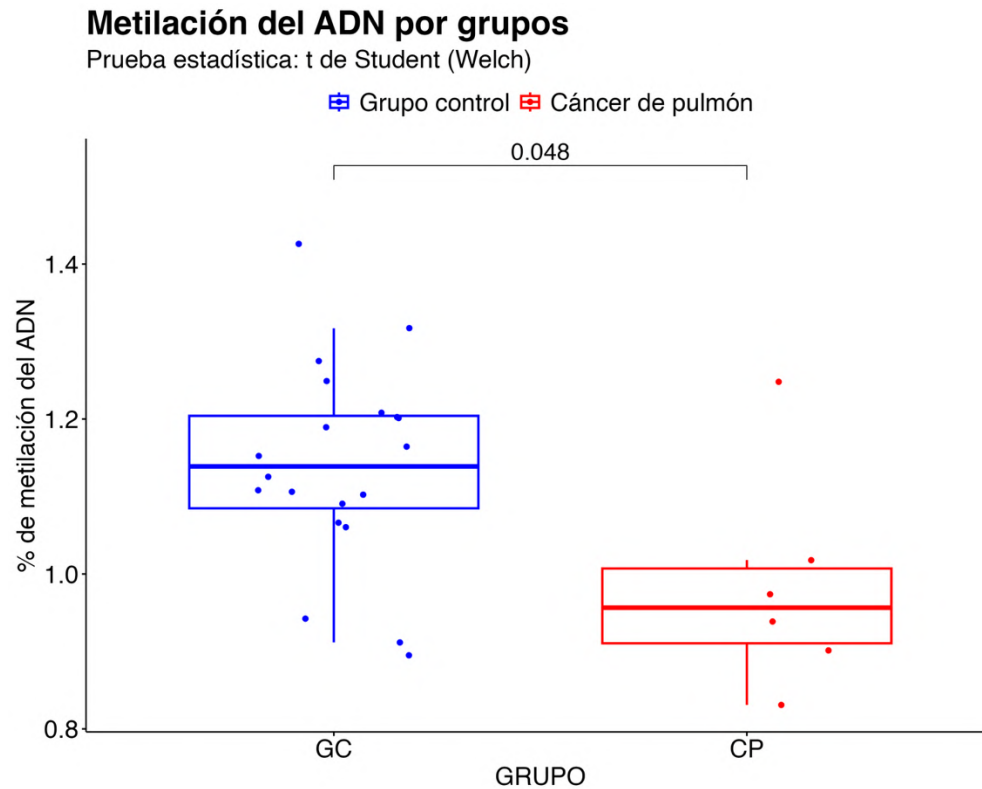


Figura 19. Comparación de medias del porcentaje de metilación global entre grupos de estudio. Prueba t de Student (Welch),  $p < 0.05$ .

#### 6.4.2 Evaluación de la asociación entre variables

Con el fin de explorar posibles asociaciones entre variables, se llevó a cabo un ANOVA en RStudio, en el cual se utilizó la metilación global como variable dependiente. Se eligió este análisis ya que no solo compara las medias de diferentes grupos, que en este caso son 2, sino que también construye un modelo estadístico que describe la relación entre los factores y la variable dependiente, haciendo ajustes para la varianza y las posibles interacciones entre los distintos

factores que se recolectaron mediante el cuestionario a los participantes (Weiss, 2005).

Para dicho análisis se consideraron las variables categóricas IMC, IPAQ y Sexo como factores, empleando la función “as.factor”. Una vez obtenido el modelo de ANOVA, se verificaron los supuestos de normalidad de residuos (prueba Shapiro Wilk) y homogeneidad de las varianzas (prueba de Levene). Se obtuvo un valor de  $p = 0.9953$  en la prueba de normalidad, indicando que si  $p > 0.05$ , se cumple el supuesto de la normalidad de los residuos del modelo.

La prueba de Levene se realizó por pares de variables, obteniéndose un valor de  $p = 0.6528$  al evaluar Metilación x Sexo,  $p = 0.925$  al evaluar Metilación x Grupo y  $p = 0.8533$  al evaluar Metilación x IPAQ. Por lo tanto, en los tres casos, siendo  $p > 0.05$ , indica que las tres varianzas son homogéneas. Cumpliendo estos supuestos se puede interpretar el ANOVA (Cuadro 10). El grupo y el nivel de actividad física son los factores que están explicando la mayor variabilidad, lo que se comprueba con los valores de  $p$  para ambos, confirmando que la metilación global del ADN se ve afectada por la condición de cáncer/no cáncer y por el nivel de actividad física.

Cuadro 10. ANOVA entre variables. IMC: Índice de Masa Corporal, IPAQ: Cuestionario Internacional de Actividad Física. GL: Grados de libertad, SC: Suma de cuadrados, MC: Media de cuadrados, F: Estadístico F. Nivel de significancia: \*0.05, \*\*0.01, \*\*\*0.001.

	GL	SC	MC	F	Valor p
<b>Género</b>	1	0.00004	0.00004	0.003	0.95676
<b>Grupo</b>	1	0.12106	0.12106	8.429	0.00911**
<b>Edad</b>	1	0.00051	0.00051	0.035	0.85279
<b>IMC</b>	1	0.00120	0.00120	0.083	0.77600
<b>IPAQ</b>	2	0.14547	0.07274	5.064	0.01726*
Residuos	19	0.27288	0.01436		



Al comparar los valores de metilación agrupando los participantes por sexo (Figura 20), no se encontraron diferencias estadísticamente significativas ( $p=0.94$ , prueba t de Student). Si bien no se ha establecido que la metilación pueda variar de acuerdo con el sexo, este análisis se incluyó para descartarlo como un factor influyente en análisis posteriores.

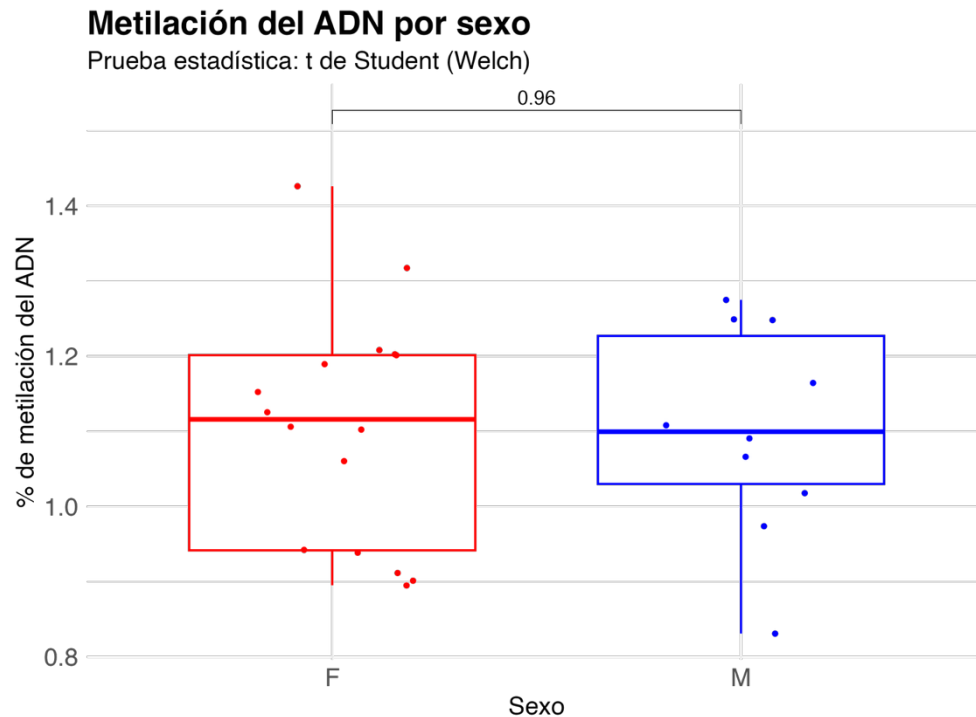


Figura 20. Comparación del porcentaje de metilación del ADN por sexo. F: Femenino, M: Masculino.

Siguiendo los resultados obtenidos en el ANOVA, se realizó una comparación del grado de metilación global del ADN en función de la clasificación por nivel de actividad física (baja, moderada y alta) para visualizar las diferencias estadísticas entre el grado de metilación. Un valor de  $p=0.025$  indica que la principal diferencia radica entre aquellas personas que realizan actividad física moderada y las que realizan actividad física alta dentro del grupo control (Figura 21). Para evaluar si el nivel de actividad física podría ser una variable confusora se llevó a cabo una prueba exacta de Fisher: Nivel de actividad física x Grupo, obteniéndose una  $p = 0.00018$ ,

resultado que indicaría que lo es, sin embargo, es importante notar que el grupo de CP tiene 5 participantes en el nivel bajo y uno en el alto, por lo que tanto el número de muestras como la baja representación de cada nivel de actividad física son factores que podrían dar lugar a interpretaciones erróneas sobre esta característica. Por lo tanto, el nivel de actividad física no es una variable confusora y tampoco es confiable para análisis posteriores.

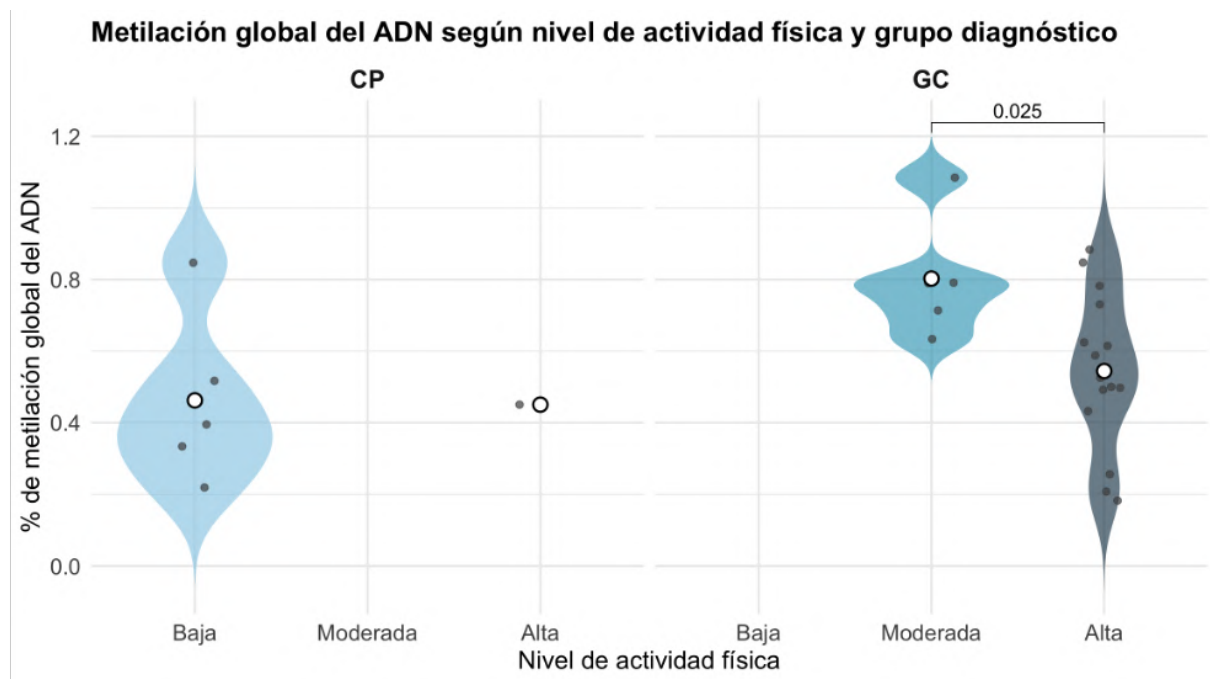


Figura 21. Comparación del porcentaje de metilación del ADN por nivel de actividad física. CP: Cáncer de pulmón. GC: Grupo control.

#### 6.5 Correlación de las variables demográficas con el porcentaje de metilación del ADN

Con el fin de conocer si alguna de las variables demográficas presentaba alguna correlación, en qué medida y dirección con el porcentaje de metilación global del ADN como se ha indicado en estudios previos (Lam et al., 2012a), se calcularon los coeficientes de correlación de Pearson, primero, tomando todas las variables como normales y posteriormente, aplicando una transformación de las variables continuas

IMC y Edad utilizando el paquete BestNormalize de R Studio, que sugirió la transformación de Yeo Johnson como la más adecuada. Después de la transformación se comprobó que los conjuntos de datos fueran iguales, se corroboró la normalidad de los datos transformados mediante la prueba Shapiro Wilk ( $p= 0.939$  para IMC y  $p= 0.06723$  para Edad) y se volvió a calcular la matriz de correlación, sin embargo, ambos resultados, con y sin transformación fueron muy similares. Estos resultados se graficaron en un mapa de calor de correlaciones (Figura 22).

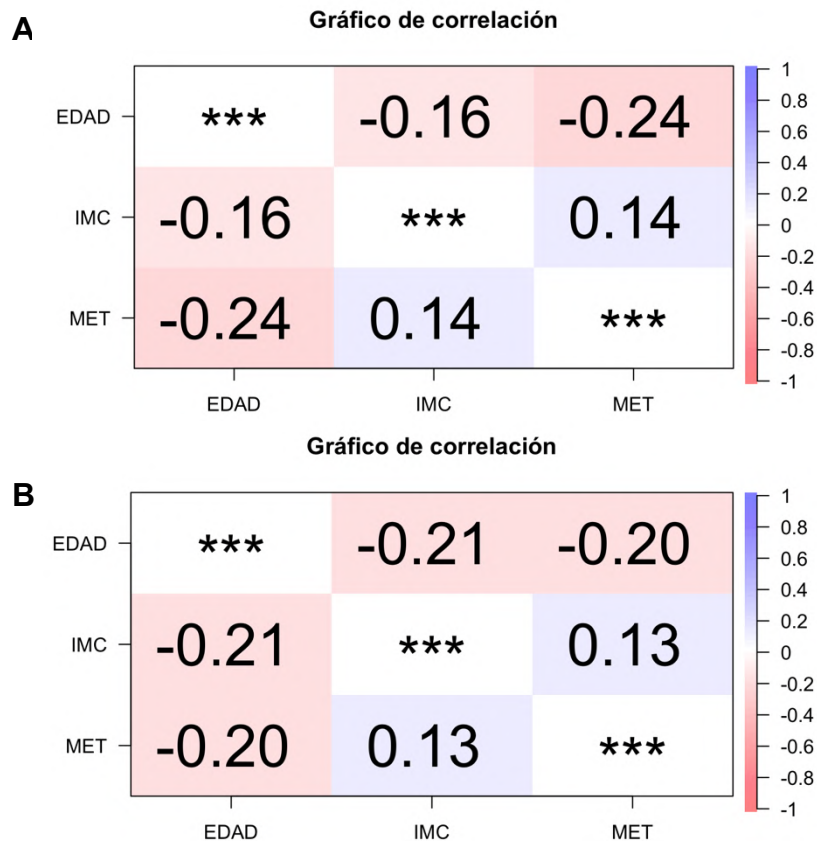


Figura 22. Matriz de correlaciones de Pearson de las variables demográficas con el porcentaje de metilación del ADN. A: Sin y B: Con transformación de variables no normales IMC y Edad. Los valores numéricos indican la magnitud y dirección de la correlación. El color representa la intensidad del coeficiente de correlación, desde azul (positiva) hasta rojo (negativa). \*Correlación con baja significancia: valor  $p < 0.05$ , \*\*Correlación significativa valor:  $p < 0.01$ , \*\*\*Correlación muy significativa (valor  $p < 0.001$ ).

La edad y el IMC muestran una correlación de -0.16, es decir, que no hay una relación fuerte entre el envejecimiento y la masa corporal. Entre la edad y la metilación del ADN existe una correlación débil y negativa (-0.24), lo que puede indicar que a medida que aumenta la edad, existe una ligera disminución del porcentaje de metilación global del ADN, esto coincide con la observación que se había reportado previamente en diversos estudios (Farré et al., 2015; Lam et al., 2012; Tollefsbol, 2023). El IMC y la metilación del ADN tienen una relación débil y positiva, que indicaría que a medida que aumenta el IMC se presenta un nivel más alto de metilación del ADN, sin embargo, ninguna de las correlaciones presentadas tiene significancia estadística. Las variables IPAQ y Sexo no se incluyeron en esta matriz de correlación al ser categóricas, por lo que la relación de ambas con la metilación del ADN se exploró en función de lo obtenido en el ANOVA considerándolos como factores.

Estos resultados preliminares permiten sugerir que en esta población muestreada no se está observando la tendencia sugerida en estudios previos de una disminución de la metilación en función del envejecimiento, pero sí se observa que existen diferencias estadísticamente significativas entre los grupos, confirmándose que en el cáncer de pulmón existe un proceso de hipometilación que es característico de etapas avanzadas de la enfermedad.

## 6.6 Análisis espectral

### 6.6.1 Pre-procesamiento de los espectros de ADN

El pre-procesamiento de los datos se realizó de manera secuencial siguiendo las recomendaciones de trabajos previos para el tratamiento de datos espectrales de muestras biológicas (Baker et al., 2014; Lee et al., 2017). El primer paso fue la corrección de la línea base (Figura 23b, Figura 24b). La corrección de la línea base tipo banda elástica (“rubberband”) se eligió ya que fue la que mejor se ajustó a los espectros obtenidos, además, este es el tipo de corrección de línea base con el que

opera el software OPUS 5.0 de Bruker, la marca del equipo que se utilizó para adquirir los espectros. El algoritmo de banda elástica consiste en dividir el espectro en  $n$  rangos ( $n$  es el número de puntos de la línea de base) de igual tamaño, en los espectros de absorbancia se determina el valor y el mínimo de cada rango, posteriormente se unen los mínimos con líneas rectas, construyendo la línea base, partiendo desde abajo, se estira una banda elástica sobre esta curva, convirtiéndose en la línea base. Si quedan puntos de la línea base que no se encuentran en la línea base, se descartan (Shen et al., 2018).

El siguiente paso del pre-procesamiento fue el suavizado de los espectros, lo que consiste en un procedimiento en el que se realiza una convolución del espectro mediante una función polinómica para mejorar la relación señal-ruido. Los métodos de suavizado, por lo tanto, eliminan las señales de alta frecuencia, que son características del ruido, y mejoran las señales de los grupos químicos. Este paso es de suma importancia ya que existe una tendencia a utilizar un alto grado de suavizado que puede llegar a distorsionar los espectros, por lo que es recomendable probar distintos niveles de suavizado. La aplicación del filtro Savitzky-Golay (con una función polinómica de segundo grado y 25 puntos para construir la función) a los espectros de ADN permitió eliminar señales no deseadas que pudieran ser ajenas a la composición de la muestra y permitir observar señales mejor definidas (Figura 23c, Figura 24c). El método de Savitzky-Golay es de los más utilizados en FTIR, consiste en el reemplazo del valor en un punto dado calculando el valor a partir del ajuste de un polinomio de mínimos cuadrados a un grupo de puntos vecinos. Para adecuar el nivel de suavizado, se debe establecer el grado del polinomio y el número de puntos vecinos (Kokalj et al., 2011; Ortega et al., 2024).

Posteriormente, se realizó la normalización de los espectros, este método establece el valor máximo de los espectros en 1 y el mínimo en 0, este paso es esencial para eliminar el efecto confusor de variar el grosor de la muestra en la intensidad de la banda. Con ese método todos los valores de intensidades se normalizaron de 0 a 1. La normalización también ajusta cambios en la potencia del equipo, diferencias

en la forma en la que la muestra dispersa la luz o variaciones en el espesor de la muestra de ADN que se colocó en cada medición (Mokari et al., 2023).

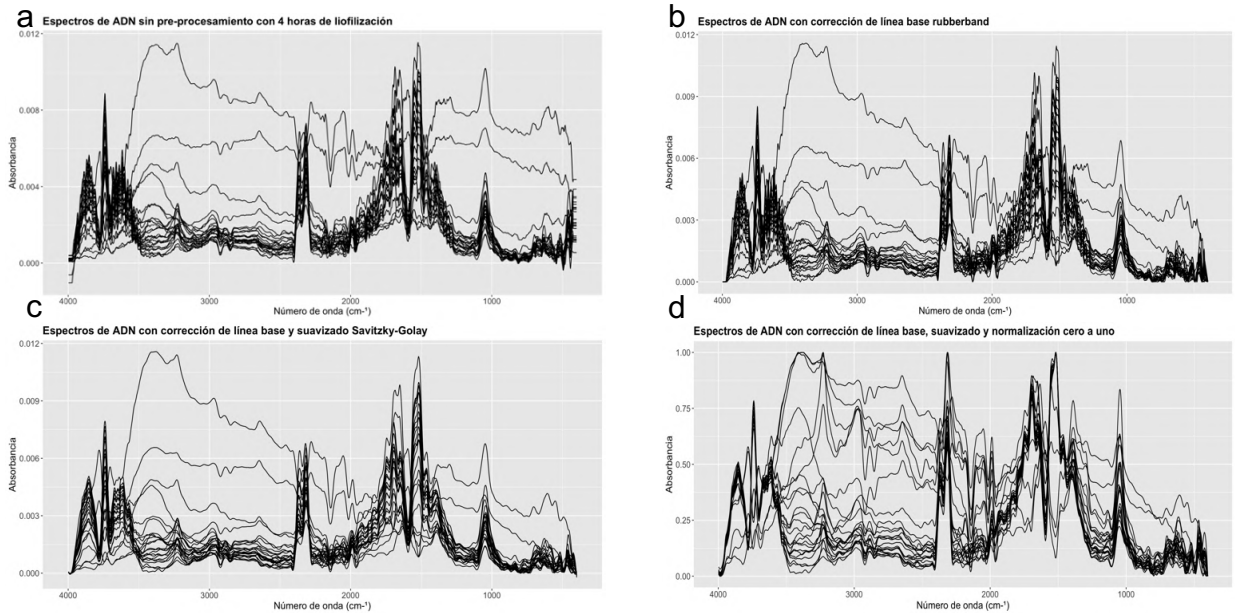


Figura 23. Pre-procesamiento de espectros a las 4 horas de liofilización.

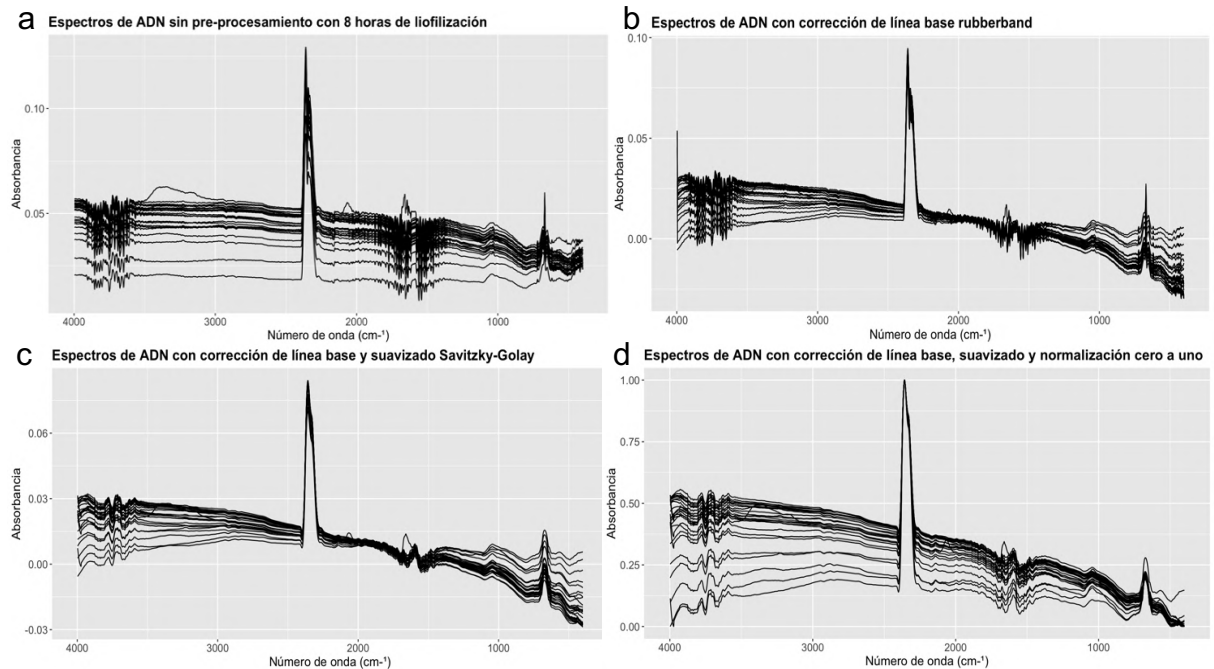


Figura 24. Pre-procesamiento de espectros de ADN a las 8 horas de liofilización.

Las vibraciones en las regiones entre 2800 y 3600  $\text{cm}^{-1}$  corresponden al agua, la primera asociada al estiramiento simétrico, mientras que la segunda, correspondiente al estiramiento asimétrico de los enlaces O-H (Mondragón-Cortez, 2020). Además, la banda cercana a los 1650  $\text{cm}^{-1}$  se asocia a la flexión en tijera del enlace H-O-H. Las bandas del agua en un espectro proveniente de una muestra biológica pueden llegar a solapar otras bandas de interés, sobre todo la banda de los 1650  $\text{cm}^{-1}$  en el caso particular de los espectros de ADN, que presentan el mayor número de bandas de interés en la región de los 1800  $\text{cm}^{-1}$  a 400  $\text{cm}^{-1}$ . Como puede notarse en los espectros, estas bandas del agua “levantan” todo el espectro (Figura 25a). Por esta razón, se optó por aumentar el tiempo de liofilización a 8 horas y repetir la adquisición espectral. Tras obtener los espectros de las muestras de ADN liofilizadas durante 8 horas, se observaron espectros completamente distintos a los obtenidos a las 4 horas de liofilización. Se notó una pérdida importante de bandas en la región de los 1800 a 400  $\text{cm}^{-1}$ . Se puede apreciar que el espectro se ve más aplanado (Figura 25b), lo cual puede atribuirse a que las bandas características del agua que se tenían a las 4 horas de liofilización ya no levantan a todo el espectro. Además, fue evidente la predominancia de la banda de los 2360  $\text{cm}^{-1}$  (Figura 25b).

Esta banda corresponde al dióxido de carbono ( $\text{CO}_2$ ), que normalmente se ve como un doblete a aproximadamente 2400  $\text{cm}^{-1}$ . Es común que al utilizar un accesorio de reflectancia total atenuada (ATR) queden algunas señales del fondo o *background*, incluso cuando previamente se ha leído el espectro ambiental o del blanco (en este caso se utilizó KBr como fondo), esto sucede especialmente en muestras en forma de polvo, debido a que una porción del  $\text{CO}_2$  queda capturado entre las partículas de la muestra del polvo de muestra que se colocó en la placa del ATR.

Otra posible explicación es que al aumentar las horas de liofilización también se prolongó el tiempo de exposición de las muestras, periodo en el que pudo haber quedado atrapada una mayor cantidad de  $\text{CO}_2$  en la muestra del ADN y pudo ser absorbido por el KBr (Fehr & Krossing, 2020; Mondragón-Cortez, 2020).

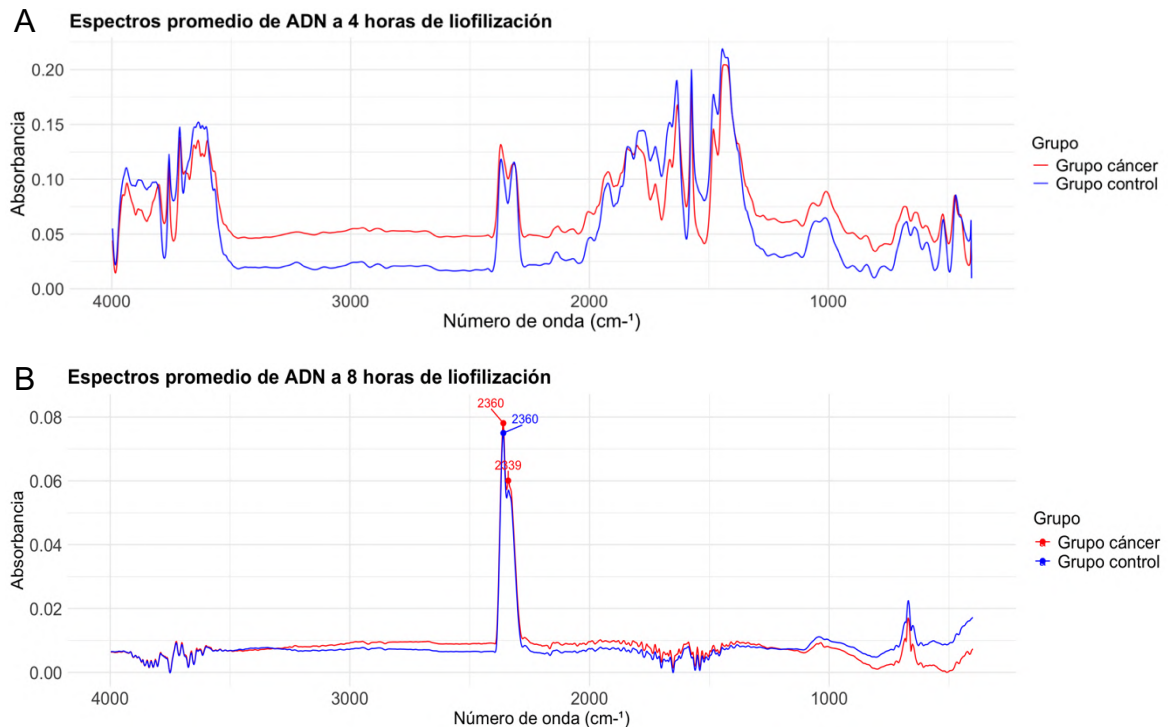


Figura 25. Comparación de espectros promedio de muestras de ADN a 4 (A) y 8 horas de liofilización (B).

#### La influencia del agua en el análisis espectral del ADN

Se propone que la humedad ha tenido una gran influencia en la calidad de los espectros obtenidos, esta observación se puede respaldar con lo reportado por (Wood, 2016), que señala que cuando el contenido de agua es bajo (< 75% de humedad relativa), el ADN adopta su conformación de tipo A, además de estar influenciado por factores ambientales como el pH y la concentración de sal. La estabilidad de la doble hélice del ADN depende también de un equilibrio de enlaces de hidrógeno entre otras bases y las moléculas de agua que la rodean.

Se ha observado que existe un aumento aparente de la absorbancia del ADN de doble cadena hidratado como posible resultado de un tipo de hipercromicidad. El efecto hipercrómico o hipercromicidad se refiere al aumento de la absorbancia de una muestra, en este caso, de ADN, en longitudes de onda específica al sufrir cambios como la desnaturalización. Este fenómeno se puede observar en los espectros de absorción UV-Vis a 260 nm, cuando el ADN pasa de un estado de



doble hélice a una sola hélice, por lo que al ser dos cadenas separadas, se absorbe una mayor cantidad de luz ultravioleta (Aufdembrink et al., 2019).

En artículos de referencia sobre los primeros análisis de ADN con espectroscopía se ha reportado una serie de cambios en los espectros de ADN sódico (neutralizado con NaOH) a medida que se aumenta el porcentaje de humedad, tales como aumentos en las intensidades de bandas entre los 1700 y 900  $\text{cm}^{-1}$ , así como desplazamientos de banda de 1240 a 1220  $\text{cm}^{-1}$ , que fue asignada a la vibración de estiramiento asimétrica  $\text{PO}_2^-$ , el desplazamiento de una banda de 1220 a 1240  $\text{cm}^{-1}$  de una muestra de ADN puro se propuso como consecuencia de la eliminación de las moléculas de agua, ya que el grupo  $\text{PO}_2^-$  queda libre de cualquier enlace hidrógeno. En estos estudios se propone que el ADN en su conformación A se produce a una humedad por debajo del 75 %, mientras que el ADN-B se forma a una humedad considerada alta o mayor al 75 % (Aufdembrink et al., 2019; Whelan et al., 2011). Además, señalan que a alta humedad las bandas en la región de 1200 a 750  $\text{cm}^{-1}$  se pueden observar con mejor nitidez y definición, demostrando que la cristalinidad de la muestra mejoraba con la absorción de agua (Figura 26). Es decir, que a menor humedad tiene lugar una transición reversible que produce una forma desordenada en la que las bases ya no están apiladas unas sobre otras ni son perpendiculares al eje de la hélice. En estudios de rehidratación del ADN (Wood, 2016) se ha propuesto la hipótesis de que el agua es absorbida primero por los grupos  $\text{PO}_2^-$ , posteriormente por los grupos C-O-P y C-O-C del enlace fosfodiéster, mientras que las bases nitrogenadas son las últimas en hidratarse. En ese estudio compararon la similitud de los cambios observados durante la deshidratación con los que sufre el ADN al desnaturalizarse por calentamiento. Por estos cambios observados en los espectros obtenidos a 4 y 8 horas de liofilización, se decidió continuar el análisis multivariado utilizando los espectros de ADN adquiridos a las 4 horas de liofilización.

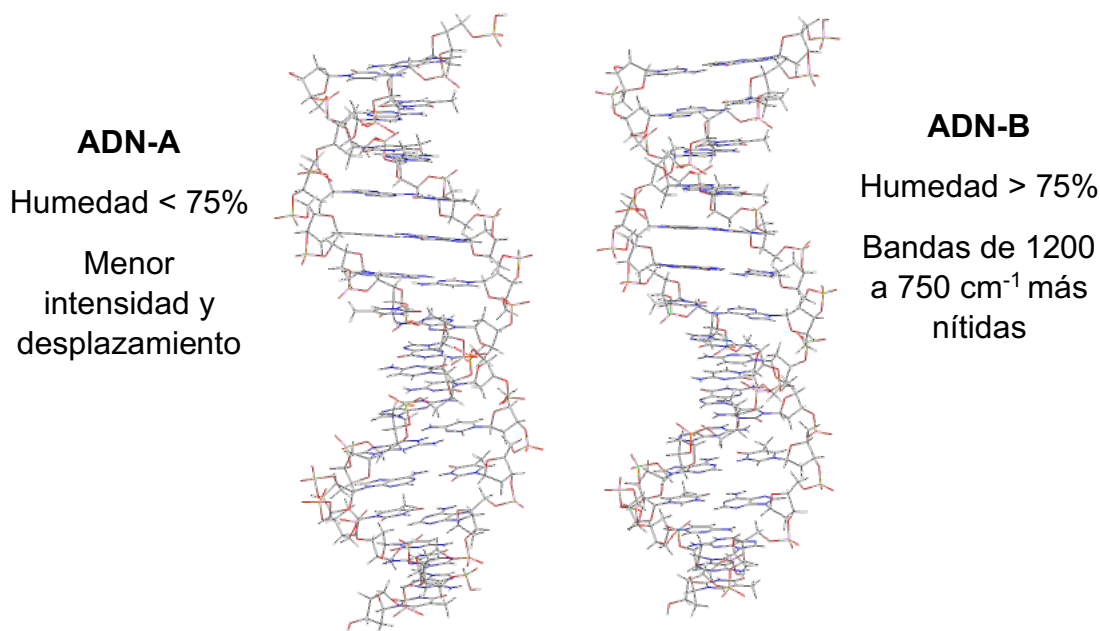


Figura 26. Efectos del agua en la conformación y el espectro infrarrojo del ADN según la literatura.

El espectro del control comercial de ADN metilado permitió identificar bandas que coincidieran con las encontradas en las muestras de ADN y de esta manera poder atribuirles correctamente a vibraciones de la estructura del ADN. En la Figura 27 se muestran los tres espectros donde puede observarse la tendencia a compartir bandas en las regiones de  $3600 - 3500 \text{ cm}^{-1}$  y  $2300 - 800 \text{ cm}^{-1}$ . De estas bandas se resaltan aquellas reportadas como indicadoras de la vibración de diferentes enlaces de la citosina presente en el ADN. Las bandas entre  $3300$  y  $3500 \text{ cm}^{-1}$  se atribuyen al enlace N-H de los grupos amino y amida, y a veces aparecen en forma de banda ancha debido a enlaces de hidrógeno, en el espectro presentado se pueden percibir especialmente desplazadas en el control de metilación ( $3466$ ,  $3524$ ,  $3607 \text{ cm}^{-1}$ ), mientras que en ambos grupos de muestras de ADN están más próximas ( $3600$  y  $3606 \text{ cm}^{-1}$ ). La banda de  $1695 \text{ cm}^{-1}$  se atribuye al enlace  $\text{C}_2=\text{O}$  de la citosina y puede observarse predominante en el espectro del control de metilación. La banda de  $1539 \text{ cm}^{-1}$  es normalmente asociada a la flexión N-H de la

citosa y también puede observarse mejor en el control de metilación aunque ligeramente desplazada. La banda de  $1010\text{ cm}^{-1}$  en cambio, se muestra claramente en el grupo cáncer y ligeramente desplazada en los otros dos espectros, esta banda se atribuye a la vibración del enlace  $\text{C}_4 - \text{C}_5$  de la citosina.

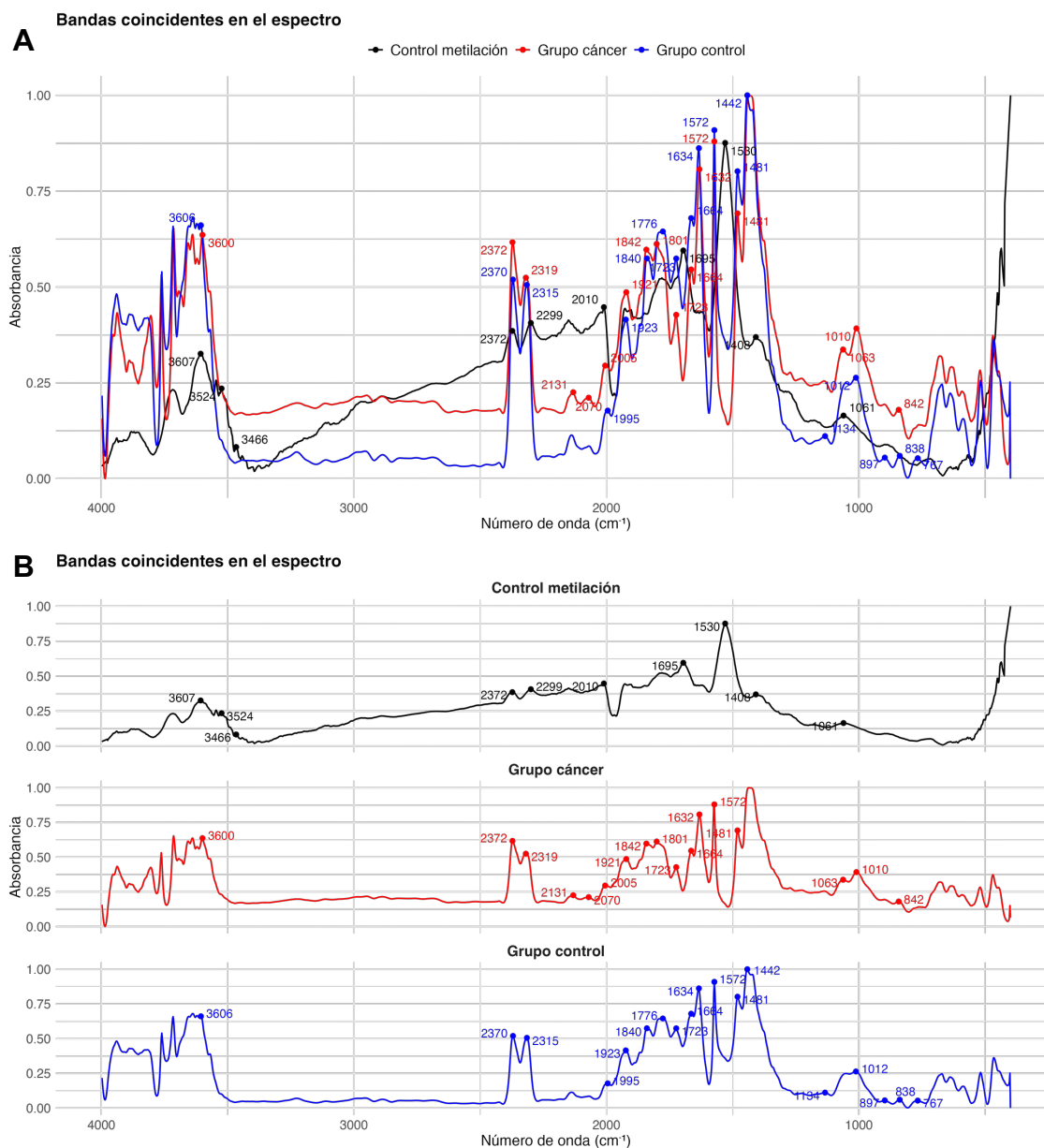


Figura 27. A: Espectros FTIR superpuestos de muestras de ADN y control de metilación. B: Espectros FTIR paralelos de muestras de ADN y control de metilación.

En la Figura 28 se muestran algunas de las bandas características de la molécula del ADN que están presentes en el espectro obtenido, como la de  $1060\text{ cm}^{-1}$ , considerada un marcador del ADN, así como bandas entre  $1020$  y  $820\text{ cm}^{-1}$ , consideradas como representativas de la región de la huella digital.

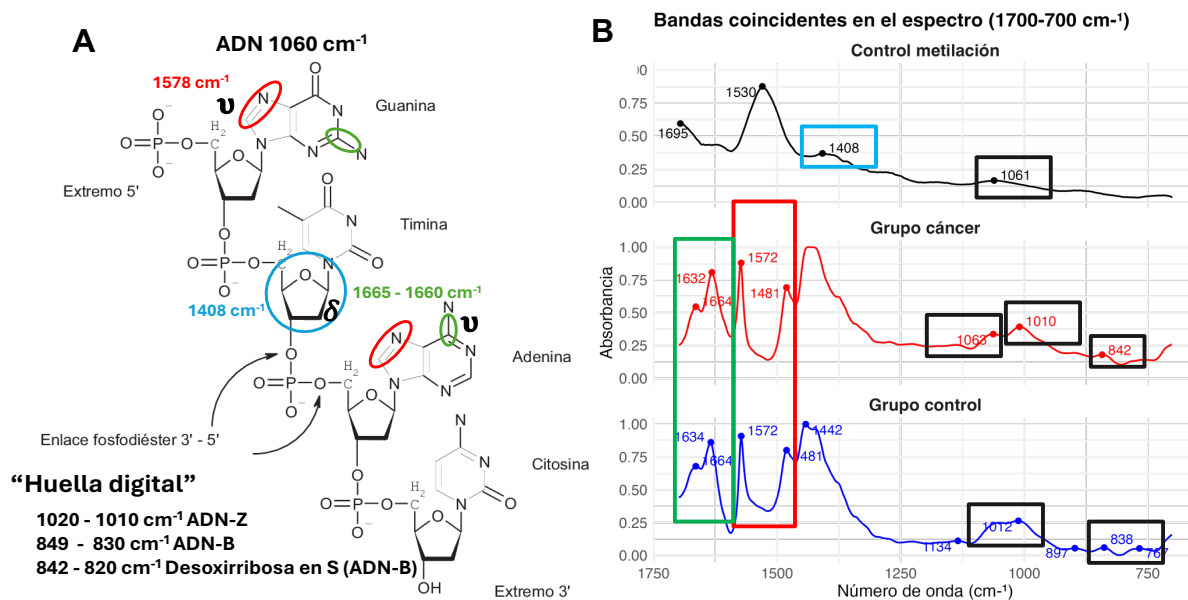


Figura 28. Bandas reportadas para enlaces presentes en el ADN. Los enlaces señalados en el panel A, se indican en los espectros del panel B con los mismos colores. Los recuadros negros señalan los marcadores espectrales de “huella digital”.

Bandas específicas de la citosina se pueden observar en la Figura 29. Entre  $3500$  y  $3300\text{ cm}^{-1}$  se reportan bandas correspondientes al estiramiento N-H de una amida primaria o secundaria, en  $1539\text{ cm}^{-1}$  correspondiente a la flexión N-H (banda Amida II) asociada a proteínas y compuestos nitrogenados, la de  $1695\text{ cm}^{-1}$  correspondiente al estiramiento C=O (Amida I), en  $1010\text{ cm}^{-1}$  asociada al estiramiento C-N común en heterociclos y amidas.

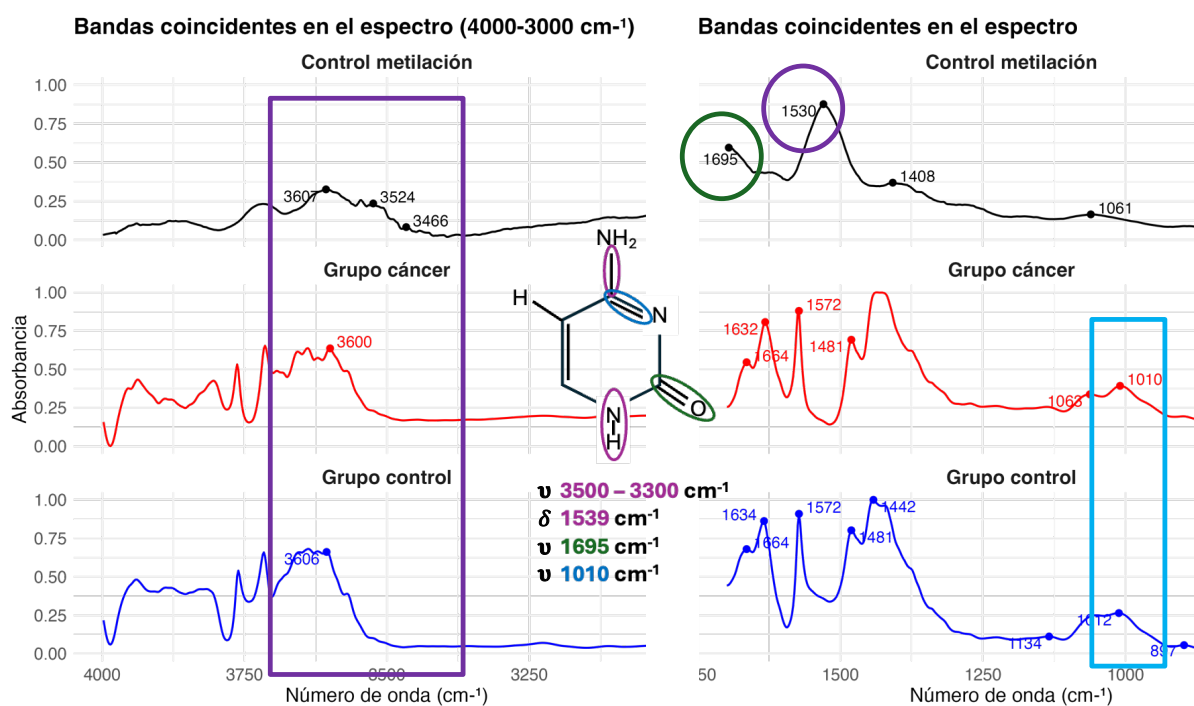


Figura 29. Bandas reportadas para enlaces de la citosina. Espectros promedio de muestras de ADN. Los enlaces señalados en la estructura química se indican en los espectros con los mismos colores.

También se buscaron bandas reportadas como específicas para la citosina metilada (Kelly et al., 2011), como 2937, 2883 y 1375 cm<sup>-1</sup>. Para esto se realizó un acercamiento en la región de interés, que se muestra en la Figura 30. Las bandas de 2937 y 2883 cm<sup>-1</sup> son características de las vibraciones del CH<sub>3</sub> y CH<sub>2</sub> de la citosina (Sarić et al., 2023), mientras que la banda de 1375 cm<sup>-1</sup> representa la flexión en el plano del CH<sub>3</sub> de la citosina y se considera una banda importante para analizar metilación del ADN aunque pueda presentarse a baja intensidad. Sin embargo, en el espectro obtenido solo se perciben las bandas de 2888 y 2886 cm<sup>-1</sup>, ambas presentes en los grupos de estudio y el control de metilación y como se anticipaba en la literatura, son menos intensas que las otras bandas características del metilo.

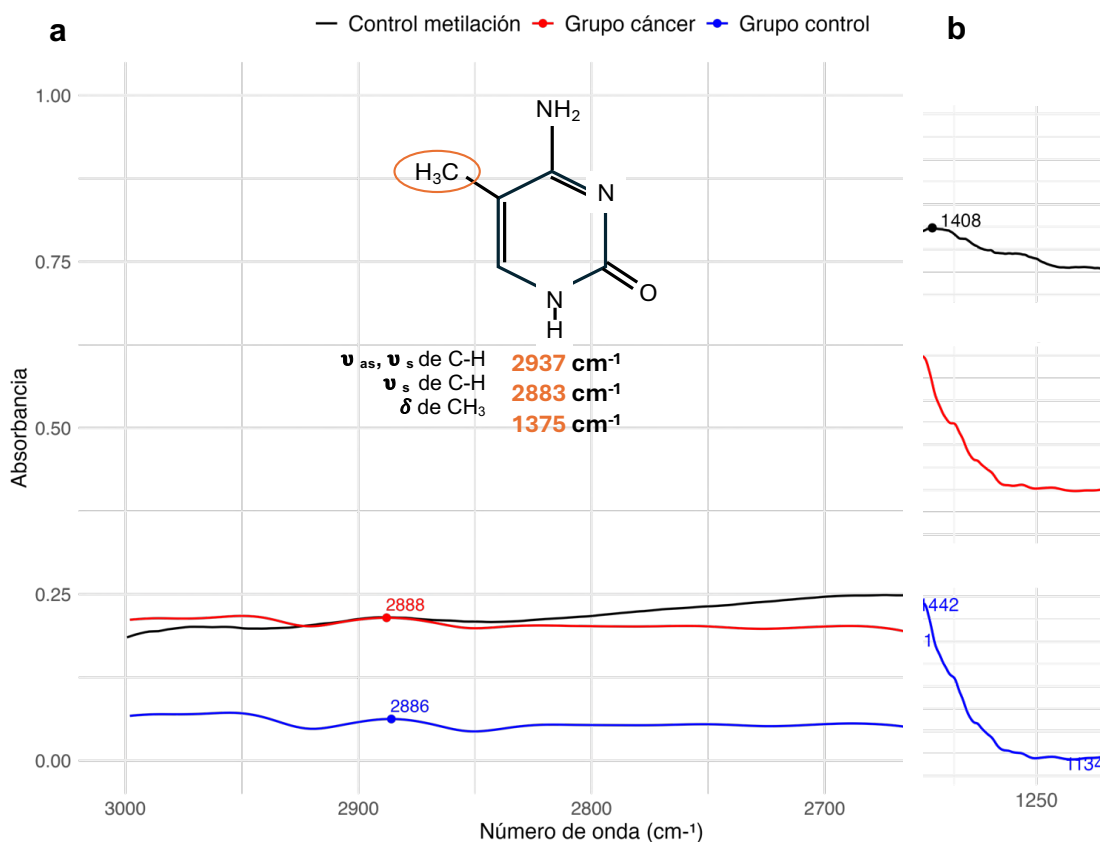


Figura 30. Bandas reportadas para la citosina metilada. Espectros promedio de muestras de ADN. A: acercamiento de la región 3000 a  $2700 \text{ cm}^{-1}$ , B: acercamiento de la región 1500 a  $1100 \text{ cm}^{-1}$ .

### 6.6.2 Análisis multivariado

Con el objetivo de reducir la dimensionalidad del conjunto de datos espectrales y poder identificar las regiones con mayor relevancia para la caracterización de las muestras, se llevó a cabo un Análisis de Componentes Principales (PCA), el primer resultado obtenido de este análisis fue un mapa de varianza explicada, que permite observar la proporción de varianza que explica cada uno de los componentes principales por región espectral, es decir, mientras el valor del coeficiente de determinación espectral ( $R^2$ ) sea más cercano a 1, tal componente principal captura la mayor variabilidad a esa región espectral en particular, lo contrario a un valor de

$R^2$  cercano a cero, que indica que esa región espectral no es representativa para ese componente principal o no captura gran parte de su variabilidad. La variabilidad de la región de 4000 a 3500  $\text{cm}^{-1}$  no fue bien captada por el primer y segundo componente (Figura 31), lo mismo sucede con la región de 1500 a 1200  $\text{cm}^{-1}$ , donde el primer componente tiene un coeficiente de determinación más alto solamente de 1350 a 1200  $\text{cm}^{-1}$  (Figura 32), una tercera región, de 1200 a 850  $\text{cm}^{-1}$  presentó una mejor distribución de los coeficientes de determinación de los tres primeros componentes principales (Figura 33), permitiendo seleccionarla como la región de interés principal para los posteriores gráficos del PCA.

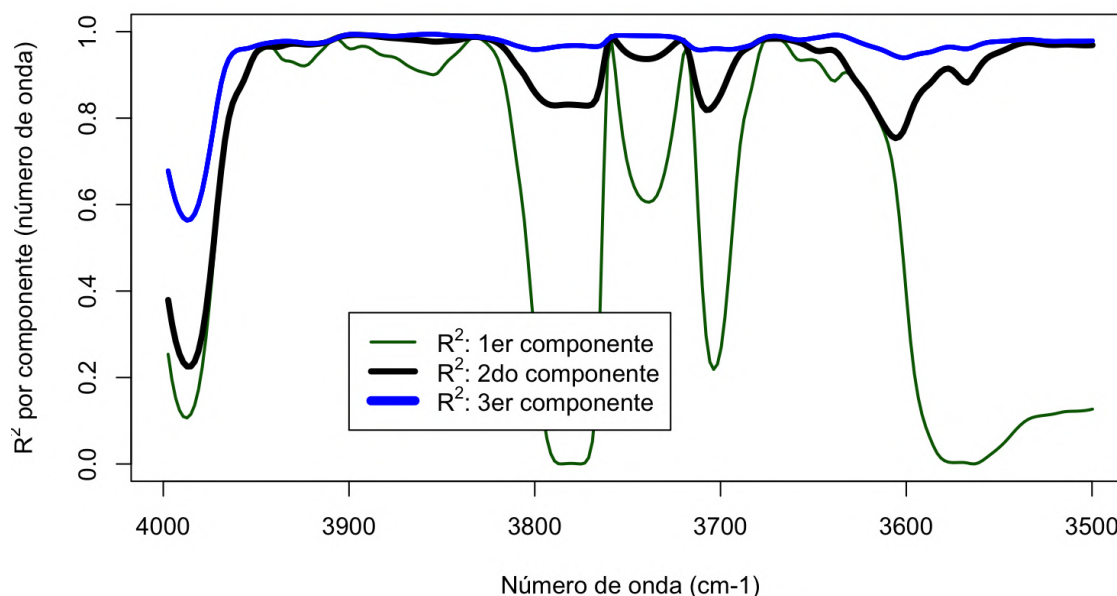


Figura 31. Mapa de varianza explicada por componentes principales en la región de 4000 a 3500  $\text{cm}^{-1}$  de espectros de muestras de ADN. El coeficiente de determinación espectral se indica en el eje Y, la región espectral con mayor contribución a cada componente principal será más cercano a 1.

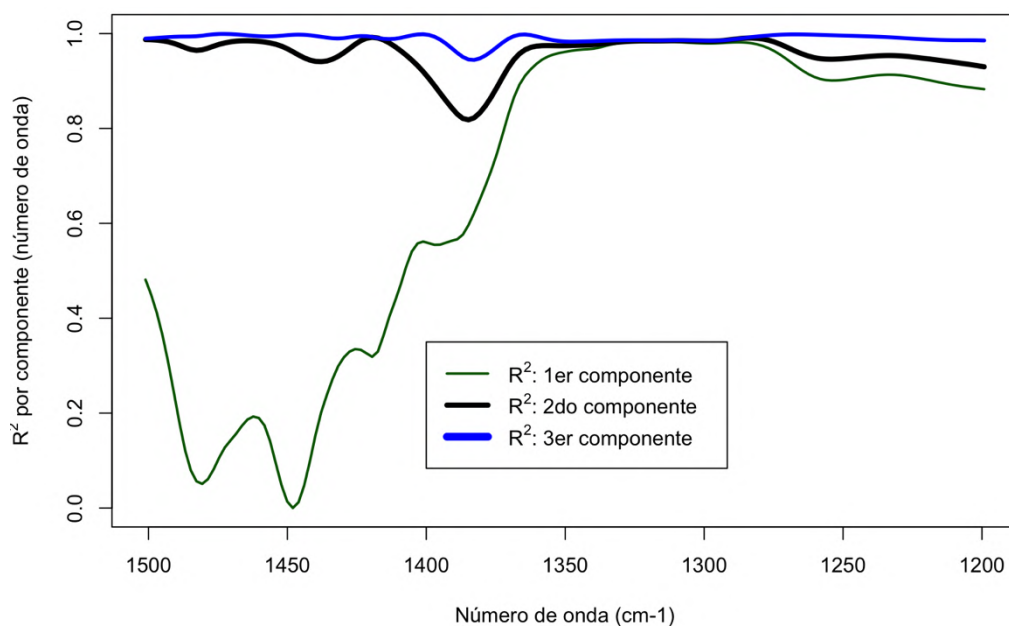


Figura 32. Mapa de varianza explicada por componentes principales en la región de 1500 a 1200  $\text{cm}^{-1}$  de espectros de muestras de ADN. El coeficiente de determinación espectral se indica en el eje Y, la región espectral con mayor contribución a cada componente principal será más cercano a 1.

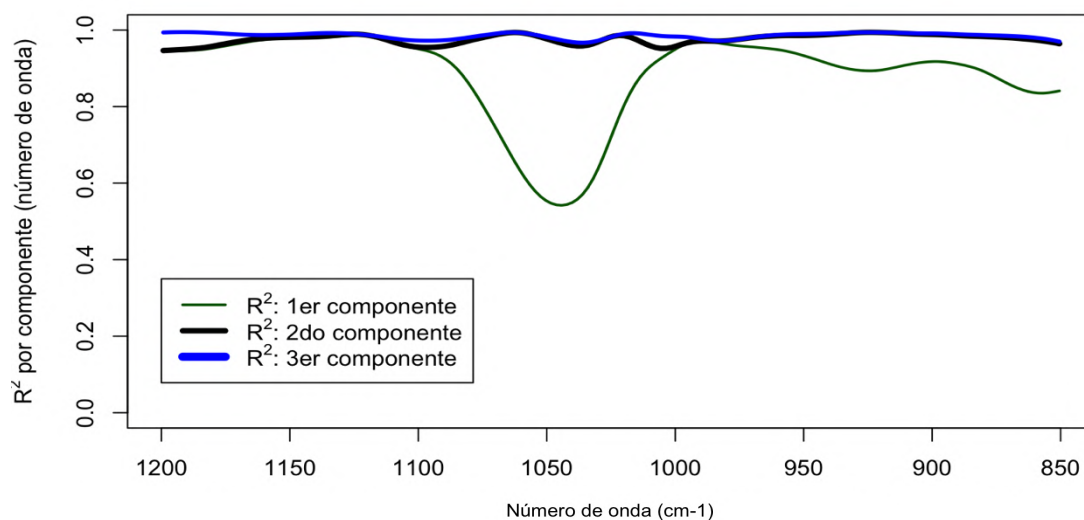


Figura 33. Mapa de varianza explicada por componentes principales en la región de 1200 a 850  $\text{cm}^{-1}$  de espectros de muestras de ADN. El coeficiente de determinación espectral se indica en el eje Y, la región espectral con mayor contribución a cada componente principal será más cercano a 1.



El primer y segundo componente principal del análisis para la región de 1200 a 850  $\text{cm}^{-1}$  explican en total el 97.6 % de la varianza del conjunto de datos espectrales (Figura 34), este gráfico permite descartar los componentes principales restantes debido a su baja contribución a la explicación de la varianza.

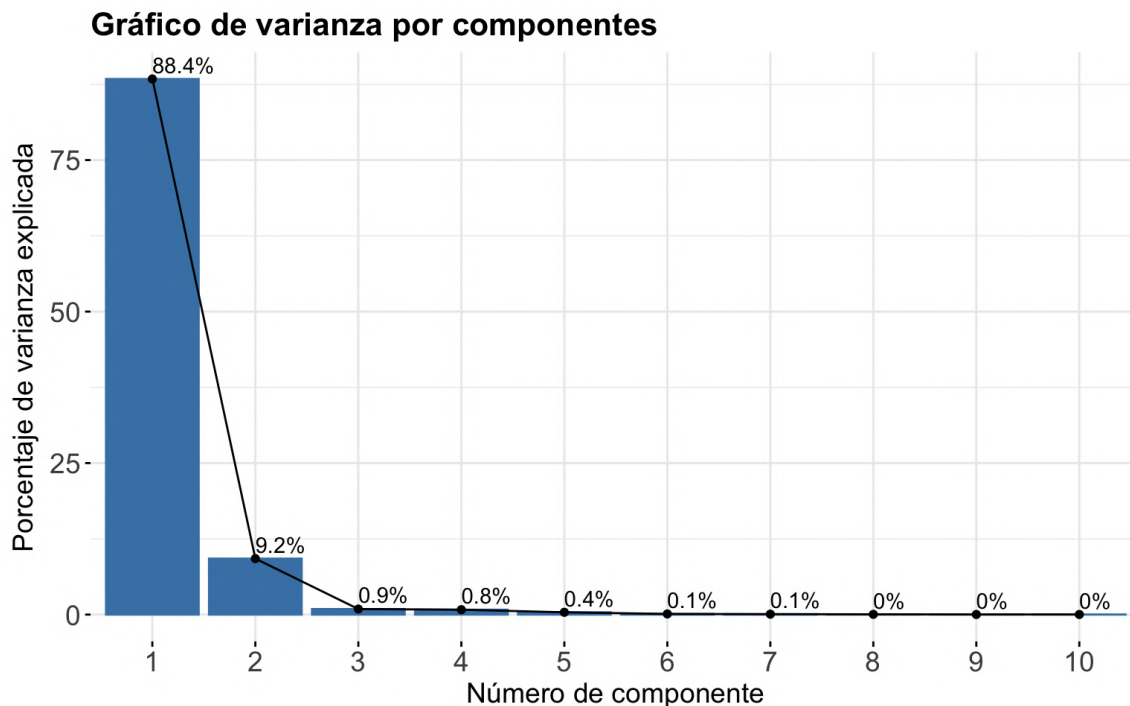


Figura 34. Porcentaje de varianza explicado por cada componente principal para la región de 1200 a 850  $\text{cm}^{-1}$ .

Se determinaron los números de onda con la mayor contribución al primer componente principal, de las cuales se pueden identificar las reportadas previamente en trabajos de análisis de ADN con espectroscopía infrarroja (Cuadro 11). Entre las bandas identificadas destacan aquellas asignadas a las conformaciones A y Z del ADN, ya que de acuerdo con los resultados obtenidos en función del tiempo de liofilización y por lo tanto, del nivel de humedad de las muestras de ADN, se puede proponer que lo observado en los espectros sean

bandas características de ADN que ha sufrido un cambio de conformación debido a la eliminación del agua.

Cuadro 11. Asignación de bandas de ADN

Número de onda (cm <sup>-1</sup> )	Asignación	Referencia
897 – 899	ADN-A	(Wood, 2016)
≈ 987	C-H, N <sub>3</sub> -C <sub>4</sub> -C <sub>5</sub> =C <sub>6</sub> (citosa)	(Kelly et al., 2011)
991-995	Vibración de la cadena principal ribosa-fosfato	(Banyay et al., 2003)
1010 - 1012	C <sub>4</sub> -C <sub>5</sub> (Citosa)	(Kelly et al., 2011)
1052- 1063	Vibración C-O de la ribosa, ADN-Z	(Wood, 2016)
1135 - 1138	C <sub>1</sub> -C <sub>2</sub> -O-C <sub>3</sub> de la ribosa	(Banyay et al., 2003)
≈ 1199	ADN-A, ribosa	(Banyay et al., 2003; Kelly et al., 2011)
1123	ADN-Z	(Kelly et al., 2011)
1135	Ribosa	(Kelly et al., 2011)
1160	Ribosa	(Wood, 2016)

Además, existen contradicciones acerca de algunas bandas que se consideran marcadores de una u otra conformación del ADN, como la banda de 899 - 890 cm<sup>-1</sup>, que se ha propuesto como representativa de la vibración del anillo de desoxirribosa y que se reporta con intensidad disminuida en el ADN-Z (Banyay et al., 2003), mientras que en otro trabajo se asigna la banda de 899 cm<sup>-1</sup> a ADN-A, la de 894 cm<sup>-1</sup> a ADN-B y la de 929 cm<sup>-1</sup> a ADN-Z (Wood, 2016). La banda de alrededor de los 1020 - 1010 cm<sup>-1</sup>, vibración de furanosa, también se reporta fuertemente incrementada en la forma ADN-Z (Kelly et al., 2011; Wood, 2016). En general, se ha sugerido que en la región de 1250 a 1000 cm<sup>-1</sup>, las vibraciones a lo largo de la cadena azúcar-fosfato dan lugar a lo que podrían considerarse marcadores de la conformación del esqueleto del ADN (Figura 35), mientras que en la región de los 1000 a 800 cm<sup>-1</sup> las vibraciones azúcar-fosfato dan lugar a marcadores del plegamiento de los azúcares, es decir, de tipo N y S, para los azúcares con plegamiento de tipo N, se reportan tres marcadores espectrales, 882 – 877 cm<sup>-1</sup>, 865 – 860 cm<sup>-1</sup> y 815 – 802 cm<sup>-1</sup>, mientras que para el plegamiento de

tipo S hay una banda ancha entre  $842 - 820 \text{ cm}^{-1}$ . La principal importancia de esta región para los fines del presente trabajo, radica en el antecedente de que la metilación de la citosina en la doble hélice del ADN ocasiona un desdoblamiento sutil de la banda de  $842 - 820 \text{ cm}^{-1}$  del plegamiento tipo S (Figura 36), sugiriendo la coexistencia de dos plegamientos del azúcar dentro de las secuencias metiladas (Banyay et al., 2003; Wood, 2016).

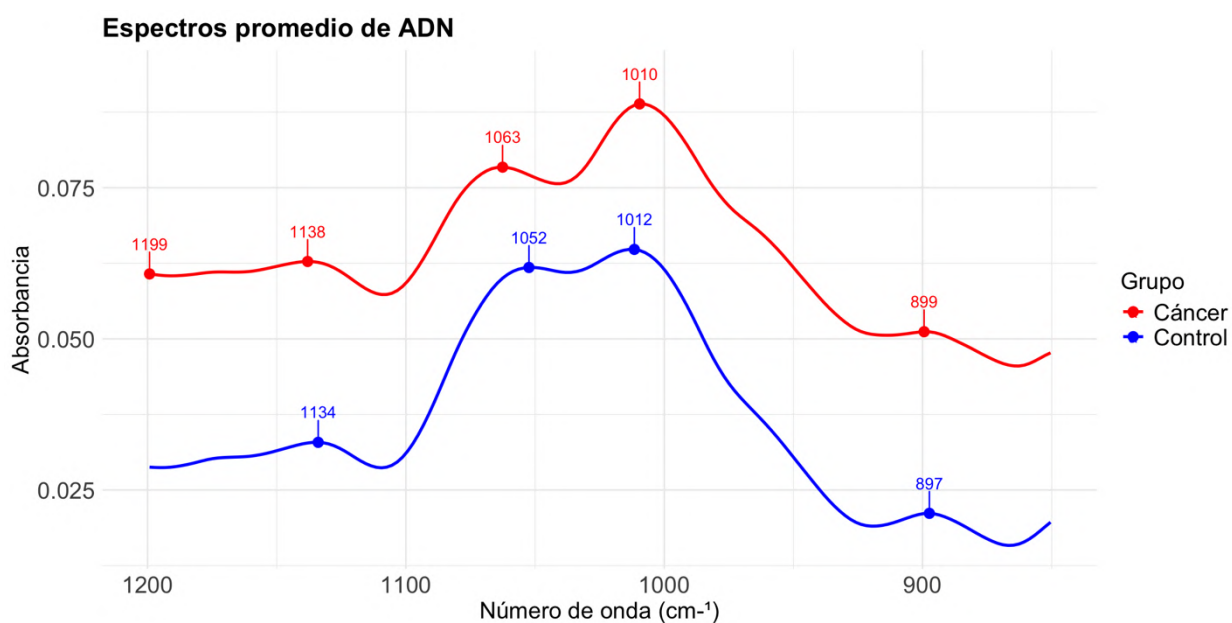


Figura 35. Acercamiento de la región  $1200 - 850 \text{ cm}^{-1}$  del espectro de ADN.

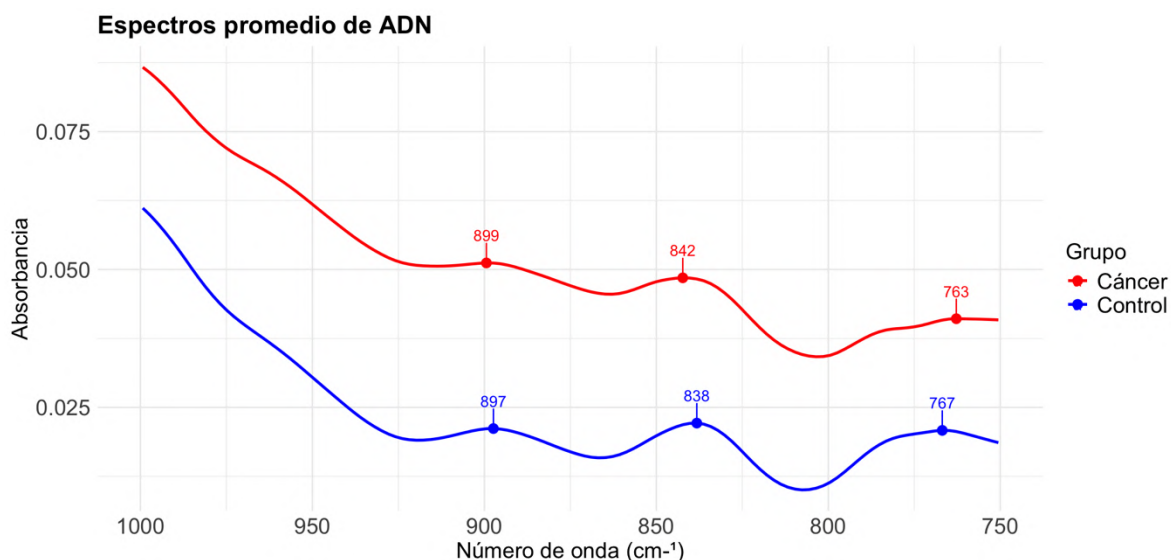


Figura 36. Acercamiento de la región 1000 - 750 cm<sup>-1</sup> del espectro de ADN.

El gráfico biplot resultante del PCA permite observar que de manera generalizada, las muestras pertenecientes al grupo control se encuentran cercanas entre sí, a excepción de la muestra A110, sin embargo, este agrupamiento no se observa en el grupo de muestras de cáncer de pulmón, que si bien la mayoría de ellas se encuentra separada del grupo control, están distribuidas de manera muy dispersa en el plano, y dos de ellas se encuentran más cercanas al grupo control que al resto de muestras de cáncer de pulmón (Figura 37). Este comportamiento puede deberse en gran parte a la heterogeneidad entre las muestras de cáncer de pulmón, lo que pudiera coincidir con las diferencias clínico-patológicas propias de cada individuo, ya que a pesar de que se muestreó un grupo homogéneo en cuanto a tipo de cáncer de pulmón y etapa, cada persona tenía un historial clínico muy particular, que incluía diferentes recaídas, comorbilidades y tratamientos específicos para síntomas tanto del cáncer de pulmón como de enfermedades concomitantes al momento de tomar la muestra sanguínea.

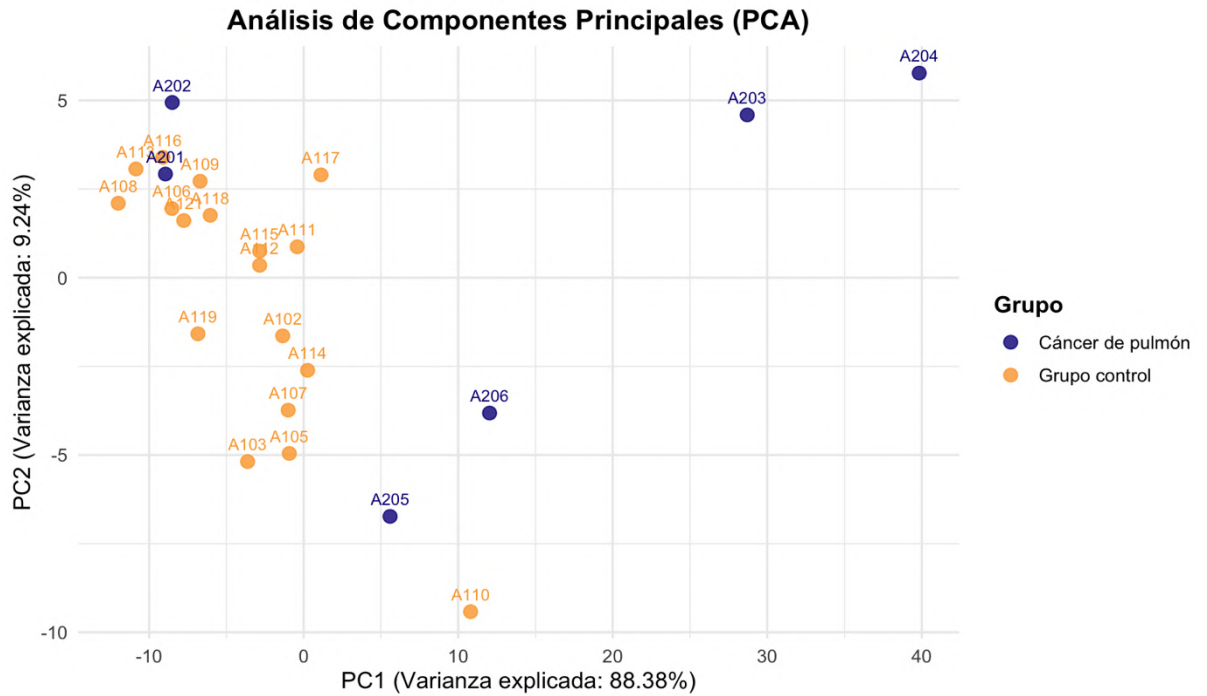


Figura 37. Biplot de PCA de espectros de ADN región 1200 - 850  $\text{cm}^{-1}$ .

Para asegurar el agrupamiento correcto de las muestras, una vez que se redujo la dimensionalidad del conjunto de datos espectrales y se generó el biplot, se aplicó un algoritmo de agrupamiento llamado DBSCAN. Los métodos de agrupamiento dividen conjuntos de datos de gran dimensión en múltiples grupos o categorías, de acuerdo con la similitud encontrada. Entre los métodos de agrupamiento no supervisado (sin especificar el agrupamiento real de los datos) más empleados se encuentran el K-medias y el DBSCAN, el primero es un algoritmo que toma puntos de datos como entrada y los agrupa en un número determinado de categorías o grupos ( $k$ ), primero de manera aleatoria y posteriormente de manera iterativa moviendo los puntos de datos al centro del conglomerado más cercano a ellos (Liu et al., 2024). El Agrupamiento Espacial Basado en Densidad de Aplicaciones con Ruido (DBSCAN) forma los grupos en función de las medidas de densidad, para lo cual primero se deben determinar dos parámetros,  $\epsilon$  (eps), el parámetro de radio y

*MinPts*, la densidad de vecindad, que se refiere al número mínimo de puntos que se encuentran dentro de un radio especificado (Liu et al., 2024). La ventaja de utilizar DBSCAN radica en que es un método que incluye el ruido y permite detectar valor atípicos, además, en conjuntos de datos de alta dimensión como los espectrales, donde los grupos no siempre presentan formas regulares, DBSCAN es mejor opción que K-medias, que por el contrario, asume que los grupos presentan la clásica forma esférica (Andriyani & Puspitarani, 2022; Chen, 2020).

El agrupamiento mediante DBSCAN permitió confirmar que las muestras no se clasifican en los dos grupos de estudio esperados (Figura 38), y que además, 4 de las 6 muestras del grupo de cáncer de pulmón son detectadas como ruido (grupo 0) por el algoritmo, mientras que las muestras del grupo control son divididas en 2 grupos (grupo 1 y 2). Este agrupamiento inconsistente con el verdadero origen de las muestras puede obedecer a las variaciones inherentes a la preparación de la muestra, su composición, así como a pequeñas variaciones en el proceso de adquisición espectral. En el caso particular de las muestras de cáncer de pulmón, como se observó en el biplot anterior, pueden ser clasificadas como ruido debido a las marcadas diferencias entre una y otra muestra.

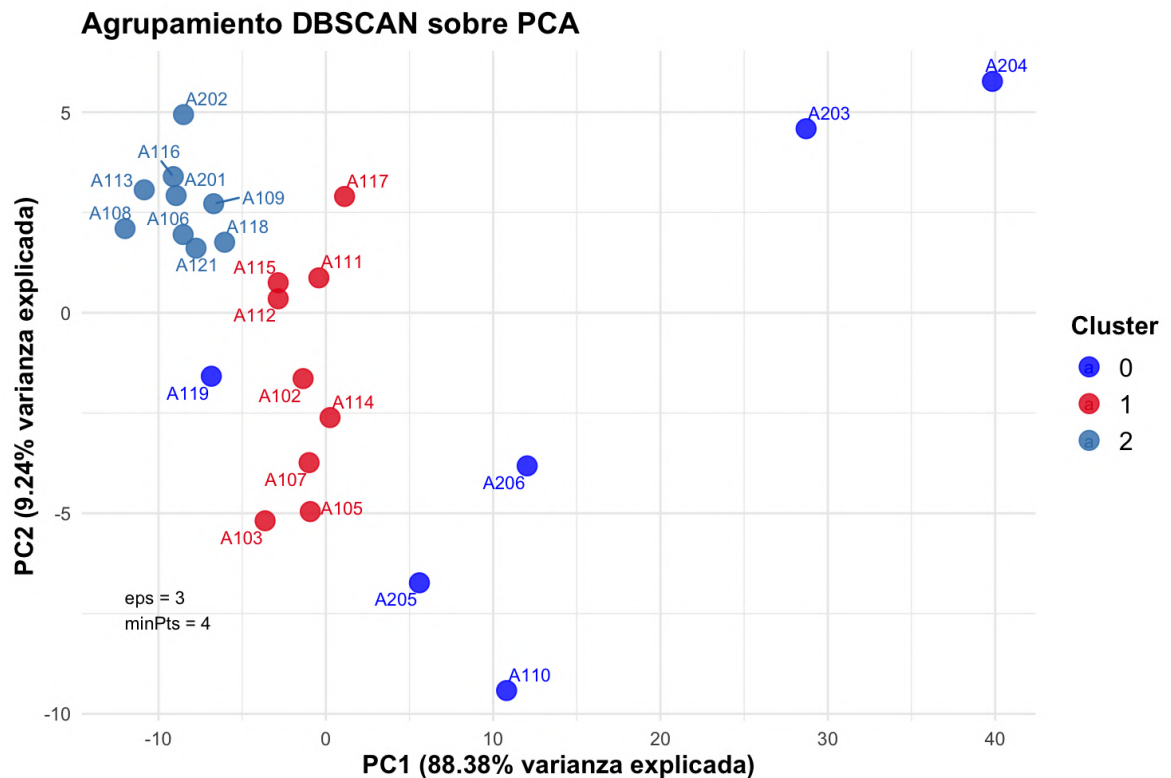


Figura 38. Biplot de PCA empleando el algoritmo de agrupamiento DBSCAN.

Es recomendable evaluar la calidad del agrupamiento obtenido con DBSCAN mediante dos índices: el Índice de Davies-Bouldin (DBI) y el Índice de Rand Ajustado (ARI). DBI se utiliza para evaluar la validez de un grupo, maximizando la distancia entre los grupos y minimizando las distancias entre puntos de un grupo. Mientras más bajo sea el valor de DBI, mejor se evalúa el agrupamiento (Muningsih et al., 2023). El ARI se emplea también en la validación de los agrupamientos, comparando la concordancia entre la clasificación por el proceso de agrupamiento y la clasificación de acuerdo con criterios externos, valores bajos de ARI indican una mala clasificación y valores altos una correcta clasificación (Santos & Embrechts, 2009). Se obtuvo un DBI = 0.5959, (un valor de DBI  $\geq 0.5$  y  $< 1$  indica cierta superposición entre los grupos, pero se consideran aceptables) y un ARI = 0.1616 (ARI = 0 indica que los agrupamientos no tienen relación con los grupos reales), lo

que podía esperarse desde los resultados obtenidos en el biplot de PCA. Para analizar el comportamiento de la muestra A110, que se observa alejada del plano en el PCA, se presenta en la Figura 39 cómo se observa su espectro individual en comparación con los espectros de las muestras restantes. A simple vista no se observan variaciones marcadas, sin embargo, en la región de 1800 a 1500  $\text{cm}^{-1}$  se observan picos más elevados, lo que podría explicar por qué está siendo apartada del resto de muestras del grupo control en el PCA.

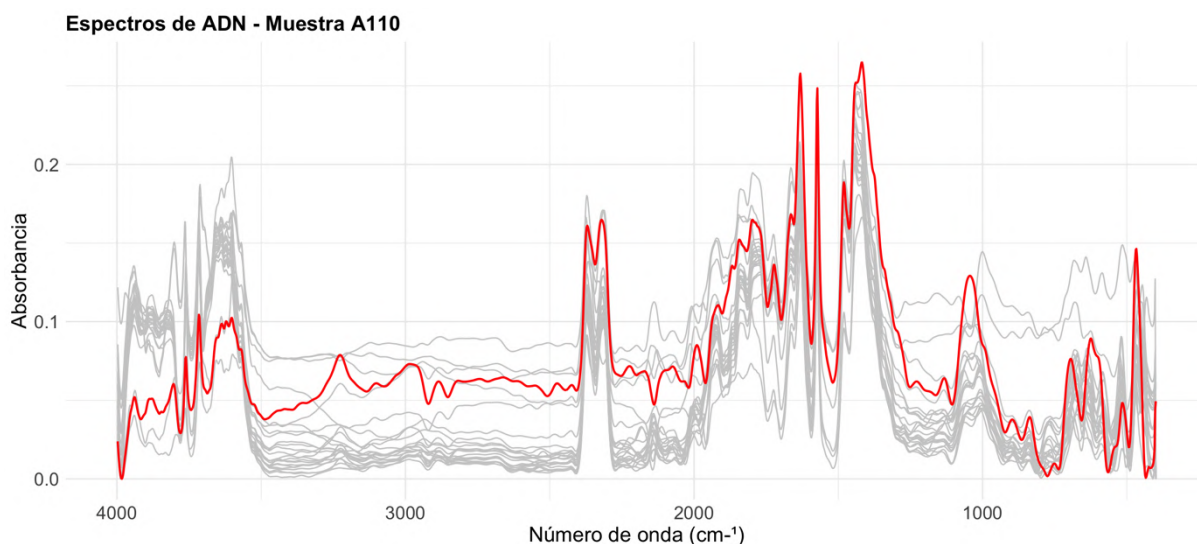


Figura 39. Espectro de muestra A110 resaltado en rojo para observar diferencias con los espectros restantes del grupo (gris).

El agrupamiento jerárquico de muestras también permitió observar de manera más organizada la clasificación de las muestras (Figura 40). Con los resultados obtenidos hasta ahora se puede plantear una posible relación del agrupamiento obtenido con el porcentaje de metilación global del ADN, debido a que las muestras del grupo de cáncer de pulmón que se encuentran en las ramas del grupo control (A201 y A202) fueron las que presentaron los porcentajes de metilación más elevados del grupo de cáncer (1.018 y 1.248 respectivamente), ya que todas las muestras del grupo control presentaron porcentajes de metilación global por mayores a 0.9. Sin embargo, la muestra A110, del grupo control, presenta un



porcentaje similar al resto de su grupo, por lo que su clasificación podría explicarse por otros factores, como la alteración de la muestra antes de su adquisición espectral o errores en la medición, lo que también pudiese explicar su variación espectral (Figura 39).

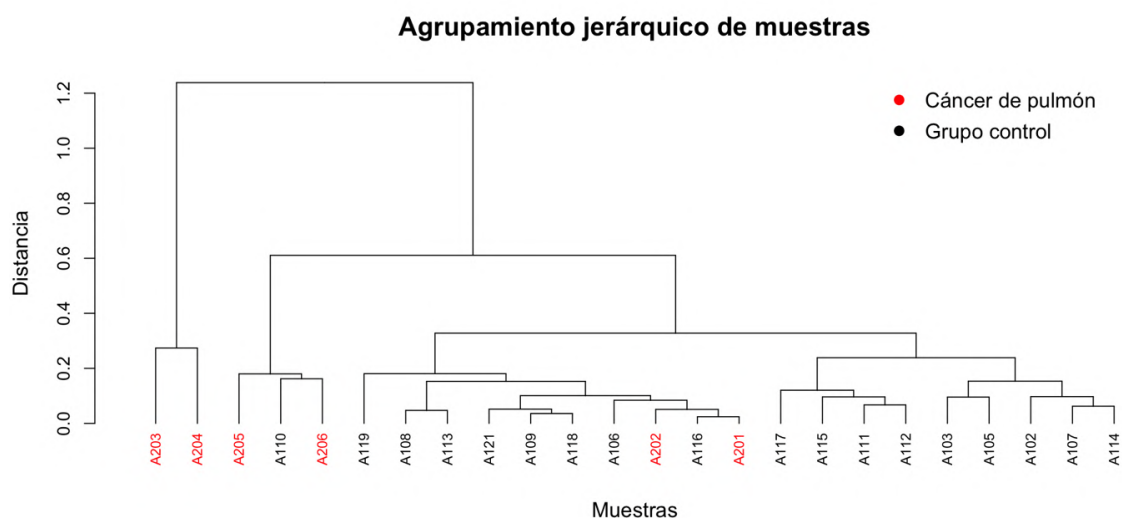


Figura 40. Dendrograma de las muestras de ADN. El eje Y indica la distancia de las muestras basado en sus perfiles espectrales, las ramas agrupan las muestras en función de la similitud, ramas que se fusionan a mayor altura reflejan mayor diferencia.

### 6.6.3 Análisis por regiones espectrales

Para evaluar las diferencias de absorbancia por regiones espectrales, se establecieron regiones de acuerdo con lo reportado previamente en estudios de ácidos nucleicos, células y tejidos con FTIR, se calcularon las absorbancias por áreas de banda de estas regiones y posteriormente se compararon mediante t de Student para buscar diferencias entre los grupos de estudio, se hicieron dos pruebas con regiones que abarcaban distintos números de onda (Figura 41 y 42). En la primer prueba se utilizó como región de Metilación el rango de 1470 a 1250  $\text{cm}^{-1}$ .

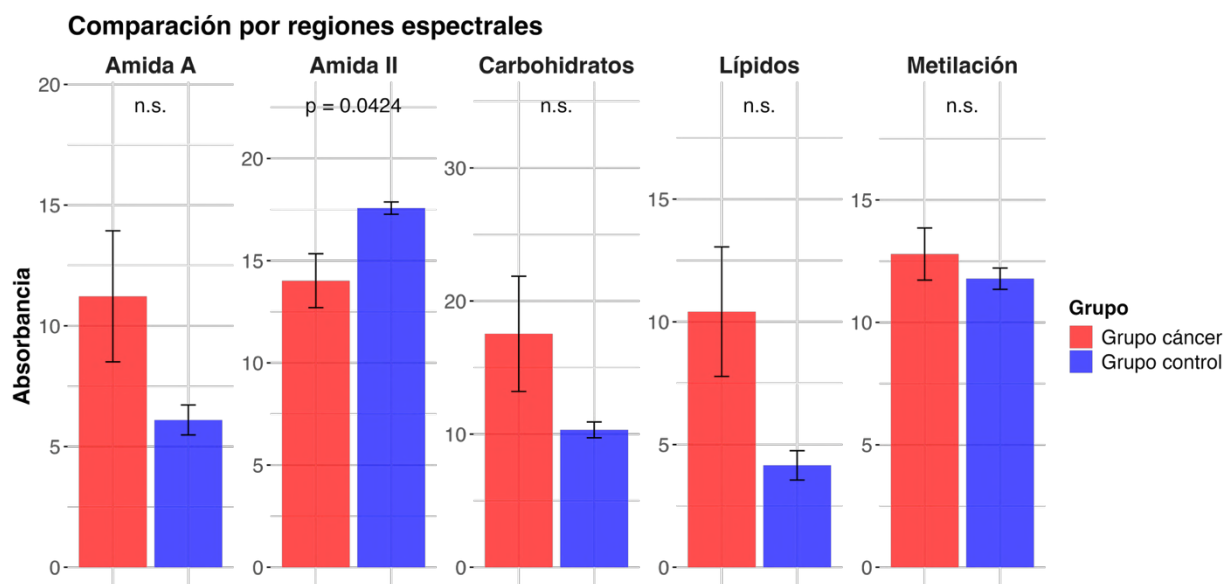


Figura 41. Comparación de absorbancias por regiones espectrales. Las barras representan la media de intensidad para cada región espectral: Amida A (3600 – 3200  $\text{cm}^{-1}$ ), Lípidos (3100 – 2700  $\text{cm}^{-1}$ ), Amida II (1800- 1500  $\text{cm}^{-1}$ ), Carbohidratos (1200 – 600  $\text{cm}^{-1}$ ), **Metilación (1470 – 1250  $\text{cm}^{-1}$ )**. Los bigotes indican el error estándar de la media. Prueba t de Student,  $p < 0.05$ , n.s. = no significativo estadísticamente.

La comparación que presenta diferencias estadísticamente significativas ( $p = 0.0468$ ), fue en la región denominada Amida II (Figura 41), que abarca de los 1800 a los 1600  $\text{cm}^{-1}$ . Aunque esta región se denomina así normalmente para cualquier tipo de muestras biológicas, en el caso particular de muestras con ácidos nucleicos aislados, esta región recibe contribuciones de bandas como 1710  $\text{cm}^{-1}$ , 1689  $\text{cm}^{-1}$ , 1632  $\text{cm}^{-1}$ , que se atribuyen a vibraciones de los enlaces carbono que conforman las bases nitrogenadas del ADN (Kelly et al., 2011).

La región denominada Metilación, recibe contribuciones de 1438  $\text{cm}^{-1}$ , 1413  $\text{cm}^{-1}$ , bandas relacionadas con la estructura Z del ADN, y 1363  $\text{cm}^{-1}$ , 1238  $\text{cm}^{-1}$ , atribuidas a vibraciones de los enlaces de las citosinas del ADN. En el primer análisis en esta

región no hay diferencias significativas, además, puede notarse que el nivel de absorbancia en esta región es mayor en el grupo cáncer, lo que se contrapone con los resultados de hipometilación obtenidos previamente con la ELISA.

La elevada absorbancia en las regiones que indican proteínas y lípidos, podrían estar reflejando la presencia de residuos de estas moléculas después de la extracción, lo que también puede cuantificarse mediante una absorbancia  $< 1.7$  de la relación 260/280 con el espectrofotómetro, sin embargo, en las pruebas de calidad realizadas, las muestras en su mayoría presentaron una absorbancia mayor a 1.7, además, la totalidad de las muestras del grupo de cáncer son las que se obtuvieron con pureza óptima, mientras que en FTIR este grupo es el que presentaría niveles más altos de absorbancia en estas regiones (Figura 41 y 42).

En el segundo análisis, para Metilación se emplearon subregiones, 3610 – 3600, 1700 – 1500, 1200 - 800  $\text{cm}^{-1}$ , con el fin de seleccionar solo aquellas bandas que contribuyeran de manera significativa a la detección de la metilación y que esta región no incluyera ruido que alterara la suma de las absorbancias. Con este cambio en la región puede observarse un resultado muy diferente al del primer análisis. En la Figura 42 la absorbancia para esta región disminuyó para el grupo cáncer, siendo congruente ahora con lo obtenido previamente con la ELISA.

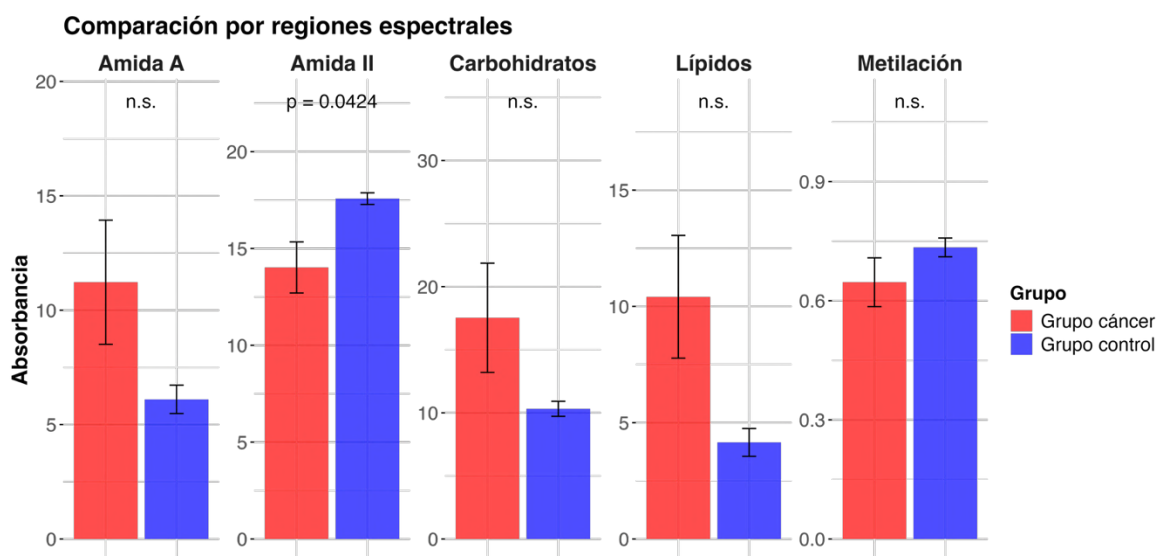


Figura 42. Comparación de absorbancias por regiones espectrales. Las barras representan la media de intensidad para cada región espectral: Amida A ( $3600 - 3200 \text{ cm}^{-1}$ ), Lípidos ( $3100 - 2700 \text{ cm}^{-1}$ ), Amida II ( $1800 - 1500 \text{ cm}^{-1}$ ), Carbohidratos ( $1200 - 600 \text{ cm}^{-1}$ ), **Metilación ( $3610 - 3600$ ,  $1700 - 1500$ ,  $1200 - 800 \text{ cm}^{-1}$ )**. Los bigotes indican el error estándar de la media. Prueba t de Student,  $p < 0.05$ , n.s. = no significativo estadísticamente.

Para explorar la posible correlación entre las técnicas utilizadas, se compararon los valores de absorbancia obtenidos mediante la ELISA de metilación y la absorbancia obtenida con FTIR en las regiones delimitadas que se consideraron aportaban información sobre la metilación del ADN.

En ambos casos la correlación es muy débil, siendo la región de la Amida II la que presenta un coeficiente de correlación más alto (Figura 43) sin embargo, no se considera aceptable. La región Metilación presentó una relación positiva y más fuerte en comparación con los valores del porcentaje de metilación global obtenidos con la ELISA (Figura 44), teniendo un coeficiente de  $r = 0.501$ , lo que se considera una correlación moderada.

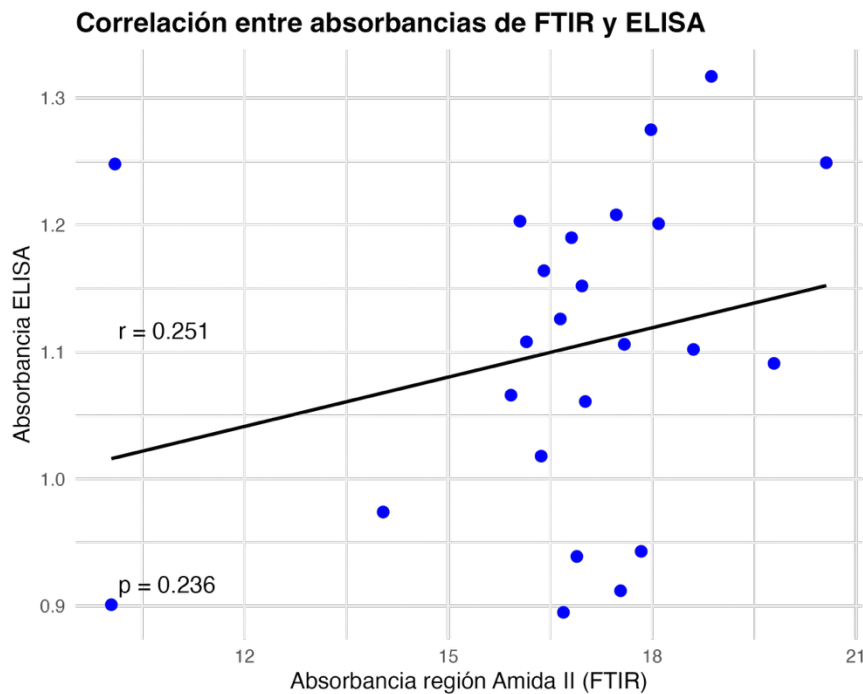


Figura 43. Correlación entre absorbancias por ELISA y FTIR en región Amida II de muestras de ADN aislado. Los puntos azules representan muestras individuales, la línea negra corresponde al ajuste por regresión lineal simple. Coeficiente de correlación de Pearson indicado con  $r$ ,  $p < 0.05$ .

Esto confirma la importancia de la selección adecuada de las bandas o regiones que se utilizarán en los análisis posteriores, además de la capacidad de la técnica para brindar información sobre la estructura del ADN. Si bien la correlación no es fuerte, puede establecerse que bajo estas condiciones ambas técnicas reportan resultados similares.

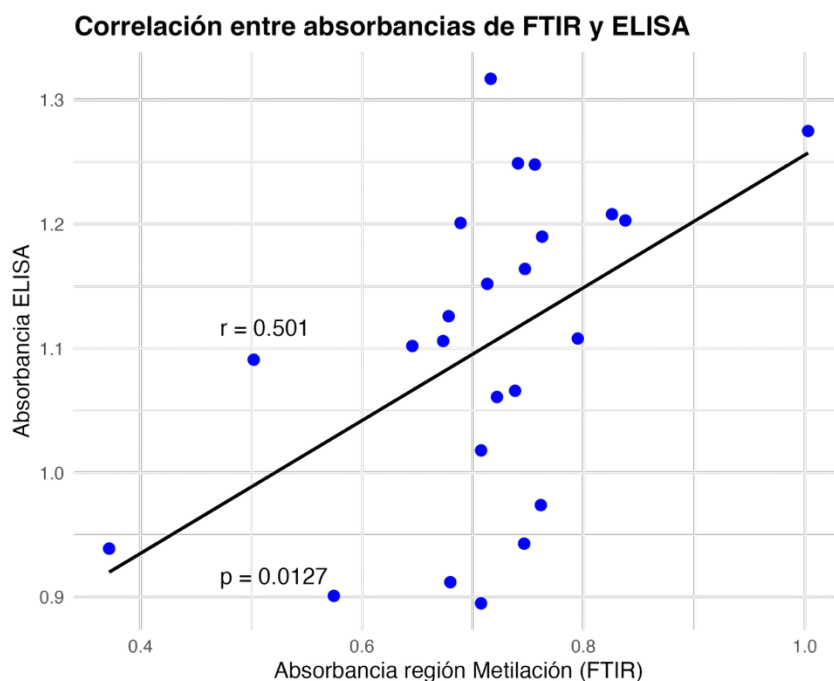


Figura 44. Correlación entre absorbancias por ELISA y FTIR en región Metilación (3610 – 3600, 1700 – 1500, 1200 - 800  $\text{cm}^{-1}$ ) de muestras de ADN aislado. Los puntos azules representan muestras individuales, la línea negra corresponde al ajuste por regresión lineal simple. Coeficiente de correlación de Pearson indicado con  $r$ ,  $p < 0.05$ .

#### 6.6.4 Pre-procesamiento de los espectros de suero

Todas las muestras de suero se liofilizaron por 16 horas, obteniendo un polvo de color dorado. El procesamiento de los espectros del suero se realizó en la misma secuencia que los de ADN, con la única diferencia del tipo de corrección de la línea base aplicada, que en este caso fue de tipo polinomial de segundo grado, ya que fue la que mejor logró corregir las desviaciones de la línea base (Figura 45).

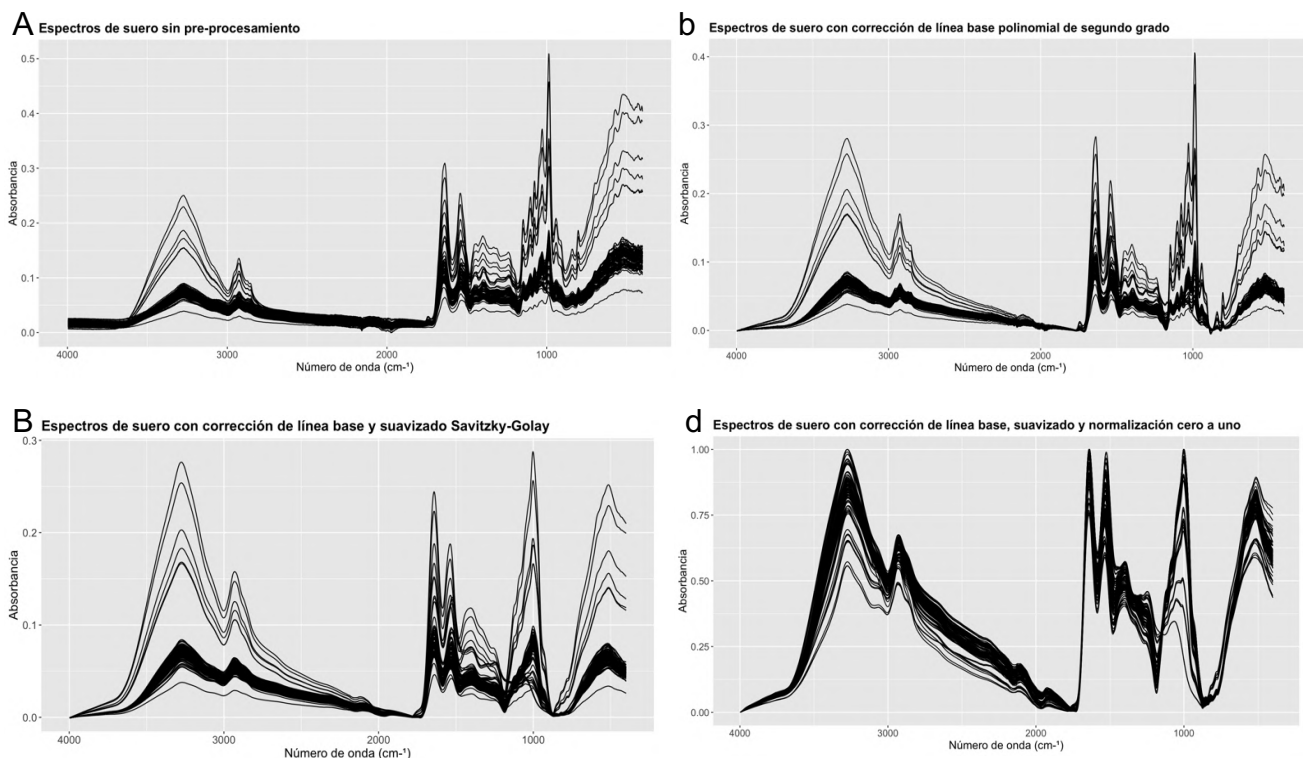


Figura 45. Pre-procesamientos de los espectros de suero a las 16 horas de liofilización.

Al graficar los espectros promedio de ambos grupos no se perciben grandes diferencias, pero notablemente, existen más bandas en la región de los 2000 a 1000  $\text{cm}^{-1}$  (Figura 46). En este tipo de muestra, formada por una matriz más compleja que las muestras de ADN aislado, se esperaría encontrar un mayor número de bandas a lo largo de toda la región del infrarrojo medio, en respuesta a la presencia de una gran cantidad de metabolitos presentes en el suero humano y por ende, a una mayor variedad de grupos funcionales.

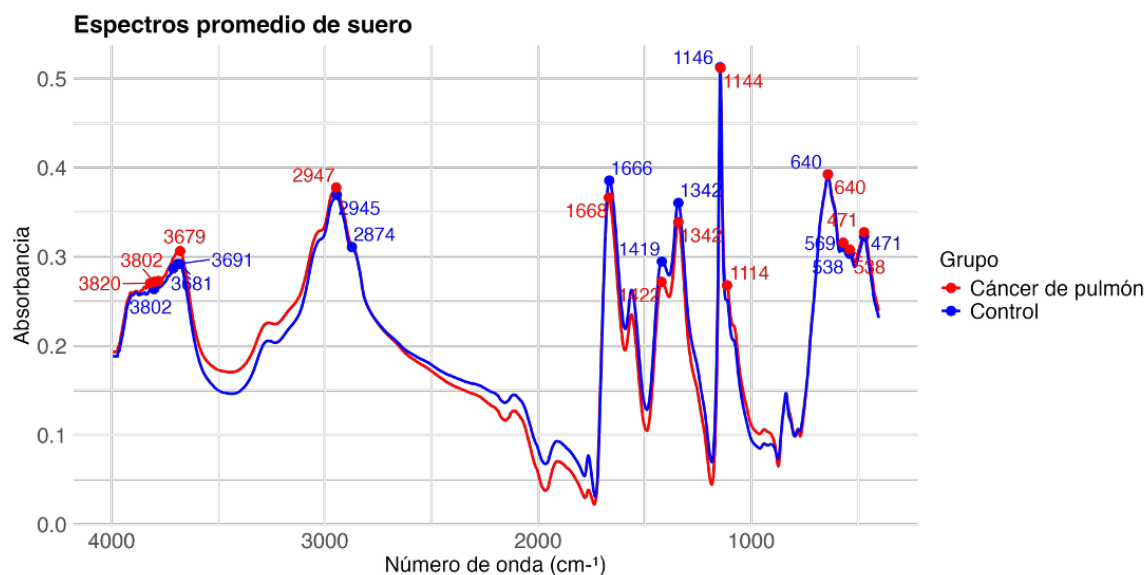


Figura 46. Espectros promedio del suero. Bandas más representativas indicadas.

#### 6.6.5 Análisis multivariado

Se llevaron a cabo los mismos análisis multivariados para las muestras de suero, comenzando con el PCA y la obtención del gráfico de coeficiente de determinación espectral, que puede observarse más heterogéneo que el del ADN, ya que en este caso el primer componente principal no mantiene una continuidad en el coeficiente de determinación a lo largo de la región espectral (Figura 47), indicando que en este tipo de muestras, la variabilidad es incluso mayor.



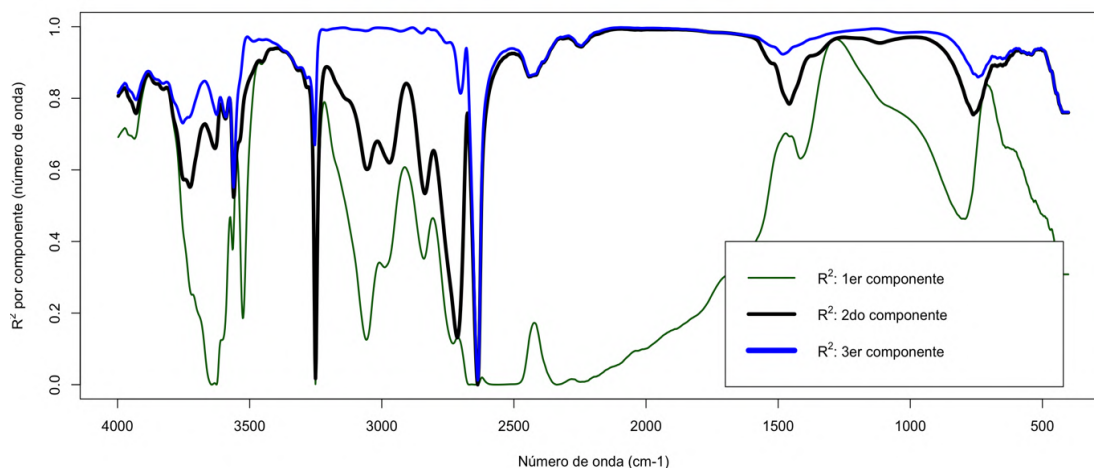


Figura 47. Mapa de varianza explicada por componentes principales en la región de 4000 a 500  $\text{cm}^{-1}$  de espectros de muestras de suero. El coeficiente de determinación espectral se indica en el eje Y, la región espectral con mayor contribución a cada componente principal será más cercano a 1.

Con la obtención de este primer gráfico, la delimitación de una región de interés en específico es más complicada que con los espectros del ADN, ya que centrarse en una sola región puede ocasionar el descarte de la contribución de otros grupos funcionales que se detecten en el resto del espectro y que aporten a la visión integral de la composición de una muestra de matriz compleja como lo es el suero humano.

Los dos primeros componentes explican un total del 84.5 % de la varianza, pero a diferencia de lo obtenido con los espectros del ADN, en este caso no presentan una gran diferencia entre el porcentaje de varianza explicado entre ellos (Figura 48).

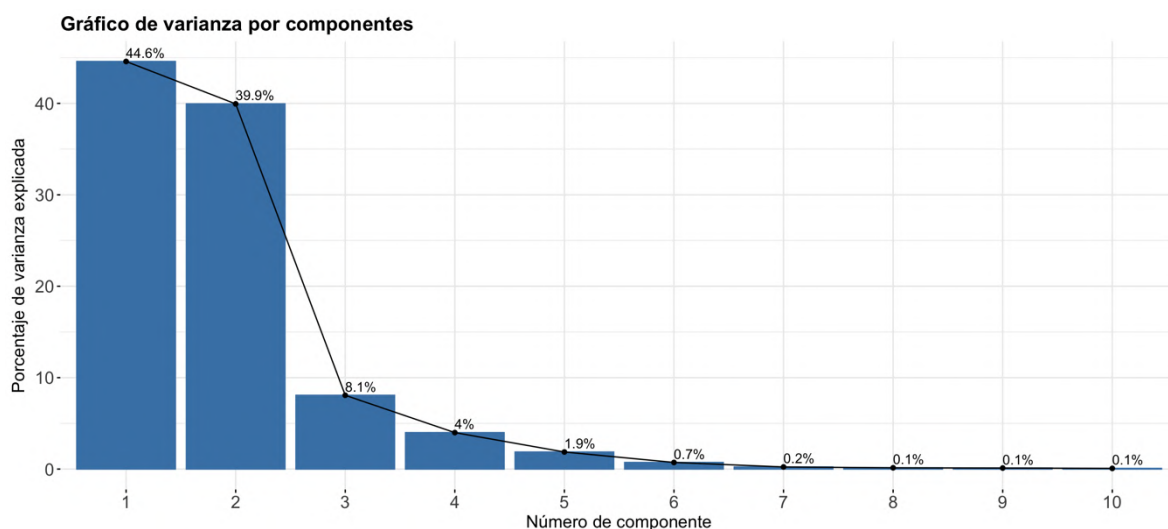


Figura 48. Porcentaje de varianza explicado por cada componente principal para todo el espectro del suero.

Al obtener el biplot del espectro del suero puede observarse que tampoco existe la tendencia al agrupamiento estricto de las muestras (Figura 49), si bien las pertenecientes al grupo de cáncer de pulmón se encuentran cercanas unas de otras, las del grupo control se encuentran más dispersas.

Esta dispersión entre muestras del grupo control es más evidente en el suero ya que este fluido es capaz de capturar distintos procesos bioquímicos que pueden variar en gran medida incluso de un individuo a otro.

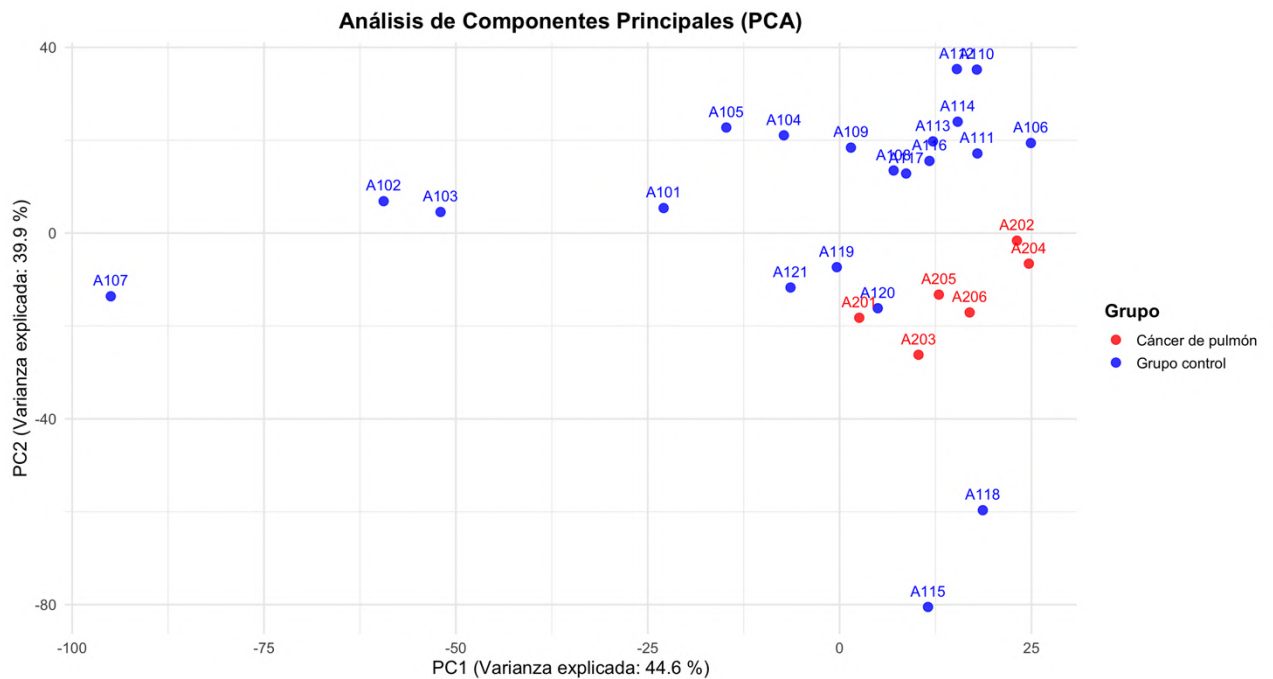


Figura 49. Biplot de PCA de espectros de suero.

Posteriormente se aplicó el algoritmo de agrupamiento DBSCAN. Esta vez 9 muestras fueron detectadas como ruido (grupo 0), mientras que 3 muestras del grupo cáncer de pulmón fueron clasificadas juntas en el grupo 2 (Figura 50). La validación del algoritmo se realizó con el DBI (0.5327, es decir, con cierta superposición entre los grupos, pero aceptable) y con el ARI (0.1172, que indica mezcla de clases entre los agrupamientos).

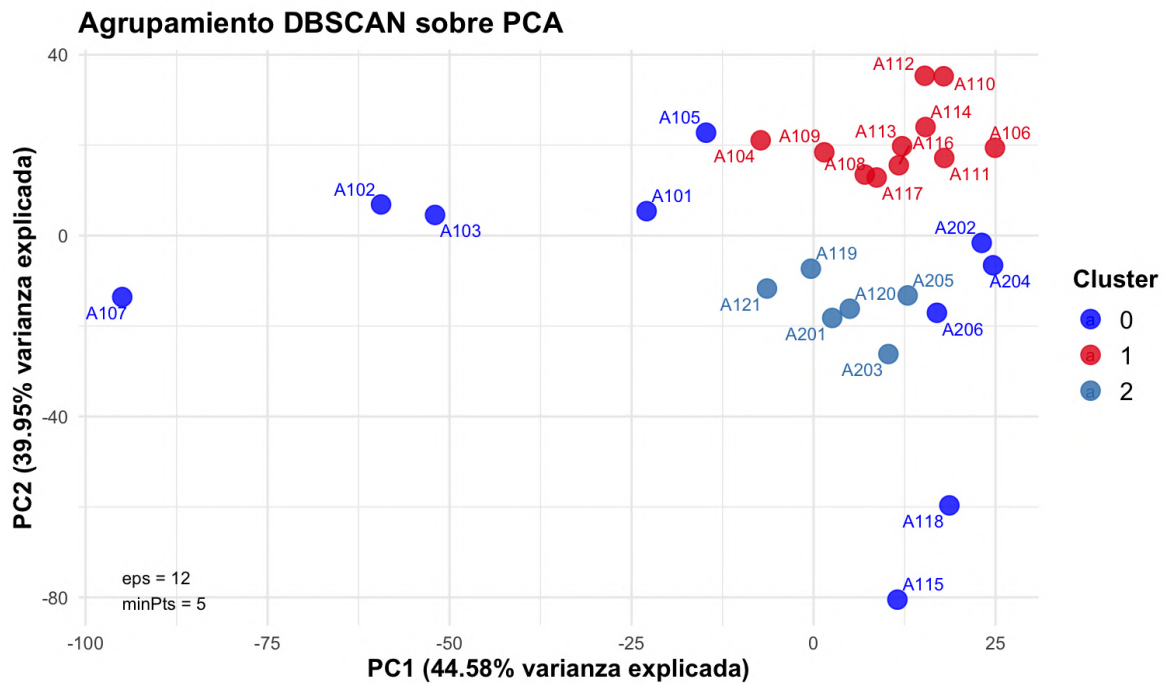


Figura 50. Biplot de PCA empleando el algoritmo de agrupamiento DBSCAN.

Al realizar el agrupamiento jerárquico de las muestras de suero, en el dendograma puede apreciarse que todas las muestras del grupo de cáncer de pulmón se encuentran hacia el lado derecho y que comparten un grupo que está separado de la mayoría del resto de muestras pertenecientes al grupo control, además, la distancia entre las muestras de este grupo es aparentemente menor que la distancia entre las muestras del grupo control (Figura 51).

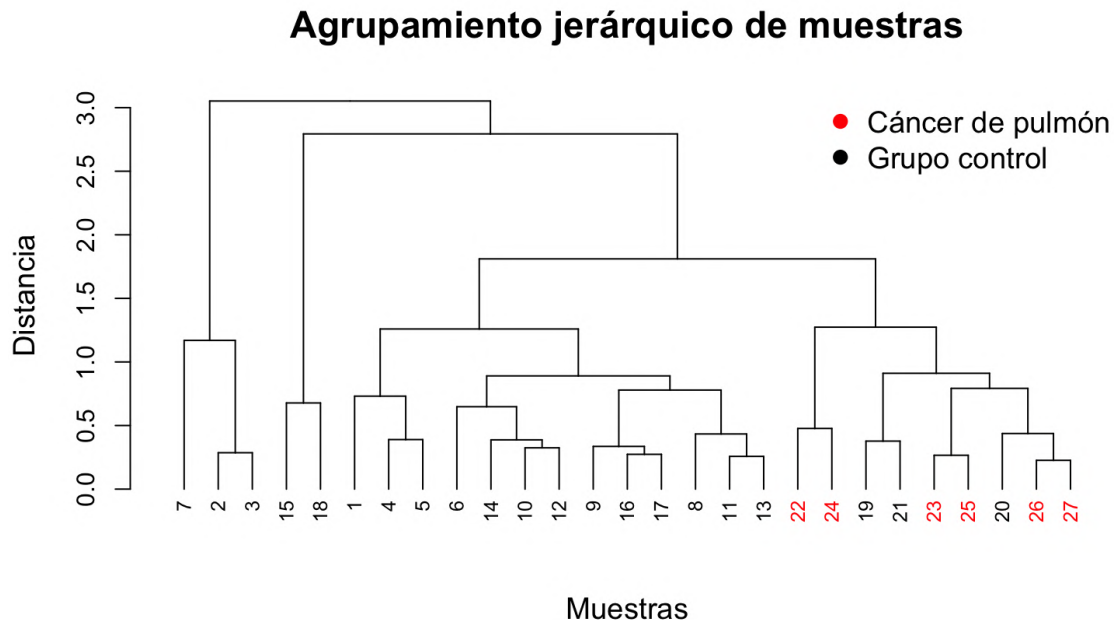


Figura 51. Dendrograma de las muestras de suero. El eje Y indica la distancia de las muestras basado en sus perfiles espectrales, las ramas agrupan las muestras en función de la similitud, ramas que se fusionan a mayor altura reflejan mayor diferencia.

La diferencia entre el agrupamiento jerárquico de ambos tipos de muestras puede radicar en las características que capturan, los cambios en los metabolitos pueden ocurrir de manera casi inmediata como respuesta a alteraciones en el metabolismo humano (Sala et al., 2023), que puede ser de diversos orígenes, relacionado directamente con enfermedades o con la forma en la que nuestros sistemas reaccionan para regular la homeostasis, asimilar medicamentos o responder ante agentes infecciosos, a diferencia de los cambios que pueda presentar el ADN, que si bien la metilación del ADN puede ser una manifestación de la respuesta celular a alteraciones en el entorno celular (como el microambiente tumoral) y a mutaciones específicas, la captura de estos cambios para su análisis sigue siendo un reto que se refleja en todas las propuestas de nuevas metodologías y herramientas para su análisis (P. Li et al., 2022). Es posible que el suero haya capturado de mejor manera

los cambios fisiológicos entre personas con y sin cáncer de pulmón, y por ello en el dendograma correspondiente permite observar una ligera separación entre ambos grupos.

Con el fin de explorar a mayor profundidad la distribución de las muestras de suero en el plano del PCA, se realizaron otros análisis, esta vez buscando la correspondencia de la distribución con alguna característica demográfica de las que fueron evaluadas (edad, género, IMC y nivel de actividad física o IPAQ). En la Figura 52 puede observarse que los grupos detectados por el algoritmo (Figura 50) tampoco se explican por características demográficas, las únicas muestras que comparten el mismo grupo y la misma característica (pertenecen al nivel alto de actividad física) son las muestra A115 y A118, del grupo control.

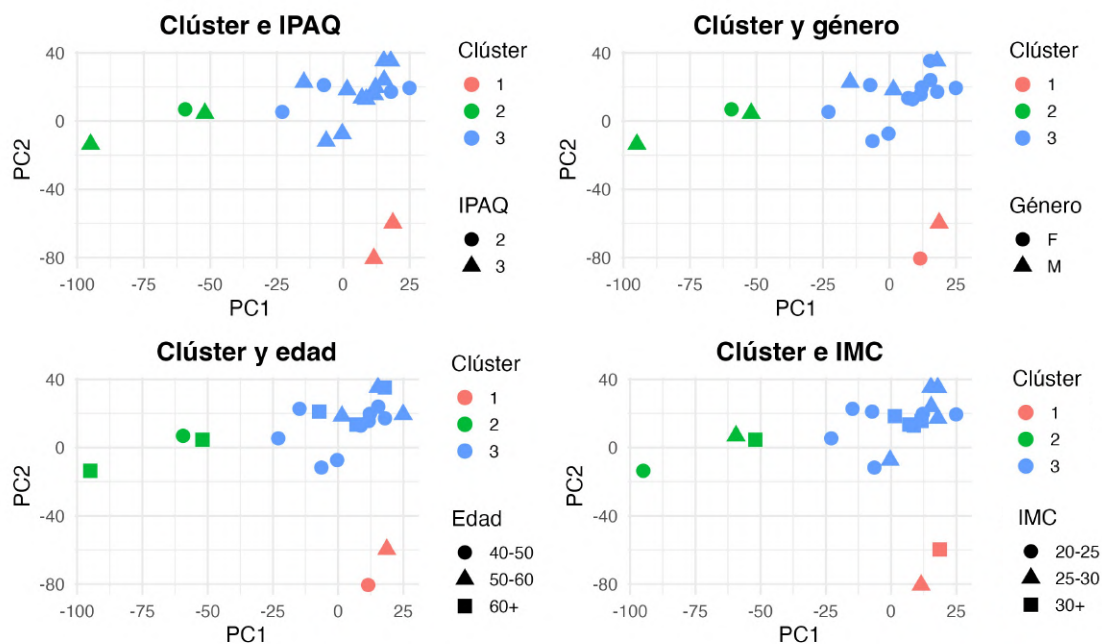


Figura 52. PCA con muestras indicadas por características demográficas. Clúster = grupo asignado por el algoritmo.

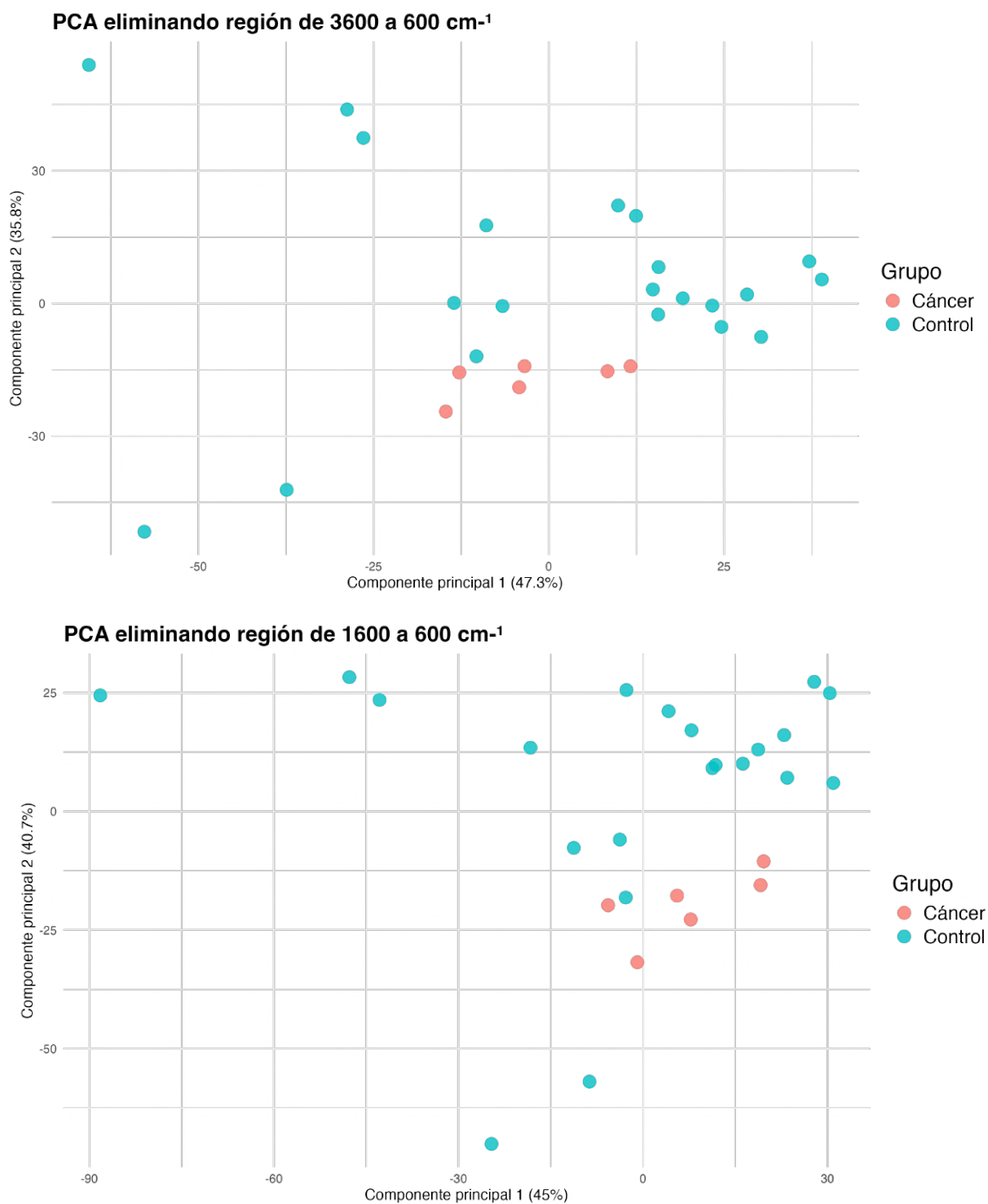


Figura 53. PCA de muestras de suero eliminando regiones espectrales antes del análisis.

Además, para demostrar la influencia del análisis espectral en esta distribución de las muestras, se realizaron dos PCA, esta vez eliminando regiones de interés que fueran determinantes para el agrupamiento. En la Figura 53 puede notarse que la distribución cambia completamente cuando se eliminan las regiones de 3600 a 600  $\text{cm}^{-1}$  y de 1600 a 600  $\text{cm}^{-1}$ , confirmando la importancia de estas regiones en la clasificación de las muestras y su capacidad para recolectar la mayor variabilidad de la base de datos espectrales.

Por lo tanto, hasta ahora la distribución de las muestras en el plano del PCA, puede explicarse únicamente por variaciones espectrales. Las características que el equipo de FTIR haya podido capturar de cada muestra puede influir directamente en los agrupamientos que se observen. Estas características pueden ser enfermedades concomitantes, alteraciones metabólicas, enfermedades desconocidas y por lo tanto no reportadas por los participantes como diabetes. Por la falta de esta información en los expedientes, no pueden realizarse más análisis buscando las posibles correlaciones.

#### 6.6.6 Análisis por regiones espectrales

Con los resultados obtenidos en el análisis multivariado se observó que las diferencias entre los grupos eran muy sutiles cuando se consideraban todas las señales obtenidas de la región completa del espectro infrarrojo medio. Con el fin de obtener una visión más clara de las bandas espectrales que se obtuvieron en el análisis y con esto, poder caracterizar las muestras de suero, se dividió el espectro en regiones de interés previamente reportadas (Yang et al., 2021a) y se calculó el área de banda espectral extrayendo los valores de absorbancia de las bandas elegidas y sumándolas. Una vez obtenidos estos valores se calcularon los promedios y desviación estándar por grupo para poder compararlos mediante una *t* de Student.

La región denominada Amida A (3600 – 3200  $\text{cm}^{-1}$ ) recibe contribución de la banda 3295 – 3293  $\text{cm}^{-1}$  reportada en muestras de suero de cáncer de pulmón y controles



sin cáncer (Wang et al., 2014). La región de los lípidos ( $3100 - 2700 \text{ cm}^{-1}$ ) abarca la banda del estiramiento asimétrico de  $\text{CH}_3$  a  $2959 - 2872 \text{ cm}^{-1}$  y  $\text{CH}_2$  a  $2930 \text{ cm}^{-1}$ . La región de Amida II ( $1800 - 1500 \text{ cm}^{-1}$ ) incluye la banda en  $1650 \text{ cm}^{-1}$  asociadas a la vibración de estiramiento del grupo  $\text{C}=\text{O}$  en los enlaces peptídicos de las proteínas, la banda del grupo funcional  $\text{N-H}$  a  $1542 \text{ cm}^{-1}$ , el estiramiento  $\text{C-N}$  a  $1533 \text{ cm}^{-1}$  y la banda a  $1546 \text{ cm}^{-1}$  que se asocia con la combinación de la flexión de  $\text{N-H}$  y este último estiramiento. La región Carbohidratos ( $1200 - 600 \text{ cm}^{-1}$ ) se denominó así principalmente por las bandas del estiramiento  $\text{C-O}$  y la flexión  $\text{C-O}$  a  $1044 - 1038 \text{ cm}^{-1}$  (Bangaoil et al., 2020). Además, se reportan bandas a  $1160 - 1120 \text{ cm}^{-1}$  correspondientes con estiramientos del enlace  $\text{C-O-C}$  y  $\text{C-O}$  en carbohidratos y lípidos y entre  $970$  y  $900 \text{ cm}^{-1}$  también correspondientes a carbohidratos como la glucosamina (Tomas et al., 2022). Se consideró que la región Metilación ( $1470 - 1250 \text{ cm}^{-1}$ ) podría incluir bandas indicadoras de metilación no solo en proteínas y lípidos, sino también en ácidos nucleicos, como las bandas entre  $1465$  y  $1450$  que corresponden a vibraciones de flexiones simétricas  $\text{CH}_3$  en proteínas y lípidos. Bandas entre  $1462$  y  $1464 \text{ cm}^{-1}$  se podrían asociar a vibraciones del  $\text{C}_4$  y  $\text{C}_5$  de la citosina, así como la vibración  $\text{C}_4\text{-NH}_2$  y  $\text{C}_5=\text{C}_6$  en la citosina ( $1363 \text{ cm}^{-1}$ ) y metilación de la citosina ( $1238 \text{ cm}^{-1}$ ) (Kelly et al., 2011). En la Figura 55 se puede observar que las regiones que presentan las mayores diferencias entre los grupos son la Amida A ( $p= 0.003$ ) y la Amida II ( $p= 0.02$ ), teniendo resultados contrastantes, ya que el contenido relacionado con la Amida A, siguiendo la ley de Beer-Lambert, se encuentra incrementado en el grupo con cáncer de pulmón, mientras que el contenido relacionado con la región de Amida II se encuentra disminuido. La región de las amidas se ha reportado previamente con valores de absorbancia más elevados en el grupo con cáncer, haciendo referencia a un incremento en la concentración de proteínas en los pacientes con cáncer, lo que en este caso podría interpretarse con el incremento que muestra la región de la denominada Amida A, cuyo aumento puede observarse claramente en la Figura 54. El traslape de las bandas en la regiones designadas que podrían no coincidir exactamente con las

regiones delimitadas por los otros autores podría explicar estos resultados contrastantes (Wang et al., 2014; Yang et al., 2021b).

Por otro lado, la región designada para los carbohidratos no presentó ninguna diferencia estadísticamente significativa entre los grupos de estudios, en congruencia con la mayoría de los artículos que tampoco reportan dicha variabilidad. Este resultado podría estar influido por numerosos factores, primero, es posible que la técnica sea menos sensible para esa región espectral, ya que las bandas predominantes o que presentan las señales más intensas son las proteínas, y estas pueden enmascarar a los carbohidratos (Araújo et al., 2022), considerando también que el suero humano presenta un elevado porcentaje de proteínas y una menor proporción de carbohidratos. Otro factor propiamente atribuible a la naturaleza de los grupos de estudio es que durante el cáncer las principales vías afectadas son las del metabolismo de las proteínas, lípidos y ácidos nucleicos. Sin embargo, estudios con resonancia magnética nuclear han encontrado alteraciones en el metabolismo de la glucosa en muestras de suero de pacientes con cáncer de pulmón (Singh et al., 2022), lo que podría sugerir una inclinación hacia la explicación de la baja sensibilidad de la espectroscopía infrarroja para detectar cambios sutiles en las concentraciones de carbohidratos en muestras complejas como el suero humano. Respecto a la región de los lípidos, los resultados también son congruentes con la mayoría de los artículos, ya que algunos si bien han reportado diferencias entre grupo de cáncer y grupo control, han sido sobre el límite del valor de significancia estadística (Yang et al., 2021b).

Finalmente, la región Metilación confirmó los resultados obtenidos por la ELISA de metilación de ADN (un menor porcentaje de metilación global en el grupo con cáncer de pulmón), con las bandas seleccionadas para conformar el área de banda. Además, este resultado de absorbancia de metilación disminuida en el grupo control fue estadísticamente significativo ( $p= 0.042$ ). En un estudio de tejido de cáncer de páncreas donde se empleó el análisis de metilación de ADN mediante espectroscopía de Raman para complementar la caracterización del tejido, se utilizó

un análisis multivariado y *machine learning* para determinar bandas espectrales que pudieran servir de indicadores de la metilación y establecer relaciones de absorbancia para determinar el porcentaje. Entre las regiones utilizadas se empleó la de 1420 a 1360  $\text{cm}^{-1}$ , y los resultados obtenidos demostraron una menor absorbancia en esta región en las muestras de cáncer pancreático en comparación con el grupo control de tejido de tumor benigno (Szymoński et al., 2023).

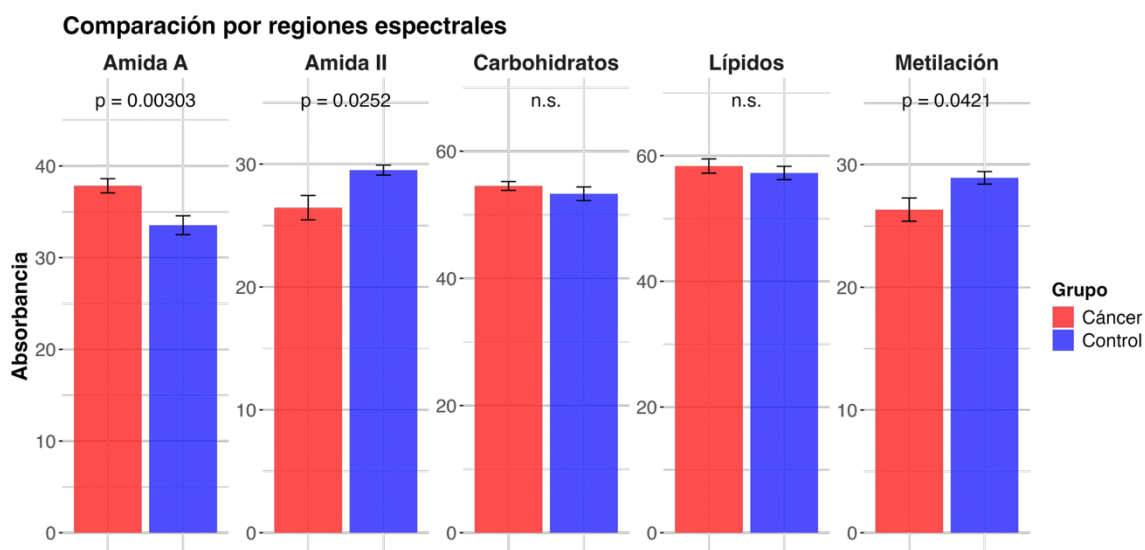


Figura 54. Área de banda espectral calculada por regiones: Amida A (3600 – 3200  $\text{cm}^{-1}$ ), Amida II (1800 – 1500  $\text{cm}^{-1}$ ), Carbohidratos (1200 – 600  $\text{cm}^{-1}$ ), Lípidos (3100 – 2700  $\text{cm}^{-1}$ ) y Metilación (1470 – 1250  $\text{cm}^{-1}$ ). Prueba t de Student,  $p < 0.05$ , n.s. = no significativo estadísticamente.

Con el fin de explorar la posible correlación entre las dos técnicas realizadas, los valores de absorbancia en la región denominada como Metilación, se compararon con los valores obtenidos con la ELISA, es decir, el porcentaje de metilación global de cada muestra. Aunque existe una tendencia entre el aumento y disminución de los valores obtenidos por cada técnica, esta relación es débil ( $r = 0.247$ ), además de presentar una gran dispersión entre los puntos, por lo que no puede asegurarse que ambas técnicas estén midiendo la metilación global de la misma manera, ni que FTIR pueda utilizarse como alternativa a la ELISA (Figura 55). Esto puede

deberse a que en primera instancia, lo capturado por FTIR es una mezcla compleja de moléculas de la muestra de suero, mientras que lo capturado por la ELISA son las citosinas metiladas de las muestras de ADN aislado, siendo sumamente difícil comparar la complejidad de ambos tipos de muestra. Además, debe considerarse que los número de onda seleccionados para formar parte de la región, pueden no estar directamente relacionados con la metilación a pesar de haber sido reportados como representativos de los ácidos nucleicos. Estos resultados se asemejan a lo reportado por Thumbovorn *et al.* en 2022, donde se compararon los valores de metilación global obtenidos por ELISA y pirosecuenciación ( $r = 0.442$ ), demostrando la variabilidad entre técnicas para detectar la metilación.

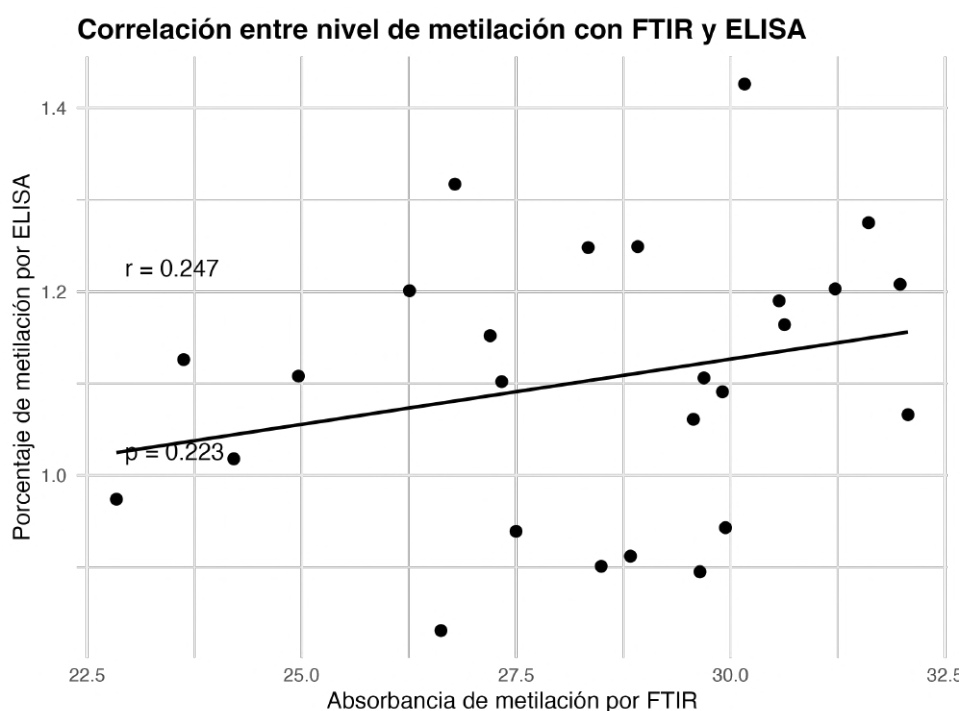


Figura 55. Correlación entre absorbancias por ELISA y FTIR en región Metilación de muestras de suero. Los puntos representan muestras individuales, la línea negra corresponde al ajuste por regresión lineal simple. Coeficiente de correlación de Pearson indicado con  $r$ ,  $p < 0.05$ .

## 7 Conclusiones

En esta investigación se confirma que el proceso de hipometilación observado en otros tipos de cáncer, como cáncer de mama y cervicouterino, también se presenta en el cáncer de pulmón y es detectable en muestras de sangre periférica. El análisis espectral de las muestras de ADN aislado de muestras biológicas permite obtener información sobre su estructura, fundamentando la propuesta de los marcadores espectrales. Este estudio, además, permite señalar la importancia de considerar la duración del tiempo del procesamiento de las muestras de ADN para obtener espectros óptimos que faciliten la identificación de bandas que proporcionen información sobre la estructura del ADN. Las muestras de suero se proponen como una mejor opción para el análisis espectral por la facilidad de manejo y porque a pesar de ser una matriz más compleja, permite obtener información sobre ácidos nucleicos. La correlación entre el grado de metilación mediante ambas técnicas aunque fue positiva y moderada, no es contundente, por lo que se concluye que no es posible proponer la detección de la metilación del ADN mediante espectroscopía FTIR en el ámbito clínico.

Para evaluar a profundidad su capacidad para la detección de la estructura del ADN se recomienda que en estudios posteriores se establezcan criterios más rigurosos para el manejo del ADN aislado, así como pruebas piloto con ADN en solución acuosa y la incorporación de controles y curvas de calibración para cuantificar. Los resultados plasmados en este trabajo deben interpretarse con las precauciones correspondientes: se ha utilizado una población limitada y homogénea de pacientes con cáncer de pulmón y no se han realizado análisis confirmatorios de la caracterización de las muestras con otro tipo de metodología espectral o de resonancia. Con estas consideraciones, es importante recalcar el cumplimiento de la hipótesis propuesta, tanto para la detección de la hipometilación como de la correlación de ambas técnicas y de la aportación al conocimiento de las múltiples aplicaciones de una técnica sencilla como la espectroscopía infrarroja por transformada de Fourier en el estudio de muestras biológicas.

## 8 Referencias

- Aguilera, O., Fernandez, A., & Fraga, M. (2010). Epigenetics and environment: A complex relationship. *Journal of applied physiology (Bethesda, Md. : 1985)*, 109, 243-251. <https://doi.org/10.1152/jappphysiol.00068.2010>
- Alcázar-Ramos, A. G. (2019). *Determinación del grado de exposición al humo del tabaco y su asociación con factores de predisposición para el desarrollo de cáncer de pulmón* [Tesis de maestría, Universidad Autónoma de Querétaro]. <https://ri-ng.uaq.mx/handle/123456789/1899>
- Alcázar-Ramos, A. G., Meraz-Ríos, M. A., Figueroa-Corona, M. D. P., García-Gutiérrez, D. G., Bertadillo-Jilote, A. D., Pérez-Ramírez, I. F., García-Mejía, K. L., Ugalde-Villanueva, B., & Lira-De León, K. I. (2019). P1.06 Determination of Exposure to Tobacco Smoke and the Percentage of Global DNA Methylation in Smokers and Patients with Lung Cancer. *Journal of Thoracic Oncology*, 14(11, Supplement 2), S1175-S1176. <https://doi.org/10.1016/j.jtho.2019.09.140>
- American Cancer Society. (2023). *Cancer Facts & Figures 2023*. American Cancer Society. <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2023/2023-cancer-facts-and-figures.pdf>
- Andriyani, F., & Puspitarani, Y. (2022). Performance Comparison of K-Means and DBScan Algorithms for Text Grouping Product Reviews. *Sinkron*, 7, 944-949. <https://doi.org/10.33395/sinkron.v7i3.11569>
- Araújo, R., Ramalheite, L., Ribeiro, E., & Calado, C. (2022). Plasma versus Serum Analysis by FTIR Spectroscopy to Capture the Human Physiological State. *BioTech*, 11(4), 56. <https://doi.org/10.3390/biotech11040056>
- Arroyo-Hernández, M., Alva-López, L. F., Rendón, A., Barroso-Villafuerte, F. R., Báez-Saldaña, R., Corona-Cruz, J. F., Farfán-Salazar, G., Guerrero-Ixtlahuac, J., Castillo-González, P., Salmon-Demongin, A., Zaldívar-Crosby, G., Bolaño-Guerra, L. M., Zavala-Cruz, G. G., Sánchez-Gutiérrez, A., Rendón-Ramírez, E. J., Magdaleno-Maldonado, G. E., Olivares-Torres, C. A., López-Saucedo, R. A., Lizardo-Rodríguez, A. E., ... Arrieta-Rodríguez, O. G. (2022). Guía de Práctica Clínica para el diagnóstico temprano y la referencia oportuna del cáncer de pulmón. *Salud Pública de México*, 64(5, sept-oct), Article 5, sept-oct. <https://doi.org/10.21149/13919>
- Aufdembrink, L. M., Hoog, T. G., Pawlak, M. R., Bachan, B. F., Heili, J. M., & Engelhart, A. E. (2019). Chapter Two—Methods for thermal denaturation studies of nucleic acids in complex with fluorogenic dyes. En A. E. Hargrove (Ed.), *Methods in Enzymology* (Vol. 623, pp. 23-43). Academic Press. <https://doi.org/10.1016/bs.mie.2019.05.029>
- Baker, M. J., Trevisan, J., Bassan, P., Bhargava, R., Butler, H. J., Dorling, K. M., Fielden, P. R., Fogarty, S. W., Fullwood, N. J., Heys, K. A., Hughes, C., Lasch, P., Martin-Hirsch, P. L., Obinaju, B., Sockalingum, G. D., Sulé-Suso, J., Strong, R. J., Walsh, M. J., Wood, B. R., ... Martin, F. L. (2014). Using Fourier transform IR

- spectroscopy to analyze biological materials. *Nature Protocols*, 9(8), 1771-1791. <https://doi.org/10.1038/nprot.2014.110>
- Balan, V., Mihai, C.-T., Cojocaru, F.-D., Uritu, C.-M., Dodi, G., Botezat, D., & Gardikiotis, I. (2019). Vibrational Spectroscopy Fingerprinting in Medicine: From Molecular to Clinical Practice. *Materials*, 12(18), Article 18. <https://doi.org/10.3390/ma12182884>
- Bangaoil, R., Santillan, A., Angeles, L. M., Abanilla, L., Jr, A. L., Ramos, M. C., Fellizar, A., Jr, L. G., & Albano, P. M. (2020). ATR-FTIR spectroscopy as adjunct method to the microscopic examination of hematoxylin and eosin-stained tissues in diagnosing lung cancer. *PLOS ONE*, 15(5), e0233626. <https://doi.org/10.1371/journal.pone.0233626>
- Banyay, M., Sarkar, M., & Gräslund, A. (2003). A library of IR bands of nucleic acids in solution. *Biophysical Chemistry*, 104(2), 477-488. [https://doi.org/10.1016/S0301-4622\(03\)00035-8](https://doi.org/10.1016/S0301-4622(03)00035-8)
- Beć, K. B., Grabska, J., & Huck, C. W. (2022). Chapter One—Physical principles of infrared spectroscopy. En D. Cozzolino (Ed.), *Comprehensive Analytical Chemistry* (Vol. 98, pp. 1-43). Elsevier. <https://doi.org/10.1016/bs.coac.2020.08.001>
- Beltran, D. M. A. (2022). *Evaluación del porcentaje de metilación del ADN en sangre periférica ocasionado por el consumo del tabaco*. <https://ring.uaq.mx/handle/123456789/3565>
- Blat, A., Wiercigroch, E., Smeda, M., Wislocka, A., Chlopicki, S., & Malek, K. (2019). Fourier transform infrared spectroscopic signature of blood plasma in the progression of breast cancer with simultaneous metastasis to lungs. *Journal of Biophotonics*, 12(10), e201900067. <https://doi.org/10.1002/jbio.201900067>
- Brody, D. J., Faust, E., & Tsai, J. (2021). Secondhand Smoke Exposure Among Nonsmoking Adults: United States, 2015-2018. *NCHS Data Brief*, 369, 1-8.
- Chen, Y. (2020). *DBSCAN Is Semi-Spectral Grouping\** (p. 264). <https://doi.org/10.1109/BigDIA51454.2020.00048>
- Departamento de control de calidad del Banco Nacional de ADN. (2024, junio). *Programa de control de calidad de muestras de ADN y ARN Banco Nacional de ADN Carlos III*. Universidad de Salamanca. [www.bancoadn.org](http://www.bancoadn.org)
- Desjardins, P., & Conklin, D. (2010). NanoDrop microvolume quantitation of nucleic acids. *Journal of Visualized Experiments: JoVE*, 45, 2565. <https://doi.org/10.3791/2565>
- Detterbeck, F. C. (2018). The eighth edition TNM stage classification for lung cancer: What does it mean on main street? *The Journal of Thoracic and Cardiovascular Surgery*, 155(1), 356-359. <https://doi.org/10.1016/j.jtcvs.2017.08.138>
- Division of Cancer Prevention and Control, Centers for Disease Control and Prevention. (2023, agosto 2). *¿Cuáles son los factores de riesgo del cáncer de pulmón? | Centros para el Control y la Prevención de Enfermedades*. [https://www.cdc.gov/cancer/lung/basic\\_info/risk\\_factors.htm](https://www.cdc.gov/cancer/lung/basic_info/risk_factors.htm)
- El-Azazy, M. (2018). Introductory Chapter: Infrared Spectroscopy - A Synopsis of the Fundamentals and Applications. En *Infrared Spectroscopy—Principles, Advances, and Applications*. IntechOpen.

- <https://doi.org/10.5772/intechopen.82210>
- EpigenTek. (2022). *MethylFlash™ Methylated DNA Quantification Kit (Colorimetric)*. <https://www.epigentek.com/catalog/methylflash-methylated-dna-quantification-kit-colorimetric-p-2094.html>
- Esteller, M. (Ed.). (2019). *Epigenetics in Biology and Medicine*. CRC Press. <https://doi.org/10.1201/9781420007077>
- Farré, P., Jones, M. J., Meaney, M. J., Emberly, E., Turecki, G., & Kobor, M. S. (2015). Concordant and discordant DNA methylation signatures of aging in human blood and brain. *Epigenetics & Chromatin*, 8(1), 19. <https://doi.org/10.1186/s13072-015-0011-y>
- Fehr, S. M., & Krossing, I. (2020). Spectroscopic Signatures of Pressurized Carbon Dioxide in Diffuse Reflectance Infrared Spectroscopy of Heterogeneous Catalysts. *ChemCatChem*, 12(9), 2622-2629. <https://doi.org/10.1002/cctc.201902038>
- Fernandez, A. F., Valledor, L., Vallejo, F., Cañal, M. J., & Fraga, M. F. (2018). Quantification of Global DNA Methylation Levels by Mass Spectrometry. *Methods in Molecular Biology (Clifton, N.J.)*, 1708, 49-58. [https://doi.org/10.1007/978-1-4939-7481-8\\_3](https://doi.org/10.1007/978-1-4939-7481-8_3)
- Fior, R., & Zilhão, R. (Eds.). (2019). *Molecular and cell biology of cancer. When Cells Break the Rules and Hijack Their Own Planet*. Springer Cham. <https://doi.org/10.1007/978-3-030-11812-9>
- Ghosal, R., Lewis, K. E. L., Kloer, P., Mehta, R., Parry, D., Llewellyn-Jones, C., Mur, L., Blaser, J., & Lewis, P. D. (2008). Using Fourier transform infrared (FTIR) spectroscopy to evaluate metabolic markers in sputum in patients with and without lung cancer. *Lung Cancer*, 60, S11. [https://doi.org/10.1016/S0169-5002\(08\)70033-3](https://doi.org/10.1016/S0169-5002(08)70033-3)
- GLOBOCAN. (2025). *Cancer Today*. <https://gco.iarc.who.int/today/>
- González de Buitrago, J. M. (2010). *Técnicas y métodos de laboratorio clínico*. Elsevier España.
- Guo, L., Yu, Y., Yang, F., Gao, W., Wang, Y., Xiao, Y., Du, J., Tian, J., & Yang, H. (2023). Accuracy of baseline low-dose computed tomography lung cancer screening: A systematic review and meta-analysis. *Chinese Medical Journal*, 136(9), 1047-1056. <https://doi.org/10.1097/CM9.0000000000002353>
- Gurbanov, R., Tunçer, S., Mingu, S., Severcan, F., & Gozen, A. G. (2019). Methylation, sugar puckering and Z-form status of DNA from a heavy metal-acclimated freshwater *Gordonia* sp. *Journal of Photochemistry and Photobiology B: Biology*, 198, 111580. <https://doi.org/10.1016/j.jphotobiol.2019.111580>
- Han, Y., Han, L., Yao, Y., Li, Y., & Liu, X. (2018). Key factors in FTIR spectroscopic analysis of DNA: The sampling technique, pretreatment temperature and sample concentration. *Analytical Methods*, 10(21), 2436-2443. <https://doi.org/10.1039/C8AY00386F>
- Hana, J., & Faizah, N. (2023). Diagnostic Approach of Lung Cancer: A Literature Review. *Jurnal Respirasi*, 9, 141-149. <https://doi.org/10.20473/jr.v9-i.2.2023.141-149>
- Heredia, D., Arrieta, O., Valezka, F., Morales, B., Magos, F., Ruiz, F., Corona-Cruz, J.,



- Francisco, J., Cruz, C., San, A., No, F., Heredia, C., Dj, V., Fv, M., Magos, M., Lozano, F., Fj, R., & Cáncer, J. (2023). *Oncoguía de cáncer de pulmón de células no pequeñas en estadios tempranos, localmente avanzados y metastásicos 2023*.
- Herráez, Á. (2012). *Biología molecular e ingeniería genética: Conceptos, técnicas y aplicaciones en ciencias de la salud* (2.<sup>a</sup> ed.). Elsevier.
- Hong, Y., & Kim, W. J. (2021). DNA Methylation Markers in Lung Cancer. *Current Genomics*, 22(2), 79-87. <https://doi.org/10.2174/1389202921999201013164110>
- Huang, M., Giese, T. J., Lee, T.-S., & York, D. M. (2014). Improvement of DNA and RNA Sugar Pucker Profiles from Semiempirical Quantum Methods. *Journal of Chemical Theory and Computation*, 10(4), 1538-1545. <https://doi.org/10.1021/ct401013s>
- International Agency for Research on Cancer. (2016). *IARC monographs on the evaluation of carcinogenic risks to humans: Vol 109, outdoor air pollution*.
- Iwasa, J., & Marshall, W. (2019). *Karp Biología celular y molecular. Conceptos y experimentos* (8.<sup>a</sup> ed.). McGraw Hill Interamericana.
- Kelly, J. G., Najand, G. M., & Martin, F. L. (2011). Characterisation of DNA methylation status using spectroscopy (mid-IR versus Raman) with multivariate analysis. *Journal of Biophotonics*, 4(5), 345-354. <https://doi.org/10.1002/jbio.201000085>
- Kim, S. H., Lee, H., Kim, B.-G., Kim, S.-H., Sohn, J. W., Yoon, H. J., & Park, D. W. (2024). Impact of family history of cancer on development of lung cancer among the Korean population: A prospective cohort study using KoGES data. *Journal of Thoracic Disease*, 16(2). <https://doi.org/10.21037/jtd-23-1077>
- Kokalj, M., Rihtarič, M., & Kreft, S. (2011). Commonly applied smoothing of IR spectra showed unappropriate for the identification of plant leaf samples. *Chemometrics and Intelligent Laboratory Systems*, 108(2), 154-161. <https://doi.org/10.1016/j.chemolab.2011.07.001>
- Krall, J. B., Nichols, P. J., Henen, M. A., Vicens, Q., & Vögeli, B. (2023). Structure and Formation of Z-DNA and Z-RNA. *Molecules*, 28(2), Article 2. <https://doi.org/10.3390/molecules28020843>
- Kuchiba, A., Iwasaki, M., Ono, H., Kasuga, Y., Yokoyama, S., Onuma, H., Nishimura, H., Kusama, R., Tsugane, S., & Yoshida, T. (2014). Global methylation levels in peripheral blood leukocyte DNA by LUMA and breast cancer: A case–control study in Japanese women. *British Journal of Cancer*, 110(11), 2765-2771. <https://doi.org/10.1038/bjc.2014.223>
- Kurdyukov, S., & Bullock, M. (2016). DNA Methylation Analysis: Choosing the Right Method. *Biology*, 5(1), 3. <https://doi.org/10.3390/biology5010003>
- Lam, L. L., Emberly, E., Fraser, H. B., Neumann, S. M., Chen, E., Miller, G. E., & Kobor, M. S. (2012a). Factors underlying variable DNA methylation in a human community cohort. *Proceedings of the National Academy of Sciences*, 109(supplement\_2), 17253-17260. <https://doi.org/10.1073/pnas.1121249109>
- Lam, L. L., Emberly, E., Fraser, H. B., Neumann, S. M., Chen, E., Miller, G. E., & Kobor, M. S. (2012b). Factors underlying variable DNA methylation in a human community cohort. *Proceedings of the National Academy of Sciences*, 109(supplement\_2), 17253-17260. <https://doi.org/10.1073/pnas.1121249109>

- Lee, L. C., Liong, C.-Y., & Jemain, A. A. (2017). A contemporary review on Data Preprocessing (DP) practice strategy in ATR-FTIR spectrum. *Chemometrics and Intelligent Laboratory Systems*, 163, 64-75. <https://doi.org/10.1016/j.chemolab.2017.02.008>
- Lewis, P. D., Lewis, K. E., Ghosal, R., Bayliss, S., Lloyd, A. J., Wills, J., Godfrey, R., Kloer, P., & Mur, L. A. (2010). Evaluation of FTIR Spectroscopy as a diagnostic tool for lung cancer using sputum. *BMC Cancer*, 10(1), 640. <https://doi.org/10.1186/1471-2407-10-640>
- Li, L., Fu, K., Zhou, W., & Snyder, M. (2019). Applying circulating tumor DNA methylation in the diagnosis of lung cancer. *Precision Clinical Medicine*, 2(1), 45-56. <https://doi.org/10.1093/pcmedi/pbz003>
- Li, L., Zheng, H., Huang, Y., Huang, C., Zhang, S., Tian, J., Li, P., Sood, A. K., Zhang, W., & Chen, K. (2017). DNA methylation signatures and coagulation factors in the peripheral blood leucocytes of epithelial ovarian cancer. *Carcinogenesis*, 38(8), 797-805. <https://doi.org/10.1093/carcin/bgx057>
- Li, P., Liu, S., Du, L., Mohseni, G., Zhang, Y., & Wang, C. (2022). Liquid biopsies based on DNA methylation as biomarkers for the detection and prognosis of lung cancer. *Clinical Epigenetics*, 14(1), 118. <https://doi.org/10.1186/s13148-022-01337-0>
- Li, S., Silvestri, V., Leslie, G., Rebbeck, T. R., Neuhausen, S. L., Hopper, J. L., Nielsen, H. R., Lee, A., Yang, X., McGuffog, L., Parsons, M. T., Andrulis, I. L., Arnold, N., Belotti, M., Borg, A., Buecher, B., Buys, S. S., Caputo, S. M., Chung, W. K., ... Antoniou, A. C. (2022). Cancer Risks Associated With BRCA1 and BRCA2 Pathogenic Variants. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 40(14), 1529-1541. <https://doi.org/10.1200/JCO.21.02112>
- Li, Y., Fan, Z., Meng, Y., Liu, S., & Zhan, H. (2023). Blood-based DNA methylation signatures in cancer: A systematic review. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1869(1), 166583. <https://doi.org/10.1016/j.bbadis.2022.166583>
- Liu, J., He, W., Wang, Y., & Zhang, B. (2024). Evaluation of dimensionality reduction and unsupervised grouping methods in breast datasets. *Applied and Computational Engineering*, 31, 218-228. <https://doi.org/10.54254/2755-2721/31/20230153>
- Lu, Z., & Yuan, K.-H. (2010). *Welch's t test* (pp. 1620-1623). <https://doi.org/10.13140/RG.2.1.3057.9607>
- Martisova, A., Holcakova, J., Izadi, N., Sebuyoya, R., Hrstka, R., & Bartosik, M. (2021). DNA Methylation in Solid Tumors: Functions and Methods of Detection. *International Journal of Molecular Sciences*, 22(8), Article 8. <https://doi.org/10.3390/ijms22084247>
- Medina, C., Monge, A., Denova-Gutiérrez, E., López-Ridaura, R., Barquera, S., Romieu, I., & Lajous, M. (2022). Validity and reliability of the International Physical Activity Questionnaire (IPAQ) long-form in a subsample of female Mexican teachers. *Salud Pública de México*, 64(1), Article 1. <https://doi.org/10.21149/12889>
- Miller, K. D., Nogueira, L., Devasia, T., Mariotto, A. B., Yabroff, K. R., Jemal, A., Kramer,

- J., & Siegel, R. L. (2022). Cancer treatment and survivorship statistics, 2022. *CA: A Cancer Journal for Clinicians*, 72(5), 409-436. <https://doi.org/10.3322/caac.21731>
- Mokari, A., Guo, S., & Bocklitz, T. (2023). Exploring the Steps of Infrared (IR) Spectral Analysis: Pre-Processing, (Classical) Data Modelling, and Deep Learning. *Molecules*, 28(19), 6886. <https://doi.org/10.3390/molecules28196886>
- Mondragón-Cortez, P. (2017). *Espectroscopia de infrarrojo para todos*. Centro de Investigación y Asistencia en Tecnología y Diseño del Estado de Jalisco, A. C. [https://ciatej.mx/files/el\\_ciatej/comunicacion/divulgacion/espectroscopia.pdf](https://ciatej.mx/files/el_ciatej/comunicacion/divulgacion/espectroscopia.pdf)
- Mondragón-Cortez, P. (2020). *Principios y aplicaciones de la espectroscopia de infrarrojo en el análisis de alimentos y bebidas*. Centro de Investigación y Asistencia en Tecnología y Diseño del Estado de Jalisco A. C. [www.ciatej.mx](http://www.ciatej.mx)
- Muningsih, E., Kesuma, C., Sunanto, Suripah, & Widayanto, A. (2023). Combination of K-Means method with Davies Bouldin index and decision tree method with parameter optimization for best performance. En *AIP Conference Proceedings*. <https://doi.org/10.1063/5.0129119>
- National Cancer Institute. (2023, septiembre 6). *Lung Cancer Screening (PDQ®)—NCI* (nciglobal,ncienterprise) [pdqCancerInfoSummary]. <https://www.cancer.gov/types/lung/hp/lung-screening-pdq>
- Organización Mundial de la Salud. (2023). *Cáncer*. Organización Mundial de la Salud. <https://www.who.int/es/news-room/fact-sheets/detail/cancer>
- Ortega, S., Banda, E., Nohra Violeta, G.-R., Paramo-Garcia, U., & Mendoza-Martínez, A. (2024). Aplicación de filtros Savitzky-Golay y transformada rápida de Fourier en el procesamiento de espectros derivados obtenidos a partir de soluciones de asfaltenos. *CienciaUAT*, 18. <https://doi.org/10.29059/cienciauat.v18i2.1812>
- Ortiz Zamora, S. I., Nieto Aguilar, R., Characu Rueda Rocha, A., Manuel Jacobo, C., Sánchez Moreno, J. A., & Aguilera Méndez, A. (2023). *La epigenética y el cáncer: Terapia epigenética, una alternativa terapéutica*. 89, 21-32. <https://doi.org/10.35830/cn.vi89.716>
- Pei, Y., Lou, X., Li, K., Xu, X., Guo, Y., Xu, D., Yang, Z., Xu, D., Cui, W., & Zhang, D. (2020). Peripheral Blood Leukocyte N6-methyladenosine is a Noninvasive Biomarker for Non-small-cell Lung Carcinoma. *OncoTargets and therapy*, 13, 11913-11921. <https://doi.org/10.2147/OTT.S267344>
- Pineda-Lancheros, L. E., Rojo Tolosa, S., Gálvez Navas, J. M., Martínez Martínez, F., Sánchez Martín, A., Jiménez Morales, A., & Pérez Ramírez, C. (2022). Effect of Single Nucleotide Polymorphisms in the Vitamin D Metabolic Pathway on Susceptibility to Non-Small-Cell Lung Cancer. *Nutrients*, 14(21), 4668. <https://doi.org/10.3390/nu14214668>
- Rojas Moreno, A. P., Bruges, R. E., Cañas Arboleda, A., & Jaramillo, L. F. (2016). *Regulación epigenética en cáncer de pulmón: Implicaciones para el clínico*. 57(3), 332-347. <http://dx.doi.org/10.11144/Javeriana.umed57-3.recip>
- Sala, A., Cameron, J. M., Brennan, P. M., Crosbie, E. J., Curran, T., Gray, E., Martin-Hirsch, P., Palmer, D. S., Rehman, I. U., Rattray, N. J. W., & Baker, M. J. (2023). Global serum profiling: An opportunity for earlier cancer detection. *Journal of Experimental & Clinical Cancer Research: CR*, 42, 207.

- <https://doi.org/10.1186/s13046-023-02786-y>
- Salazar Montes, A. M., Sandoval Rodríguez, A. S., & Armendáriz Borunda, J. S. (2016). *Biología molecular. Fundamentos y aplicaciones en las ciencias de la salud* (2.<sup>a</sup> ed.). McGraw Hill Interamericana.
- Santos, J., & Embrechts, M. (2009). On the Use of the Adjusted Rand Index as a Metric for Evaluating Supervised Classification. En *Artificial Neural Networks-ICANN* (Vol. 2009, p. 184). [https://doi.org/10.1007/978-3-642-04277-5\\_18](https://doi.org/10.1007/978-3-642-04277-5_18)
- Sarić, A., Rajić, J., Tolić, A., Dučić, T., & Vidaković, M. (2023). Synchrotron-based FTIR microspectroscopy reveals DNA methylation profile in DNA-HALO structure. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 302, 123090. <https://doi.org/10.1016/j.saa.2023.123090>
- Shen, X., Ye, S., Xu, L., Hu, R., Jin, L., Xu, H., Liu, J., & Liu, W. (2018). Study on baseline correction methods for the Fourier transform infrared spectra with different signal-to-noise ratios. *Applied Optics*, 57(20), 5794-5799. <https://doi.org/10.1364/AO.57.005794>
- Shirahata, A., & Hibi, K. (2014). Serum vimentin methylation as a potential marker for colorectal cancer. *Anticancer Research*, 34(8), 4121-4125.
- Singh, A., Prakash, V., Gupta, N., Kumar, A., Kant, R., & Kumar, D. (2022). Serum Metabolic Disturbances in Lung Cancer Investigated through an Elaborative NMR-Based Serum Metabolomics Approach. *ACS Omega*, 7(6), 5510-5520. <https://doi.org/10.1021/acsomega.1c06941>
- Souza, V. G. P., Forder, A., Brockley, L. J., Pewarchuk, M. E., Telkar, N., de Araújo, R. P., Trejo, J., Benard, K., Seneda, A. L., Minutentag, I. W., Erkan, M., Stewart, G. L., Hasimoto, E. N., Garnis, C., Lam, W. L., Martinez, V. D., & Reis, P. P. (2023). Liquid Biopsy in Lung Cancer: Biomarkers for the Management of Recurrence and Metastasis. *International Journal of Molecular Sciences*, 24(10). <https://doi.org/10.3390/ijms24108894>
- Staniszewska-Slezak, E., Wiercigroch, E., Fedorowicz, A., Buczek, E., Mateuszuk, L., Baranska, M., Chlopicki, S., & Malek, K. (2018). A possible Fourier transform infrared-based plasma fingerprint of angiotensin-converting enzyme inhibitor-induced reversal of endothelial dysfunction in diabetic mice. *Journal of Biophotonics*, 11(2). <https://doi.org/10.1002/jbio.201700044>
- Su, K.-Y., & Lee, W.-L. (2020). Fourier Transform Infrared Spectroscopy as a Cancer Screening and Diagnostic Tool: A Review and Prospects. *Cancers*, 12(1), Article 1. <https://doi.org/10.3390/cancers12010115>
- Szigeti, K. A., Barták, B. K., Nagy, Z. B., Zsigrai, S., Papp, M., Márkus, E., Igaz, P., Takács, I., Molnár, B., & Kalmár, A. (2022). Methodological and Biological Factors Influencing Global DNA Methylation Results Measured by LINE-1 Pyrosequencing Assay in Colorectal Tissue and Liquid Biopsy Samples. *International Journal of Molecular Sciences*, 23(19), Article 19. <https://doi.org/10.3390/ijms231911608>
- Szymoński, K., Lipiec, E., Sofińska, K., Skirlińska-Nosek, K., Czaja, M., Seweryn, S., Wilkosz, N., Birarda, G., Piccirilli, F., Vaccari, L., Chmura, Ł., Szpor, J., Adamek, D., & Szymoński, M. (2023). Variabilities in global DNA methylation and  $\beta$ -sheet richness establish spectroscopic landscapes among subtypes of pancreatic

- cancer. *European Journal of Nuclear Medicine and Molecular Imaging*, 50(6), 1792. <https://doi.org/10.1007/s00259-023-06121-7>
- Temiz, N. A., Donohue, D. E., Bacolla, A., Luke, B. T., & Collins, J. R. (2012). The Role of Methylation in the Intrinsic Dynamics of B- and Z-DNA. *PLOS ONE*, 7(4), e35558. <https://doi.org/10.1371/journal.pone.0035558>
- Thai, A. A., Solomon, B. J., Sequist, L. V., Gainor, J. F., & Heist, R. S. (2021). Lung cancer. *Lancet (London, England)*, 398(10299), 535-554. [https://doi.org/10.1016/S0140-6736\(21\)00312-3](https://doi.org/10.1016/S0140-6736(21)00312-3)
- Thumbovorn, R., Bhattarakosol, P., & Chaiwongkot, A. (2022). Detection of Global DNA Methylation in Cervical Intraepithelial Neoplasia and Cancerous Lesions by Pyrosequencing and Enzyme-Linked Immunosorbent Assays. *Asian Pacific Journal of Cancer Prevention: APJCP*, 23(1), 143-149. <https://doi.org/10.31557/APJCP.2022.23.1.143>
- Tollefsbol, T. (2023). *Handbook of Epigenetics The New Molecular and Medical Genetics* (3.<sup>a</sup> ed.). Academic Press.
- Tomas, R. C., Sayat, A. J., Atienza, A. N., Danganan, J. L., Ramos, M. R., Fellizar, A., Notarte, K. I., Angeles, L. M., Bangaol, R., Santillan, A., & Albano, P. M. (2022). Detection of breast cancer by ATR-FTIR spectroscopy using artificial neural networks. *PLOS ONE*, 17(1), e0262489. <https://doi.org/10.1371/journal.pone.0262489>
- Vitorino, R., Barros, A. S., Guedes, S., Caixeta, D. C., & Sabino-Silva, R. (2023). Diagnostic and monitoring applications using near infrared (NIR) spectroscopy in cancer and other diseases. *Photodiagnosis and Photodynamic Therapy*, 42, 103633. <https://doi.org/10.1016/j.pdpdt.2023.103633>
- Wang, X., Shen, X., Sheng, D., Chen, X., & Liu, X. (2014). FTIR spectroscopic comparison of serum from lung cancer patients and healthy persons. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 122, 193-197. <https://doi.org/10.1016/j.saa.2013.11.049>
- Weber, G. F. (2016). Metabolism in cancer metastasis. *International Journal of Cancer*, 138(9), 2061-2066. <https://doi.org/10.1002/ijc.29839>
- Weiss, D. J. (2005). Multifactor Designs. En D. J. Weiss (Ed.), *Analysis of Variance and Functional Measurement: A Practical Guide* (p. 0). Oxford University Press. <https://doi.org/10.1093/oso/9780195183153.003.0006>
- Whelan, D. R., Bambery, K. R., Heraud, P., Tobin, M. J., Diem, M., McNaughton, D., & Wood, B. R. (2011). Monitoring the reversible B to A-like transition of DNA in eukaryotic cells using Fourier transform infrared spectroscopy. *Nucleic Acids Research*, 39(13), 5439-5448. <https://doi.org/10.1093/nar/gkr175>
- Wielscher, M., Vierlinger, K., Kegler, U., Ziesche, R., Gsur, A., & Weinhäusel, A. (2015). Diagnostic Performance of Plasma DNA Methylation Profiles in Lung Cancer, Pulmonary Fibrosis and COPD. *EBioMedicine*, 2(8), 929-936. <https://doi.org/10.1016/j.ebiom.2015.06.025>
- Wood, B. R. (2016). The importance of hydration and DNA conformation in interpreting infrared spectra of cells and tissues. *Chemical Society Reviews*, 45(7), 1980-1998. <https://doi.org/10.1039/C5CS00511F>
- Yang, X., Ou, Q., Qian, K., Yang, J., Bai, Z., Yang, W., Shi, Y., & Liu, G. (2021a).

- Diagnosis of Lung Cancer by ATR-FTIR Spectroscopy and Chemometrics. *Frontiers in Oncology*, 11, 753791. <https://doi.org/10.3389/fonc.2021.753791>
- Yang, X., Ou, Q., Qian, K., Yang, J., Bai, Z., Yang, W., Shi, Y., & Liu, G. (2021b). Diagnosis of Lung Cancer by ATR-FTIR Spectroscopy and Chemometrics. *Frontiers in Oncology*, 11. <https://www.frontiersin.org/articles/10.3389/fonc.2021.753791>
- Zapata, F., López-Fernández, A., Ortega-Ojeda, F., Quintanilla, G., García-Ruiz, C., & Montalvo, G. (2021). Introducing ATR-FTIR Spectroscopy through Analysis of Acetaminophen Drugs: Practical Lessons for Interdisciplinary and Progressive Learning for Undergraduate Students. *Journal of Chemical Education*, 98(8), 2675-2686. <https://doi.org/10.1021/acs.jchemed.0c01231>
- Zhang, F., Huang, Q., Yan, J., & Chen, Z. (2016). Histone Acetylation Induced Transformation of B-DNA to Z-DNA in Cells Probed through FT-IR Spectroscopy. *Analytical Chemistry*, 88(8), 4179-4182. <https://doi.org/10.1021/acs.analchem.6b00400>
- Zhu, W., Love, K., Gray, S. W., & Raz, D. J. (2023). Liquid Biopsy Screening for Early Detection of Lung Cancer: Current State and Future Directions. *Clinical Lung Cancer*, 24(3), 209-217. <https://doi.org/10.1016/j.clcc.2023.01.006>
- Zhu, Z.-Z., Hou, L., Bollati, V., Tarantini, L., Marinelli, B., Cantone, L., Yang, A. S., Vokonas, P., Lissowska, J., Fustinoni, S., Pesatori, A. C., Bonzini, M., Apostoli, P., Costa, G., Bertazzi, P. A., Chow, W.-H., Schwartz, J., & Baccarelli, A. (2012). Predictors of global methylation levels in blood DNA of healthy subjects: A combined analysis. *International Journal of Epidemiology*, 41(1), 126-139. <https://doi.org/10.1093/ije/dyq154>

## **Anexo 1: Consentimiento informado**



**Universidad Autónoma de Querétaro**

Facultad de Química

### **Consentimiento informado para la realización del estudio sobre la**

**“Caracterización de perfiles de infrarrojo séricos y su asociación con el porcentaje de metilación global del ADN en pacientes con cáncer de pulmón”**

Reciba un cordial saludo, antes de decidir participar voluntariamente en este estudio de investigación científica, por favor, dedique el tiempo necesario para leer conscientemente este documento. Tenga la confianza de preguntar y pedir aclaración de cualquier detalle, es importante que antes de firmar este documento comprenda totalmente todos los aspectos mencionados.

**Investigador:** Lic. María Arellano Sosa

**Dirigido por:** Dra. Karla Isabel Lira de León

**Centro de estudio:** Facultad de Química de la Universidad Autónoma de Querétaro.

**Programa de estudio:** Maestría en Química Clínica Diagnóstica

#### **Justificación del estudio**

El cáncer de pulmón representa un importante problema de salud mundial ya que es uno de los tipos de cáncer más frecuentes y el que presenta la mayor tasa de mortalidad. En México, el 75 % de los pacientes con cáncer de pulmón que acuden al servicio de salud, al momento del diagnóstico ya padecen la enfermedad en etapa avanzada, lo que disminuye su probabilidad de supervivencia. Actualmente no hay métodos de diagnóstico temprano para esta enfermedad, por lo tanto, es necesario explorar nuevas herramientas que mejoren el rendimiento diagnóstico y pronóstico a través de técnicas accesibles para toda la población.

### **¿Cuál es el propósito de este estudio?**

Este estudio tiene como objetivo caracterizar perfiles infrarrojos de suero y su asociación con el porcentaje de metilación global del ADN en pacientes con cáncer de pulmón.

### **¿Cuáles son las características que debo tener para participar?**

#### **Puedo participar si:**

- Tengo 40 años o más y llevo más de la mitad de mi vida viviendo en México.
- *Participantes con cáncer de pulmón:* tengo diagnosticado cáncer de pulmón por primera vez y mi tumor primario fue en pulmón y no por metástasis.

#### **No puedo participar si:**

- Tengo una exposición diaria a humo de tabaco (mayor a 1 hora a la semana) o una exposición mayor a 100horas/año a humo de leña, humo de fábricas y chimeneas.
- Previamente me diagnosticaron con otro tipo de cáncer.
- Uno de mis padres ha sido diagnosticado con cáncer de pulmón.
- He sido diagnosticado con asma, fibrosis pulmonar o cualquier otra enfermedad pulmonar crónica.
- *Participantes con cáncer de pulmón:* mi cáncer de pulmón es producto de la metástasis de otro tipo de cáncer. He desarrollado otras enfermedades pulmonares (asma, fibrosis pulmonar, EPOC) a causa del cáncer de pulmón.

### **¿En qué tipo de estudio voy a participar y en qué consiste mi participación?**

El presente estudio es de tipo observacional, transversal y comparativo, lo que significa que no implica ninguna intervención experimental, solo el análisis de datos obtenidos directamente de la muestra sanguínea. Se incluyen dos grupos de estudio: pacientes con cáncer de pulmón y un grupo control sin la enfermedad.

Su participación será en el rol de sujeto de estudio y constará de los siguientes pasos:



1. Lectura y comprensión de este consentimiento, seguido de su firma si está de acuerdo con todo.
2. Realización de un cuestionario que le será proporcionado para recolectar datos.
3. Una extracción de 9 ml en total de muestra sanguínea periférica del área del antebrazo, utilizando el procedimiento convencional de extracción, con las agujas y 2 tubos de recolección completamente nuevos y estériles para cada participante.

La duración de su participación se limita a los pasos anteriores, estimando un tiempo aproximado de 35 minutos como máximo. Su muestra de sangre será tomada por la Q.F.B. Alitzel Arcos de la Cruz con cédula profesional 13357730, y sus datos serán tomados por la Lic. María Arellano Sosa, con cédula profesional 13057599, estos procedimientos se llevarán a cabo en el Laboratorio de enfermedades crónico-degenerativas del Edificio 9 de la Facultad de Química de la Universidad Autónoma de Querétaro.

### **¿Recibiré beneficios por mi participación?**

La participación en este estudio de investigación no será remunerada económicamente, pero tenga en cuenta que posterior a la obtención de los resultados éstos le serán comunicados y podrán consistir en un perfil espectral de su muestra de suero sanguíneo y el resultado de la evaluación del porcentaje global de metilación de su ADN, estos hallazgos puede comunicarlos con su médico tratante y pudieran serle de utilidad. Además, se le comunicará si fue posible la comprobación de la hipótesis de este estudio, lo que significaría que esta investigación ha aportado en el conocimiento del diagnóstico de cáncer de pulmón y usted ha sido parte importante de esa contribución.

### **¿Cuáles son los posibles riesgos que puedo sufrir?**

Su participación en este estudio consta de un único procedimiento de toma de muestra de sangre antes mencionado. Aunque es un procedimiento rutinario, puede

implicar riesgos como la formación de hematoma, dolor y/o malestar en el área de punción, reacción alérgica, incluso mareo o desmayo e infección del área de punción si se expone a microorganismos. Estos riesgos, así como la facilidad o dificultad para hallar las venas para la punción, dependerán de las características individuales.

En caso de que usted lo requiera, se le brindará atención médica por parte de los médicos del consultorio clínico universitario “FARMA-UAQ” a cargo de la Dra. Irhe Sue Acevedo Baltazar y la Dra. Karen Itzel Suárez Torrez.

### **Confidencialidad de los datos**

Los datos obtenidos en esta investigación serán tratados con absoluta confidencialidad, y sólo se utilizarán para los fines antes establecidos de este estudio. Sus datos serán identificados a través de la asignación de un número, para respetar su derecho a la protección y privacidad. Por lo tanto, la información obtenida con esta investigación se dará a conocer manteniendo la confidencialidad de sus datos personales.

### **Derecho a no participar y/o a retirarse en cualquier momento**

Al ser su participación voluntaria, también debe saber que es libre de retirarse en cualquier momento si así lo desea, aún después de haber participado. Su decisión será completamente respetada y no implicará ninguna consecuencia desfavorable para usted. Si lo desea, puede retirarse del estudio permitiendo que se utilicen los datos y muestra obtenidas previamente para los fines antes expuestos o en caso contrario, puede pedir que sus datos sean eliminados y su muestra sea destruida. Para confirmar su retirada del estudio deberá firmar una carta de revocación del consentimiento informado que se le proporcionará en cuanto comunique su deseo de abandonar el estudio a las siguientes responsables: Dra. Karla Isabel Lira de León (correo: [liraleonki@gmail.com](mailto:liraleonki@gmail.com), teléfono: 4421560731) o a la investigadora Lic. María Arellano Sosa ([msosa352@gmail.com](mailto:msosa352@gmail.com), teléfono 9331103874).

### **Comunicación de resultados de la prueba**

Los resultados personales de este estudio le serán comunicados directamente a usted o en su defecto a la persona que usted designe para recibirlos por la Dra. Berenice Lira de León con cédula profesional 6118894 (médico general).

**Nombre de la persona a quien autorizo para recibir el resultado en caso de que yo no pueda:** \_\_\_\_\_

**Parentesco:** \_\_\_\_\_

### **Destino de la muestra biológica**

Las muestras de sangre recolectadas serán resguardadas en congelamiento a -80 °C en un espacio reservado del biobanco de la Unidad de Servicios Clínicos y Diagnóstico Molecular perteneciente a la Facultad de Química de la Universidad Autónoma de Querétaro, Centro Universitario, durante todo el proceso de realización de esta investigación (desde el momento de recolección hasta julio del 2025) y solo serán accesibles para los responsables directos, quienes harán uso estrictamente confidencial de las muestras. Una vez utilizadas serán destruidas siguiendo los protocolos de bioseguridad correspondientes.

### **Aspectos éticos**

Este estudio se apegará a lo manifestado en la Ley Federal de Salud en materia de Investigación para la salud en los capítulos I y V, en referencia a los principios de la no maleficencia, autonomía, beneficencia y justicia para con los participantes, así como con la responsabilidad de protección de la integridad, dignidad, confidencialidad y autodeterminación de su información personal, según lo que establecen las Pautas Éticas internacionales para la Investigación Biomédica en Seres Humanos y la Declaración de Helsinki. El protocolo de este estudio será previamente sometido a la aprobación del Comité de Bioética de la Universidad Autónoma de Querétaro.

Yo, \_\_\_\_\_ he leído conscientemente la información plasmada en este documento acerca del estudio al que he sido invitado a formar parte. Habiendo tenido la oportunidad de expresar mis dudas y preocupaciones sobre los procedimientos, y después de obtener explicaciones y respuestas satisfactorias, afirmo mi decisión de participar voluntariamente en este estudio de investigación, por lo que estoy de acuerdo con que se me tome una muestra sanguínea para los fines que han sido expresados.

- Afirmo mi participación voluntaria en el estudio “Caracterización de perfiles infrarrojos séricos y su asociación con el porcentaje de metilación global del ADN en pacientes con cáncer de pulmón”.
- He leído el consentimiento informado, por lo que entiendo los riesgos que implica la toma de muestra.
- Comprendo que mi participación es voluntaria y tengo el derecho de retirarme en el momento que lo desee, con previo aviso a los responsables, con quienes sé cómo ponerme en contacto.
- Entiendo que no recibiré ninguna remuneración económica por mi participación.
- Comprendo que mis datos serán tratados de manera confidencial, y que solo serán accesibles a los responsables de este estudio.

**Firma del participante**

\_\_\_\_\_

**Nombre y firma de quien ha informado (investigador responsable)**\_\_\_\_\_

**Nombre y firma de testigo 1:**

\_\_\_\_\_

**Nombre y firma de testigo 2:**

\_\_\_\_\_

**Fecha:** \_\_\_\_/\_\_\_\_/\_\_\_\_

## Anexo 2: Carta de revocación del consentimiento

Universidad Autónoma de Querétaro



Facultad de Química

Fecha: \_\_/\_\_/\_\_

### Carta de revocación del consentimiento informado para la participación en el estudio

“Caracterización de perfiles de infrarrojo séricos y su asociación con el porcentaje de metilación global del ADN en pacientes con cáncer de pulmón”

**Investigador:** Lic. María Arellano Sosa

**Dirigido por:** Dra. Karla Isabel Lira de León

**Centro de estudio:** Facultad de Química de la Universidad Autónoma de Querétaro.

**Programa de estudio:** Maestría en Química Clínica Diagnóstica

Por medio de la presente expreso mi decisión de retirar mi participación del estudio que lleva por nombre “Caracterización de perfiles de infrarrojo séricos y su asociación con el porcentaje de metilación global del ADN en pacientes con cáncer de pulmón”, expresando como motivo de mi decisión lo siguiente (puede dejar en blanco si lo desea):

---

---

---

Deseo que puedan hacer uso de los datos que ya han recabado a través de mi muestra de sangre. **Sí** \_\_\_\_ **No** \_\_\_\_

No autorizo el uso de mis datos y deseo la destrucción de mi muestra. **Sí**\_\_**No**\_\_

---

Nombre y firma del paciente

---

Nombre y firma del investigador

### Anexo 3. Cuestionario para participantes



Universidad Autónoma de Querétaro

Facultad de Química

Este formulario está basado en el Cuestionario Internacional de Actividad Física (IPAQ), además, contiene otras preguntas para conocer su hábitos alimenticios y antecedentes clínicos, por favor, tómese el tiempo necesario para responder. Si tiene alguna duda siéntase libre de comentarlo con cualquier persona de su confianza y/o preguntar directamente a los investigadores responsables.

#### Datos personales

Nombre: \_\_\_\_\_

Edad: \_\_\_\_\_

Sexo: \_\_\_\_\_

Ocupación: \_\_\_\_\_

¿Cuánto tiempo lleva dedicándose a ello? \_\_\_\_\_

Lugar de procedencia: \_\_\_\_\_

Colonia y municipio actual en Querétaro: \_\_\_\_\_

Correo electrónico: \_\_\_\_\_

Actividad física
Las preguntas se referirán al tiempo que usted destinó a estar físicamente activo en los últimos 7 días. Por favor responda a cada pregunta aún si no se considera una persona activa. Por favor, piense acerca de las actividades que realiza en su trabajo, como parte de sus tareas en el hogar o en el jardín, moviéndose de un lugar a otro, o en su tiempo libre para la recreación, el ejercicio o el deporte.
Piense en todas las actividades <b>intensas</b> que usted realizó en los últimos 7 días. Las actividades físicas intensas se refieren a aquellas que implican un esfuerzo físico intenso y que lo hacen <b>respirar</b> mucho <b>más intensamente</b> que lo normal.

<p>Piense solo en aquellas actividades físicas que realizó durante por lo menos <b>10 minutos seguidos</b>.</p>		
<p>1. Durante los últimos 7 días, ¿en cuantos realizó actividades físicas intensas tales como levantar pesos pesados, cavar, hacer ejercicios aeróbicos o andar rápido en bicicleta?</p>		
<p>___ días a la semana</p>	<p>___ <b>Ninguna actividad física intensa</b> (vaya a la pregunta 3)</p>	
<p>2. Habitualmente, ¿cuánto tiempo en total dedicó a una actividad física <b>intensa</b> en uno de esos días?</p>		
<p>___ horas por día</p>	<p>___ minutos por día</p>	<p>___ <b>No sabe/no está seguro</b></p>
<p>Piense en todas las actividades <b>moderadas</b> que usted realizó en los últimos 7 días. Las actividades moderadas son aquellas que requieren un esfuerzo físico moderado que lo hace <b>respirar algo más intensamente</b> que lo normal. Piense solo en aquellas actividades físicas que realizó durante por lo menos <b>10 minutos seguidos</b>.</p>		
<p>3. Durante los últimos 7 días, ¿en cuántos días hizo actividades físicas moderadas como transportar pesos livianos, andar en bicicleta a velocidad regular o jugar tenis? No incluya caminar.</p>		
<p>___ días por semana</p>	<p>___ <b>Ninguna actividad física moderada</b> (vaya a la pregunta 5)</p>	
<p>4. Habitualmente, ¿cuánto tiempo en total dedicó a una actividad física moderada en uno de esos días?</p>		
<p>___ horas por día</p>	<p>___ minutos por día</p>	<p>___ <b>No sabe/no está seguro</b></p>
<p>Piense en el tiempo que usted dedicó a <b>caminar</b> en los últimos 7 días. Esto incluye caminar en el trabajo o en la casa, para trasladarse de un lugar a otro, o cualquier otra caminata que usted podría hacer solamente para la recreación, el deporte, el ejercicio o el ocio.</p>		
<p>5. Durante los últimos 7 días, ¿en cuántos caminó por lo menos <b>10 minutos seguidos</b>?</p>		
<p>___ días por semana</p>	<p>___ <b>Ninguna caminata</b> (vaya a la pregunta 7)</p>	
<p>6. Habitualmente, ¿cuánto tiempo en total dedicó a caminar en uno de esos días?</p>		
<p>___ horas por día</p>	<p>___ minutos por día</p>	<p>___ <b>No sabe/no está seguro</b></p>
<p>La última pregunta es acerca del tiempo que pasó usted <b>sentado</b> durante los días hábiles de los últimos 7 días. Esto incluye el tiempo dedicado al trabajo, en la casa, en una clase, y durante el tiempo libre. Puede incluir el tiempo que pasó sentado ante un escritorio, visitando amigos, leyendo, viajando en autobús, sentado o recostado mirando la televisión.</p>		
<p>7. Durante los últimos 7 días ¿cuánto tiempo pasó sentado durante un día hábil</p>		
<p>___ horas por día</p>	<p>___ minutos por día</p>	<p>___ <b>No sabe/no está seguro</b></p>

<b>Hábitos alimenticios</b>		
8. ¿Conoce usted su índice de masa corporal (IMC)?		
<input type="checkbox"/> <b>No</b>	<input type="checkbox"/> <b>Sí, es _____</b>	<input type="checkbox"/> <b>No, pero sé que mi IMC se encuentra en sobrepeso u obesidad.</b>
<b>Antecedentes clínicos</b>		
9. Indique si ha sido diagnosticado con alguna de las enfermedades/condiciones siguientes		
<input type="checkbox"/> <b>Artritis reumatoide</b>	<input type="checkbox"/> <b>Hiperlipidemia/dislipidemia</b>	<input type="checkbox"/> <b>Diabetes tipo I</b>
<input type="checkbox"/> <b>Diabetes tipo II</b>	<input type="checkbox"/> <b>Enfermedad de Parkinson</b>	
13. Esta pregunta es para pacientes con cáncer de pulmón. Por favor indique:		
Estadio en que se encuentra: _____	Tiempo que lleva diagnosticado: _____	Tipo de cáncer: _____



#### **Anexo 4: Consentimiento informado (ampliado)**



**Universidad Autónoma de Querétaro**

Facultad de Química

#### **Consentimiento informado ampliado para la realización del estudio sobre la**

**“Caracterización de perfiles de infrarrojo séricos y su asociación con el porcentaje de metilación global del ADN en pacientes con cáncer de pulmón”**

Reciba un cordial saludo, antes de decidir participar voluntariamente en este estudio de investigación científica, por favor, dedique el tiempo necesario para leer conscientemente este documento. Este consentimiento ampliado se le proporciona para obtener su autorización de conservar su muestra sanguínea y datos para investigaciones posteriores relacionadas con este estudio. Tenga la confianza de preguntar y pedir aclaración de cualquier detalle, es importante que antes de firmar este documento comprenda totalmente todos los aspectos mencionados.

**Investigador:** Lic. María Arellano Sosa

**Dirigido por:** Dra. Karla Isabel Lira de León

**Centro de estudio:** Facultad de Química de la Universidad Autónoma de Querétaro.

**Programa de estudio:** Maestría en Química Clínica Diagnóstica

#### **Justificación del estudio**

El cáncer de pulmón representa un importante problema de salud mundial ya que es uno de los tipos de cáncer más frecuentes y el que presenta la mayor tasa de mortalidad. En México, el 75 % de los pacientes con cáncer de pulmón que acuden al servicio de salud, al momento del diagnóstico ya padecen la enfermedad en etapa avanzada, lo que disminuye su probabilidad de supervivencia. Actualmente no hay métodos de diagnóstico temprano para esta enfermedad, por lo tanto, es necesario

explorar nuevas herramientas que mejoren el rendimiento diagnóstico y pronóstico a través de técnicas accesibles para toda la población.

### **¿Cuál es el propósito de este estudio?**

Este estudio tiene como objetivo caracterizar perfiles infrarrojos de suero y su asociación con el porcentaje de metilación global del ADN en pacientes con cáncer de pulmón.

### **¿Cuáles son las características que debo tener para participar?**

#### **Puedo participar si:**

- Tengo 40 años o más y llevo más de la mitad de mi vida viviendo en México.
- *Participantes con cáncer de pulmón:* tengo diagnosticado cáncer de pulmón por primera vez y mi tumor primario fue en pulmón y no por metástasis.

#### **No puedo participar si:**

- Tengo una exposición diaria a humo de tabaco (mayor a 1 hora a la semana) o una exposición mayor a 100horas/año a humo de leña, humo de fábricas y chimeneas.
- Previamente me diagnosticaron con otro tipo de cáncer.
- Uno de mis padres ha sido diagnosticado con cáncer de pulmón.
- He sido diagnosticado con asma, fibrosis pulmonar o cualquier otra enfermedad pulmonar crónica.
- *Participantes con cáncer de pulmón:* mi cáncer de pulmón es producto de la metástasis de otro tipo de cáncer. He desarrollado otras enfermedades pulmonares (asma, fibrosis pulmonar, EPOC) a causa del cáncer de pulmón.

### **¿En qué tipo de estudio voy a participar y en qué consiste mi participación?**

El presente estudio es de tipo observacional, transversal y comparativo, lo que significa que no implica ninguna intervención experimental, solo el análisis de datos obtenidos directamente de la muestra sanguínea. Se incluyen dos grupos de estudio: pacientes con cáncer de pulmón y un grupo control sin la enfermedad.

Su participación será en el rol de sujeto de estudio y constará de los siguientes pasos:

4. Lectura y comprensión de este consentimiento, seguido de su firma si está de acuerdo con todo.
5. Realización de un cuestionario que le será proporcionado para recolectar datos.
6. Una extracción de 9 ml en total de muestra sanguínea periférica en el área del antebrazo, utilizando el procedimiento convencional de extracción, con las agujas y 2 tubos de recolección completamente nuevos y estériles para cada participante.

La duración de su participación se limita a los pasos anteriores, estimando un tiempo aproximado de 35 minutos como máximo. Su muestra de sangre será tomada por la Q.F.B. Alitzel Arcos de la Cruz con cédula profesional 13357730, y sus datos serán tomados por la Lic. María Arellano Sosa, con cédula profesional 13057599, estos procedimientos se llevarán a cabo en el Laboratorio de enfermedades crónico-degenerativas del Edificio 9 de la Facultad de Química de la Universidad Autónoma de Querétaro.

#### **¿Recibiré beneficios por mi participación?**

La participación en este estudio de investigación no será remunerada económicamente, pero tenga en cuenta que posterior a la obtención de los resultados éstos le serán comunicados y podrán consistir en un perfil espectral de su muestra de suero y el resultado de la evaluación del porcentaje global de metilación de su ADN, estos hallazgos puede comunicarlos con su médico tratante y pudieran serle de utilidad. Además, se le comunicará si fue posible la comprobación de la hipótesis de este estudio, lo que significaría que esta investigación ha aportado en el conocimiento del diagnóstico de cáncer de pulmón y usted ha sido parte importante de esa contribución.

#### **¿Cuáles son los posibles riesgos que puedo sufrir?**

Su participación en este estudio consta de un único procedimiento de toma de muestra de sangre antes mencionado. Aunque es un procedimiento rutinario, puede implicar riesgos como la formación de hematoma, dolor y/o malestar en el área de punción, reacción alérgica, incluso mareo o desmayo e infección del área de punción si se expone a microorganismos. Estos riesgos, así como la facilidad o dificultad para hallar las venas para la punción, dependerán de las características individuales.

En caso de que usted lo requiera, se le brindará atención médica por parte de los médicos del consultorio clínico universitario “FARMA-UAQ”, a cargo de la Dra. Irhe Sue Acevedo Baltazar y la Dra. Karen Itzel Suárez Torrez.

### **Confidencialidad de los datos**

Los datos obtenidos en esta investigación serán tratados con absoluta confidencialidad, y sólo se utilizarán para los fines antes establecidos de este estudio y proyectos posteriores dirigidos por la Dra. Karla Isabel Lira de León, de la misma línea de investigación (cáncer de pulmón) e institución educativa (Facultad de Química de la UAQ). Sus datos serán identificados a través de la asignación de un número, para respetar su derecho a la protección y privacidad. Por lo tanto, la información obtenida con esta investigación se dará a conocer manteniendo la confidencialidad de sus datos personales.

### **Derecho a no participar y/o a retirarse en cualquier momento**

Al ser su participación voluntaria, también debe saber que es libre de retirarse en cualquier momento si así lo desea, aún después de haber participado. Su decisión será completamente respetada y no implicará ninguna consecuencia desfavorable para usted. Si lo desea, puede retirarse del estudio permitiendo que se utilicen los datos y muestra obtenidas previamente para los fines antes expuestos o en caso contrario, puede pedir que sus datos sean eliminados y su muestra sea destruida. Para confirmar su retirada del estudio deberá firmar una carta de revocación del consentimiento informado que se le proporcionará en cuanto comunique su deseo de

abandonar el estudio a las siguientes responsables: Dra. Karla Isabel Lira de León ([liraleonki@gmail.com](mailto:liraleonki@gmail.com), teléfono: 4421560731) o a la investigadora Lic. María Arellano Sosa ([msosa352@gmail.com](mailto:msosa352@gmail.com), teléfono 9331103874).

### **Comunicación de resultados de la prueba**

Los resultados personales de este estudio le serán comunicados directamente a usted o en su defecto a la persona que usted designe para recibirlos por la Dra. Berenice Lira de León con cédula profesional 6118894 (médico general).

**Nombre de la persona a quien autorizo para recibir el resultado en caso de que yo no pueda:** \_\_\_\_\_

**Parentesco:** \_\_\_\_\_

### **Destino de la muestra biológica**

Las muestras de sangre recolectadas serán resguardadas en congelamiento a -80 °C en un espacio reservado del biobanco de la Unidad de Servicios Clínicos y Diagnóstico Molecular perteneciente a la Facultad de Química de la Universidad Autónoma de Querétaro, Centro Universitario, por un periodo máximo de 2 años y solo serán accesibles para los responsables directos, quienes harán uso estrictamente confidencial de las muestras. Una vez utilizadas serán destruidas siguiendo los protocolos de bioseguridad correspondientes.

### **Aspectos éticos**

Este estudio se apegará a lo manifestado en la Ley Federal de Salud en materia de Investigación para la salud en los capítulos I y V, en referencia a los principios de la no maleficencia, autonomía, beneficencia y justicia para con los participantes, así como con la responsabilidad de protección de la integridad, dignidad, confidencialidad y autodeterminación de su información personal, según lo que establecen las Pautas Éticas internacionales para la Investigación Biomédica en Seres Humanos y la Declaración de Helsinki. El protocolo de este estudio será

previamente sometido a la aprobación del Comité de Bioética de la Universidad Autónoma de Querétaro.

Yo, \_\_\_\_\_ he leído conscientemente la información plasmada en este documento acerca del estudio al que he sido invitado a formar parte. Habiendo tenido la oportunidad de expresar mis dudas y preocupaciones sobre los procedimientos, y después de obtener explicaciones y respuestas satisfactorias, afirmo mi decisión de participar voluntariamente en este estudio de investigación, por lo que estoy de acuerdo con que se me tome una muestra sanguínea para los fines que han sido expresados.

- Afirmo mi participación voluntaria en el estudio “Caracterización de perfiles infrarrojos séricos y su asociación con el porcentaje de metilación global del ADN en pacientes con cáncer de pulmón”.
- He leído el consentimiento informado, por lo que entiendo los riesgos que implica la toma de muestra.
- Comprendo que mi participación es voluntaria y tengo el derecho de retirarme en el momento que lo desee, con previo aviso a los responsables, con quienes sé cómo ponerme en contacto.
- Entiendo que no recibiré ninguna remuneración económica por mi participación.
- Comprendo y autorizo que mi muestra sea almacenada por un plazo máximo de 2 años y que la información obtenida puede ser utilizada en investigaciones posteriores (proyectos de tesis) relacionados con el mismo tema de investigación para el mismo programa de estudio de la institución educativa (UAQ) y dirigidos por la Dra. Karla Isabel Lira de León.
- Comprendo que mis datos serán tratados de manera confidencial, y que solo serán accesibles a los responsables del estudio.

**Firma del participante**

---

**Nombre y firma de quien ha informado (investigador responsable)**\_\_\_\_\_

**Nombre y firma de testigo 1:**


\_\_\_\_\_

**Nombre y firma de testigo 2:**

\_\_\_\_\_

**Fecha:** \_\_\_\_/\_\_\_\_/\_\_\_\_

# Anexo 5. Análisis de plagio



Página 2 of 158 - Descripción general de integridad

Identificador de la entrega trn:oid::7696:477927373

2% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

Filtrado desde el informe

- Bibliografía
- Coincidencias menores (menos de 20 palabras)
- Trabajos entregados

Fuentes principales

2% Fuentes de Internet

0% Publicaciones

0% Trabajos entregados (trabajos del estudiante)

Marcas de integridad

N.º de alertas de integridad para revisión

No se han detectado manipulaciones de texto sospechosas.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.