



UNIVERSIDAD DE CONCEPCIÓN  
FACULTAD DE CIENCIAS FORESTALES  
INGENIERIA EN BIOTECNOLOGIA VEGETAL

**APLICACIÓN DE TÉCNICAS ESPECTROSCÓPICAS PARA LA PREDICCIÓN  
DE LA CAPACIDAD MADURATIVA EN CULTIVOS EMBRIOGÉNICOS DE  
*PINUS RADIATA***

Tesis presentada a la Facultad de Ciencias Forestales de la Universidad de  
Concepción para optar al título profesional de Ingeniero en Biotecnología  
Vegetal

POR: Emilio Morales Navarro

Profesor Guía: Rodrigo Hasbún Zaror

Marzo, 2025

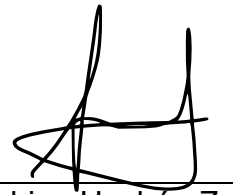
Concepción, Chile

© 2025, Emilio Antonio Morales Navarro

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento.

**APLICACIÓN DE TÉCNICAS ESPECTROSCÓPICAS PARA LA PREDICCIÓN  
DE LA CAPACIDAD MADURATIVA EN CULTIVOS EMBRIOGÉNICOS DE  
*PINUS RADIATA***

Profesor Guía



---

Rodrigo Hasbún Zaror

Profesor Asociado

Ingeniero Forestal, Dr.

Profesora Guía



---

Rosario Castillo Felices

Colaboradora Externa

Química, Dra.

## TABLA DE CONTENIDO

RESUMEN .....	viii
ABSTRACT .....	1
I. INTRODUCCIÓN .....	2
II. METODOLOGÍA .....	5
2.1 Material Vegetal.....	5
2.2 Captura de imágenes hiperespectrales .....	6
2.3 Pretratamiento de datos .....	7
2.4 Análisis de imágenes hiperespectrales .....	8
2.5 Espectroscopia FT-IR.....	9
2.6 Contabilización de embriones.....	10
2.7 Análisis estadístico .....	11
III. RESULTADOS .....	13
3.1. Imágenes RGB.....	13
3.2. Perfiles espectrales brutos .....	13
3.3. Análisis multivariado mediante componentes principales .....	18
3.4. Identificación de bandas espectrales relevantes.....	19
3.5. Generalización del diagrama de dispersión simple de dos variables (Biplot) .....	22
3.6. Resultados de análisis PLS-DA .....	23
3.7 Análisis incluyendo todos los genotipos.....	25
3.7.1. Análisis Multivariado Mediante Componentes Principales .....	25
3.7.2. Identificación de bandas espectrales relevantes.....	27

3.8 Resultados del análisis FTIR .....	30
3.8.1 Perfiles espectrales brutos .....	30
3.8.2 Identificación de bandas espectrales relevantes .....	30
3.9 Contabilización de embriones .....	32
IV.    DISCUSIÓN .....	33
V.    CONCLUSIONES .....	38
VI.    BIBLIOGRAFÍA .....	42

## TABLA DE FIGURAS

- Figura 3. 1:** Análisis espectral y diferenciación de los genotipos 1, 2 y 3 de *Pinus radiata* mediante imágenes hiperespectrales. (A) Imágenes RGB correspondientes a las masas embriogénicas en estado de multiplicación mediante el método de disgregación. (B) Espectros medios brutos obtenidos a partir de los datos hiperespectrales. (C) y (D) Análisis de componentes principales (PCA) utilizando la segunda derivada (ventana = 25 puntos) de la escala cromática para visualizar las diferencias espectrales. (E) Gráfico de carga de las longitudes de onda más relevantes para el PC2, que explica el 13,08% de la varianza total. El análisis se realizó utilizando la segunda derivada con una ventana de 25 puntos. (F) Biplot del PC1 versus el PC2. Las muestras representan tres masas embriogénicas por genotipo, y el triángulo rojo resalta los píxeles clave para identificar patrones de caracterización. .... 14
- Figura 3. 2:** Análisis espectral y diferenciación de genotipos de *Pinus radiata* mediante imágenes hiperespectrales. (A) Collage de imágenes RGB capturadas con una cámara de 50 MP, correspondientes a los genotipos 4, 5, 6 y 7 de masas embriogénicas en estado de multiplicación mediante el método de aglomeración. (B) Espectros medios brutos de los genotipos 4, 5, 6 y 7, obtenidos a partir de los datos hiperespectrales. (C) Análisis de componentes principales (PCA) utilizando la segunda derivada (ventana = 35 puntos) mostrando el PC1 y (D) PC2. (E) Biplot del PC1 versus el PC2. (F) Gráfico de carga de las longitudes de onda más relevantes para PC1 y PC2, destacando las bandas espectrales clave. .... 16
- Figura 3. 3:** Resultados de clasificación PLS-DA para las muestras de masas embriogénicas de *Pinus radiata* en etapa de multiplicación, diferenciadas por el método de propagación. (A) Clasificación PLS-DA de los genotipos 1, 2 y 3, multiplicados mediante el método de disgregación. Se establecen tres clases: Class 1 (genotipo 1), Class 2 (genotipo 2) y Class 3 (genotipo 3). (B) Clasificación PLS-DA de los genotipos 4, 5, 6 y 7, multiplicados mediante el método de aglomerado. Se definen cuatro clases: Class 1 (genotipo 4), Class 2 (genotipo 5), Class 3 (genotipo 6) y Class 4 (genotipo 7). .... 24
- Figura 3. 4:** Análisis de componentes principales (PCA) aplicado a masas embriogénicas de *Pinus radiata*. (A) PC1 de los genotipos 1, 2, 3, 4, 5, 6 y 7, calculado con segunda derivada (ventana de 35 puntos), explicando el 72,01% de la varianza total. Los genotipos 1, 2 y 3 corresponden a la etapa de maduración, mientras que los genotipos 4, 5, 6 y 7 se encuentran en la etapa de multiplicación. (B) PC2 de los mismos genotipos, con segunda derivada (ventana de 35 puntos), explicando el 10,89% de la varianza total, manteniendo la diferenciación entre etapas. (C) Gráfico de cargas de PC1, resaltando las bandas

espectrales más relevantes en la variabilidad de las muestras. (D) Biplot de PC1 genotipos, ya que los genotipos 1, 2 y 3 generaron embriones, mientras que los genotipos 4, 5, 6 y 7 no lo hicieron versus PC2, visualizando los patrones de distribución de los genotipos en función de sus espectros, permitiendo identificar tendencias y agrupamientos .....26

**Figura 3. 5** Análisis espectroscópico de extractos de metanol obtenidos de masas embriogénicas de *Pinus radiata*. (A) Gráfico de absorbancia bruta de las muestras, donde las masas embriogénicas capaces de generar embriones se representan en rojo, y aquellas que no generan embriones en verde. (B) Biplot de PC1 (38,71% de varianza explicada) vs. PC2 (24,28% de varianza explicada), mostrando patrones de diferenciación entre las muestras que generan embriones (rojo) y las que no (verde). (C) Gráfico de absorbancia con segunda derivada (ventana de 35 puntos), resaltando las bandas espectrales más relevantes en los extractos de metanol de masas embriogénicas de *Pinus radiata*.....31

**Figura 3. 6:** Diagramas de caja que muestran el promedio del número de embriones por gramo de peso fresco en masas embriogénicas de *Pinus radiata* en etapa de maduración. Se incluyen los genotipos 1, 2, 3, diferenciando los métodos de propagación: disgregación (genotipos 1, 2 y 3). (A) Diagrama de caja con el promedio del número de embriones normales por gramo de peso fresco en los distintos genotipos. (B) Diagrama de caja con el promedio del número de embriones amorfos por gramo de peso fresco en los mismos genotipos y condiciones experimentales.....32

## RESUMEN

La fase de maduración en los cultivos embriogénicos de *Pinus radiata* es un punto crítico para la producción de plantas, ya que actualmente el número de embriones generados solo puede determinarse al final del proceso. Este período, que dura entre 4 y 5 meses, incrementa significativamente los costos y la complejidad de los programas de mejoramiento genético y las operaciones silvícolas.

En este estudio, se evaluó la aplicación de técnicas espectroscópicas, específicamente imágenes hiperespectrales y espectroscopía FTIR, para predecir la capacidad madurativa de los cultivos embriogénicos mediante análisis multivariados. Se implementaron distintos tratamientos para la normalización de los datos espectrales y se aplicaron análisis de componentes principales (PCA) y regresión por mínimos cuadrados parciales discriminante (PLS-DA) con el fin de identificar patrones espectrales en los genotipos evaluados.

Los resultados indicaron que las imágenes hiperespectrales presentan una capacidad superior para discriminar entre grupos de genotipos en comparación con FTIR. El análisis PCA permitió diferenciar claramente dos conjuntos de genotipos: los grupos 1, 2 y 3 frente a los grupos 4, 5, 6 y 7, con una varianza acumulada de hasta el 72,01%. Por otro lado, el modelo PLS-DA mostró un desempeño destacado, con valores de especificidad de 1,00, sensibilidad de 0,99 y precisión de 0,99 al distinguir masas embriogénicas capaces de generar embriones de aquellas no generadoras. En contraste, el análisis mediante FTIR presentó limitaciones significativas en términos de sensibilidad y resolución para la identificación de patrones claros.

## ABSTRACT

The maturation phase in *Pinus radiata* embryogenic cultures is a critical point for plant production since the number of embryos generated can only be determined at the end of the process. This period, lasting between 4 and 5 months, significantly increases the costs and complexity of genetic improvement programs and silvicultural operations.

This study evaluated the application of spectroscopic techniques, specifically hyperspectral imaging and FTIR spectroscopy, to predict the maturation capacity of embryogenic cultures through multivariate analysis. Various treatments were applied to normalize spectral data, and principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA) were used to identify spectral patterns in the evaluated genotypes.

The results indicated that hyperspectral imaging has a superior ability to discriminate between genotype groups compared to FTIR. PCA clearly differentiated two sets of genotypes: groups 1, 2, and 3 versus groups 4, 5, 6, and 7, with an accumulated variance of up to 72.01%. Meanwhile, the PLS-DA model showed outstanding performance, with specificity values of 1.00, sensitivity of 0.99, and accuracy of 0.99 in distinguishing embryogenic masses capable of generating embryos from non-generating ones. In contrast, FTIR analysis showed significant limitations in terms of sensitivity and resolution for identifying clear patterns.

## I.INTRODUCCIÓN

En el campo de la biotecnología vegetal, el uso eficiente de herramientas biotecnológicas es crucial para mejorar la producción de plantas. La embriogénesis somática permite la generación de plantas a partir de células somáticas, ha demostrado su relevancia en áreas como la agronomía y silvicultura (Akhtar, 2012). Esta técnica facilita la obtención de embriones, semillas sintéticas y la generación de plantas de élite (Aquea *et al.*, 2008). Sin embargo, algunos de los desafíos de esta herramienta es la estabilidad genética en plantas, crio-preservación y producción de embriones viables.

La embriogénesis somática está influenciada por factores como la temperatura, tiempo de subcultivo, la composición del medio de cultivo, concentración de macroelementos, microelementos, azúcares, vitaminas, hormonas y gelificantes, que son cruciales para la obtención de plantas élite (Nic-Can *et al.*, 2016). El proceso completo abarca las etapas de inducción, maduración, crecimiento y germinación de los embriones somáticos (García-Mendiguren *et al.*, 2016). La fase de maduración es crítica y prolongada, en la que los embriones deben desarrollarse para germinar y convertirse en plantas viables. El principal problema es que la capacidad del material vegetal para generar embriones viables solo se puede evaluar al final de la fase de maduración. Esto significa

que, si el material no es adecuado, se habrá invertido tiempo y recursos en un proceso que puede no resultar exitoso (Lineros *et al.*, 2018).

Las herramientas espectroscópicas han emergido como soluciones prometedoras para predecir el comportamiento de las plantas, permitiendo una evaluación a través de "huellas dactilares" espectroscópicas. Técnicas como la espectroscopia infrarroja por transformada de Fourier (FT-IR) y análisis de imágenes hiperespectrales han demostrado ser útiles en diversos estudios vegetales, incluyendo la medición del contenido de nitrógeno (Miao *et al.*, 2023), el análisis de nutrientes (Prananto *et al.*, 2020), la determinación de la calidad de tubérculos (Su *et al.*, 2021), el monitoreo de cultivos (Polder *et al.*, 2024), la detección de metales (Junmeng *et al.*, 2022), e incluso la predicción de oligoelementos en plantas (Cozzolino, 2004). . En el caso de *Pinus radiata*, especie de gran relevancia forestal, económica e industrial (INFOR, 2023), la embriogénesis somática enfrenta desafíos que limitan su escalabilidad, especialmente debido a la dificultad para predecir y controlar la capacidad madurativa de los embriones. La variabilidad en el rendimiento de las plantas propagadas y la prolongada fase de maduración generan costos elevados, haciendo crucial la selección precisa del material embriogénico (Linero *et al.*, 2023). En este contexto, el uso de espectroscopía para predecir con precisión la capacidad madurativa durante la fase de multiplicación se presenta como una

alternativa prometedora para mejorar la eficiencia en programas de reforestación y propagación vegetativa.

**Pregunta de investigación:** ¿Cuál es la capacidad predictiva de las técnicas espectroscópicas en la determinación del potencial madurativo de cultivos embriogénicos somáticos de *Pinus radiata*?

### **Hipótesis**

- Hipótesis nula ( $H_0$ ): Las técnicas espectroscópicas basadas en imágenes no permiten predecir la capacidad madurativa de los cultivos embriogénicos somáticos de *Pinus radiata* durante la fase de multiplicación.
- Hipótesis alternativa ( $H_1$ ): Las técnicas espectroscópicas basadas en imágenes permiten predecir la capacidad madurativa de los cultivos embriogénicos somáticos de *Pinus radiata* durante la fase de multiplicación.

**Objetivo general:** Comparar el desempeño de distintas técnicas espectroscópicas aplicadas en fase de multiplicación para la discriminación de genotipos con capacidades madurativas contrastan.

## II. METODOLOGÍA

### 2.1 Material Vegetal

Se utilizaron cultivos embriogénicos de *Pinus radiata* de siete genotipos, proporcionados por la empresa Bioforest Spa, Arauco (Chile). Los análisis se realizaron en dos momentos del proceso embriogénico: multiplicación y final de la etapa de maduración. El medio de multiplicación consistió en 30 ml de medio Litvay semisólido modificado (mLM) contenido en placas de Petri de 16 mm. El medio de cultivo se complementó con 4,5  $\mu\text{M}$  Ácido 2,4-diclorofenoxiacético (2,4-D), 2,7 $\mu\text{M}$  6 bencilaminopurina (BAP), 30 g  $\text{L}^{-1}$  sacarosa y 3 g  $\text{L}^{-1}$  Phytigel®. Los genotipos 1, 2 y 3 fueron multiplicados por el método de disgregación en medio líquido y disposición en papel filtro sobre el medio de cultivo semisólido. En cambio, los genotipos 4, 5, 6 y 7 fueron multiplicados usando el método de división de aglomerados y disposición directa en el medio de cultivo semisólido. El medio de cultivo de maduración utilizado fue mLM suplementado 90 $\mu\text{M}$  ácido abscísico, 60 g  $\text{L}^{-1}$  sacarosa y 10 g  $\text{L}^{-1}$  Phytigel™. Para ambos medios el pH se ajustó a 5,6 antes de la esterilización en autoclave a 121°C y 1 atmósfera de presión (Linerros *et al.*, 2023). Las masas embriogénicas fueron mantenidas en

condiciones estériles y en oscuridad a 25°C durante la multiplicación y maduración.

## **2.2 Captura de imágenes hiperespectrales**

Para la captura de imágenes hiperespectrales, se utilizó el software Spectronon™ junto a una cámara PIKA L (Resonon), la cual opera en un rango de 400-1000 nm, abarcando los espectros visibles (VIS) e infrarrojo cercano (NIR). La cámara fue montada en un pedestal, manteniendo una distancia de 30 cm respecto a la muestra. Las imágenes se capturaron a temperatura ambiente y bajo condiciones controladas de iluminación.

Para evitar interferencias por la reflectancia de la luz en el plástico, la placa de Petri que contenía el material se mantuvo abierta durante la captura. Una vez finalizado el proceso, el material quedó disponible para requerimientos posteriores. Las imágenes de los genotipos 1, 2 y 3 se tomaron con fondo negro, mientras que las imágenes de los genotipos 4, 5, 6 y 7 se capturaron con fondo blanco.

### 2.3 Pretratamiento de datos

El preprocesamiento general aplicado a todas las masas embriogénicas consistió en generar una máscara o recorte de las imágenes espectrales para seleccionar únicamente los datos correspondientes a las masas embriogénicas de *Pinus radiata* de cada genotipo. Estas masas se organizaron en tres composiciones de imágenes: una para los genotipos 1, 2 y 3; otra para los genotipos 4, 5, 6 y 7; y una última composición que incluyó los siete genotipos.

Posteriormente, se aplicó un pretratamiento espectral adicional en cada análisis, con el objetivo de eliminar ruidos en las bandas y optimizar la calidad de los datos para su procesamiento. Para análisis de los genotipos 1, 2 y 3, se aplicó un pretratamiento de segunda derivada con una ventana de 25 puntos ( $W = 25$ ). En el caso de los genotipos 4, 5, 6 y 7, se utilizó un pretratamiento de segunda derivada con una ventana de 35 puntos ( $W = 35$ ). Y para la composición que incluía todos los genotipos, se aplicó un pretratamiento uniforme mediante una segunda derivada con una ventana de 35 puntos ( $W = 35$ ). Estos pretratamientos permitieron mejorar la claridad de las señales espectrales, reducir el ruido y normalizar los datos espectrales, lo que facilitó la identificación de patrones y diferencias entre los grupos analizados.

## **2.4 Análisis de imágenes hiperespectrales**

Una vez aplicados los pretratamientos, se realizaron análisis de componentes principales (PCA), diagrama de dispersión simple para comparar dos variantes (Biplot) y análisis de clasificación mediante el modelo de PLS-DA (Partial Least Squares Discriminant Analysis) para evaluar la capacidad discriminativa de los datos hiperespectrales.

Para las imágenes correspondientes a los genotipos 1, 2 y 3, se utilizó el software Spectronon para generar collage con imágenes, se visualizaron las bandas espectrales brutas, se generaron análisis de componentes principales (PCA) y se obtuvo el gráfico de carga correspondiente al segundo componente principal (PC2). Además, mediante la herramienta HYPER-Tools, se creó un diagrama de dispersión simple para comparar dos variantes (Biplot) y análisis PLS-DA.

Para la imagen de los genotipos 4, 5, 6 y 7, también se empleó el software Spectronon. En este caso, se realizó el recorte de imagen, visualización de bandas espectrales brutas, se generaron componentes principales PCA1 y PCA2 con su respectivo gráfico de carga y se creó Biplot, Asimismo, mediante Hypertools se ejecutó análisis PLS-DA

Finalmente, para el análisis que incluyó todos los genotipos, se utilizó HYPER-Tools. En este caso, se generaron los componentes principales PCA1 y PCA2, se obtuvieron los gráficos de carga correspondiente a PC1, y se creó un Biplot para visualizar las relaciones entre los genotipos.

## **2.5 Espectroscopia FT-IR**

Después de capturar las imágenes de las masas embriónicas, se prepararon las muestras para su análisis mediante espectroscopia infrarroja por transformada de Fourier (FT-IR). Cada muestra consistió en un duplicado de 100 mg de masa embriónica molida en fase de multiplicación de cada uno de los 7 genotipos, mezclada con 1 mL de metanol. Este análisis se realizó utilizando un espectrofotómetro IRAffinity-1S (SHIMADZU) en modo de absorbancia. Las mediciones espectroscópicas se llevaron a cabo bajo condiciones controladas, a temperatura ambiente. Inicialmente, se depositó una gota de la mezcla en estado líquido sobre el porta muestras y se realizó la primera medición. Luego, se esperó entre 5 y 10 minutos hasta que la muestra estuviera seca, y se realizó una segunda medición. Todas las mediciones se realizaron en el rango espectral de 4000 a 400  $\text{cm}^{-1}$ .

En total, se obtuvieron 28 datos espectrales (4 por genotipo: 2 en estado líquido y 2 en estado seco). Estos datos se procesaron utilizando el software LabSolutions IR, lo que permitió identificar los picos característicos de los compuestos presentes en cada muestra, creando así una huella dactilar basada en las vibraciones moleculares. Adicionalmente, se generó un gráfico con los datos espectrales brutos, en el cual se diferenciaron las masas embriogénicas que generan embriones (genotipos 1, 2 y 3) de las que no generan embriones (genotipos 4, 5, 6 y 7), utilizando colores distintos para cada grupo. Con estos datos, también se realizó un análisis de Componentes Principales (PCA), aplicando un pretratamiento de segunda derivada con una ventana de 35 puntos ( $W = 35$ ). Finalmente, se generó biplot que incluyó los datos espectrales de cada muestra y su duplicado, permitiendo visualizar las relaciones y diferencias entre los genotipos en función de su composición molecular.

## **2.6 Contabilización de embriones**

Para validar y comparar los resultados hiperespectrales tomados en fase de multiplicación, con los datos reales de la capacidad madurativa de los embriones, se realizó una contabilización de embriones al cabo de 5 meses desde que el material ingresó a la etapa de maduración. Este proceso fue llevado a cabo por la empresa Arauco, en un ambiente estéril, utilizando una lupa para diferenciar

entre embriones normales y embriones amorfos. Una vez finalizada la contabilización, se generó un promedio de la cantidad de embriones normales y amorfos por genotipo. Con estos datos, se elaboró un diagrama de caja que permitió visualizar la variabilidad en la producción de embriones entre los genotipos generadores de embriones (genotipos 1, 2 y 3).

## **2.7 Análisis estadístico**

Los análisis estadísticos incluyeron pruebas multivariadas para la reducción de dimensionalidad y la clasificación. Se aplicaron técnicas como PCA y PLS-DA para identificar patrones espectrales específicos y sus relaciones con variables biológicas (capacidad madurativa y diferencias genotípicas). Se analizaron las métricas de desempeño (sensibilidad, especificidad, precisión y tasa de error) para validar los resultados de los modelos aplicados. Además, se utilizó el valor de  $p$  para evaluar la significancia estadística de las diferencias observadas en conjunto a diagrama de caja entre los grupos analizados. Un valor de  $p < 0.05$  se consideró estadísticamente significativo, lo que indica que las diferencias no son producto del azar.

En cuanto al diseño experimental, se utilizaron réplicas para garantizar la robustez de los análisis. En la fase de multiplicación, para los genotipos 1, 2 y 3,

se trabajó con una masa por placa, aplicando el método de disgregación, utilizando 3 placas por genotipos. Por otro lado, para los genotipos 4, 5, 6 y 7, se utilizó 1 placa con 8 masas por genotipos, aplicando el método de aglomerado. En la fase de maduración, se utilizaron 7 placas por genotipo para el conteo de embriones. Este diseño permitió obtener datos representativos y asegurar la confiabilidad de los resultados.

### **III. RESULTADOS**

#### **3.1. Imágenes RGB**

El análisis inicial de los genotipos 1, 2 y 3 se realizó mediante la observación de imágenes RGB (Figura 3.1A), donde se identificaron diferencias leves en los tonos de coloración de las masas embriogénicas entre los genotipos. En el caso de los genotipos 1 y 2, sus masas embriogénicas mostraron una coloración estable y uniforme, indicando probablemente un desarrollo similar. En contraste, el genotipo 3 presentó una notoria diferenciación, con dos masas más oscuras y una más clara, lo que podría indicar una mayor variabilidad en su composición o esta de desarrollo. El genotipo presentó masas de tamaño superior al resto.

#### **3.2. Perfiles espectrales brutos**

Los espectros medios no procesados (Figura 3.1B) exhibieron patrones de absorción característicos en el rango visible (400-700 nm) e infrarrojo cercano (700-1000 nm). En la región azul-verde (400-500 nm), se observó variabilidad intergenotípica, lo que podría estar asociado con diferencias en la concentración de pigmentos fotosintéticos, como clorofilas y carotenoides (Huang *et al.*, 2015).

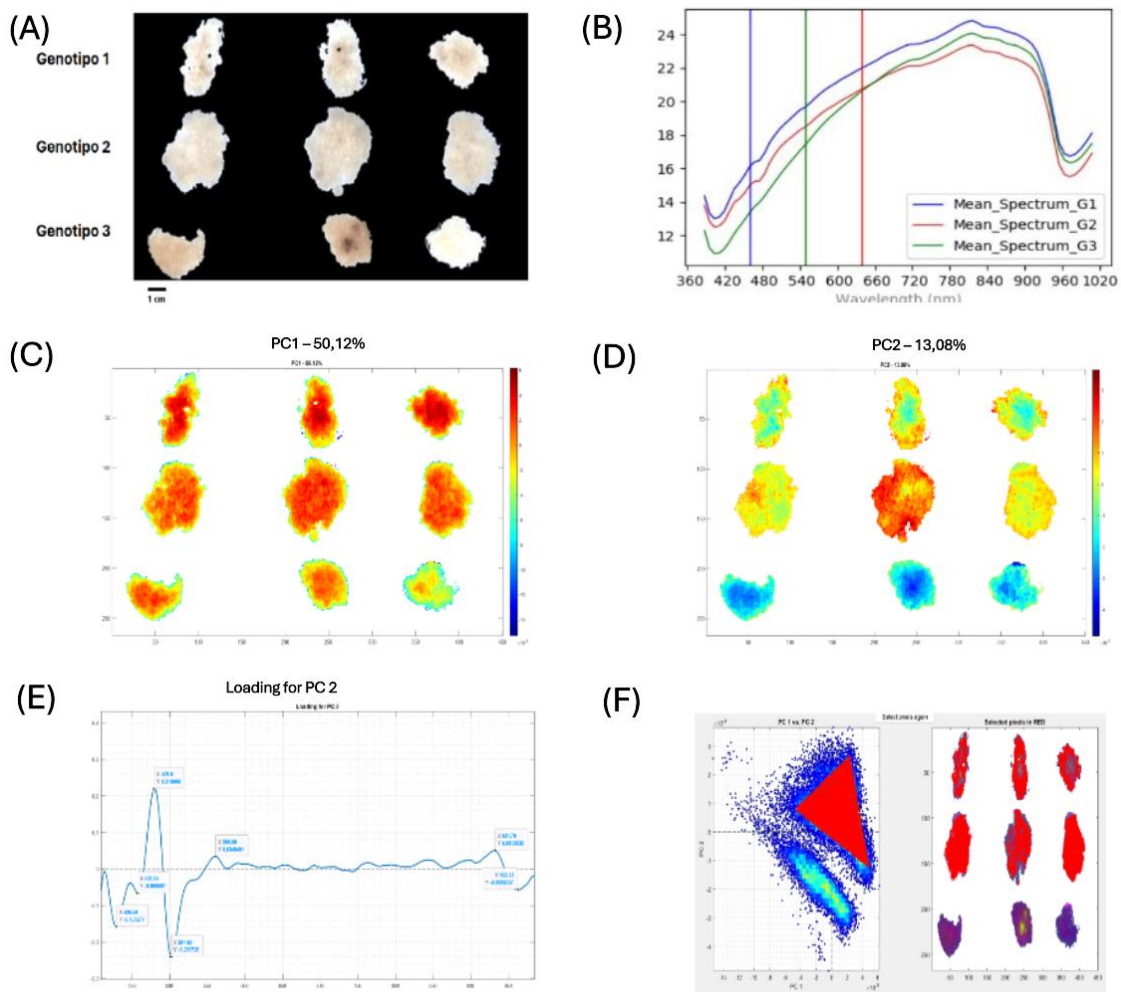


Figura 3.1: Análisis espectral y diferenciación de los genotipos 1, 2 y 3 de *Pinus radiata* mediante imágenes hiperespectrales. (A) Imágenes RGB correspondientes a las masas embriogénicas en estado de multiplicación mediante el método de disgregación. (B) Espectros medios brutos obtenidos a partir de los datos hiperespectrales. (C) y (D) Análisis de componentes principales (PCA) utilizando la segunda derivada (ventana = 25 puntos) de la escala cromática para visualizar las diferencias espectrales. (E) Gráfico de carga de las longitudes de onda más relevantes para el PC2, que explica el 13,08% de la varianza total. El análisis se realizó utilizando la segunda derivada con una ventana de 25 puntos. (F) Biplot del PC1 versus el PC2. Las muestras representan tres masas embriogénicas por genotipo, y el triángulo rojo resalta los píxeles clave para identificar patrones de caracterización.

Por otro lado, los espectros de los genotipos 1, 2 y 3 en el rango del infrarrojo cercano (700-1000 nm), presentan firmas espectrales que mostraron superposiciones entre los grupos, sugiriendo similitudes en la composición hídrica y la estructura molecular de las masas embriogénicas (Huang *et al.*, 2024). Este hallazgo resalta la necesidad de aplicar transformaciones matemáticas, como la segunda derivada, para realzar diferencias críticas no evidentes en los espectros brutos, una práctica ampliamente recomendada en el análisis de datos hiperespectrales complejos (LI *et al* 2022).

El análisis visual de los genotipos 4, 5, 6 y 7 se llevó a cabo mediante la observación de imágenes RGB (Figura 3.2A). En este caso, no se identificaron diferencias significativas en los colores entre los genotipos, aunque se observó que los genotipos 4 y 5 presentaron una apariencia más opaca y oscura en comparación con los genotipos 6 y 7, cuyas masas embriogénicas mostraron una apariencia más cristalina y brillante. Estas diferencias en la tonalidad y brillo podrían estar relacionadas con variaciones en la composición o el estado de desarrollo de las masas embriogénicas entre los genotipos analizados. Los tamaños de las masas fueron bastante uniformes.

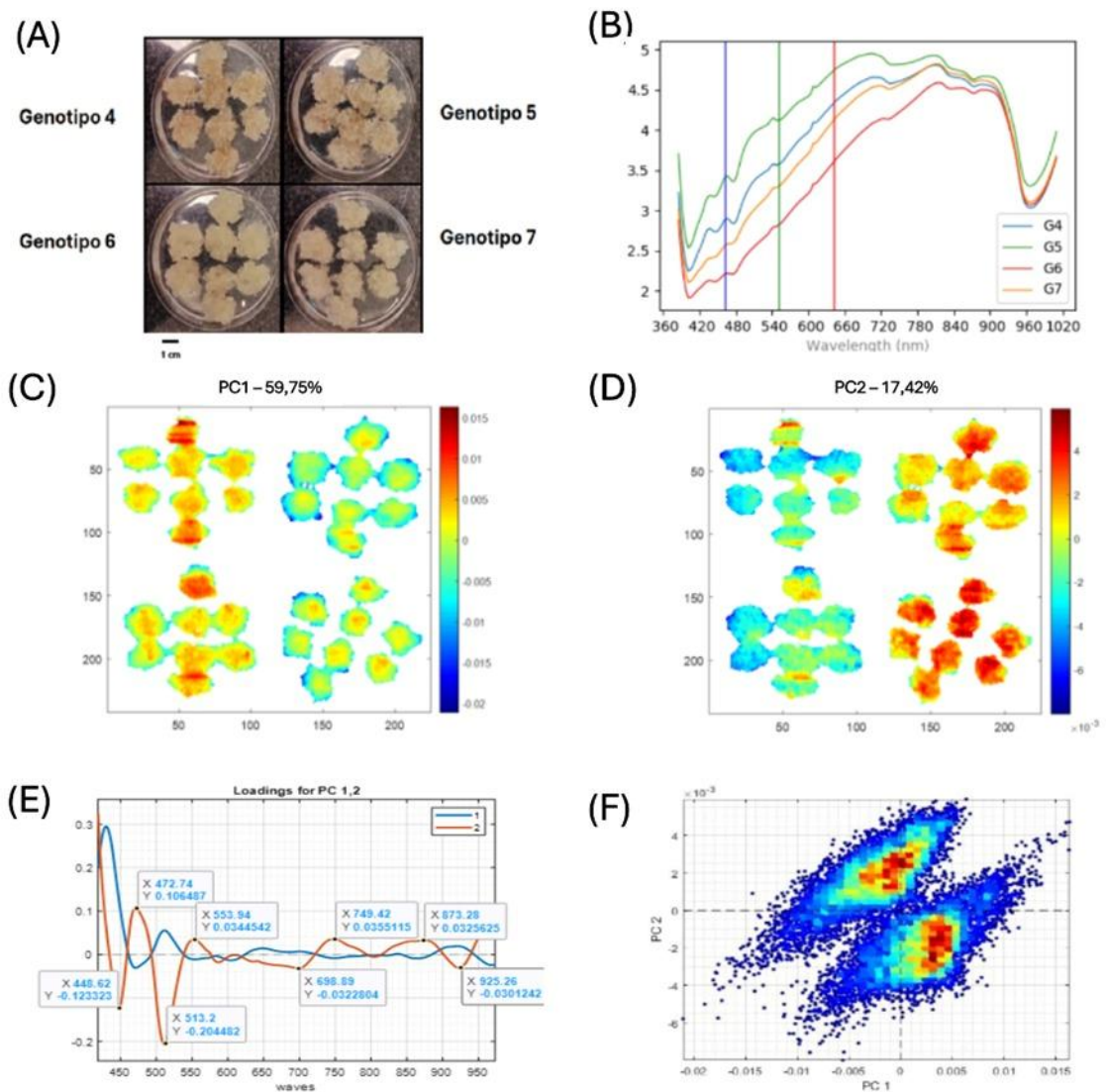


Figura 3.2: Análisis espectral y diferenciación de genotipos de *Pinus radiata* mediante imágenes hiperespectrales. (A) Collage de imágenes RGB capturadas con una cámara de 50 MP, correspondientes a los genotipos 4, 5, 6 y 7 de masas embriogénicas en estado de multiplicación mediante el método de aglomeración. (B) Espectros medios brutos de los genotipos 4, 5, 6 y 7, obtenidos a partir de los datos hiperespectrales. (C) Análisis de componentes principales (PCA) utilizando la segunda derivada (ventana = 35 puntos) mostrando el PC1 y (D) PC2. (E) Biplot del PC1 versus el PC2. (F) Gráfico de carga de las longitudes de onda más relevantes para PC1 y PC2, destacando las bandas espectrales clave.

Los espectros medios no procesados de los genotipos 4, 5, 6 y 7 (Figura 3.2B) mostraron patrones de absorción característicos en el rango visible (400-700 nm) e infrarrojo cercano (700-1000 nm). En la región azul-verde (400-500 nm), los espectros se podrían asociar pigmentos fotosintéticos, como clorofilas y carotenoides (Huang *et al.*, 2015), los cuales tienen una señal baja probablemente asociada al cultivo de las masas celulares en oscuridad. Se observaron picos de absorción en 442 nm y 471 nm, lo que sugiere variabilidad intergenotípica posiblemente asociada a diferencias en la concentración del rango del rojo (600-700 nm), y se identificaron picos en 552 nm y 618 nm, indicando posibles diferencias en la actividad fotosintética entre los genotipos.

En el rango del infrarrojo cercano (700-1000 nm), los picos observados en 721 nm, 813 nm y 901 nm mostraron superposiciones entre los grupos, lo que sugiere similitudes en la composición hídrica y la estructura molecular de las masas embriogénicas (Huang *et al.*, 2024). Estos hallazgos resaltan la importancia de aplicar transformaciones matemáticas, como la segunda derivada, para realzar diferencias críticas no evidentes en los espectros brutos, una práctica ampliamente recomendada en el análisis de datos hiperespectrales complejos (Li *et al.*, 2022).

### **3.3. Análisis multivariado mediante componentes principales**

El análisis de componentes principales (PCA) aplicado a las imágenes hiperespectrales de masas embriogénicas de los genotipos 1, 2 y 3, preprocesadas con segunda derivada (ventana de 25 puntos), reveló dos componentes principales (PC1 y PC2) que, en conjunto, explican el 63,2% de la varianza total. El PC1 representa la principal fuente de variabilidad en la composición bioquímica o fisiológica de los callos (Figura 3.1C). La distribución espacial de colores dentro de cada genotipo mostró gradientes cromáticos continuos, con tonos que varían desde el azul (valores bajos) hasta el rojo (valores altos). Cabe destacar que los genotipos no presentaron diferencias significativas en PC1, lo que sugiere una composición bioquímica o fisiológica similar en este componente principal. En cambio el PC2 mostró una mayor capacidad discriminativa entre los genotipos, evidenciando discontinuidades en la distribución espacial de colores dentro de cada grupo (Figura 3.1D). Esta variabilidad podría estar relacionada con diferencias en la estructura celular, el contenido de agua o la presencia de compuestos secundarios, aspectos no capturados por PC1 (Huang *et al.*, 2024). La combinación de ambos componentes permite una comprensión más completa de la variabilidad entre diferentes masas embriogénicas, destacando su utilidad en el diseño de sistemas de clasificación automática.

El análisis de componentes principales (PCA) aplicado a las imágenes hiperespectrales de las masas embriogénicas de los genotipos 4, 5, 6 y 7, preprocesadas con segunda derivada (ventana de 35 puntos), reveló dos componentes principales (PC1 y PC2) que, en conjunto, explican el 77,17% de la varianza total. El PC1 mostró una separación robusta y consistente entre los cuatro genotipos analizados (Figura 3.2C). Esta variabilidad se reflejó en la distribución espacial de colores al comparar los genotipos, donde se observaron gradientes cromáticos continuos, con tonos que varían desde el azul (valores bajos) hasta el rojo (valores altos). El PC2 demostró similar capacidad discriminativa, separando claramente a los genotipos 4 y 6 de los genotipos 5 y 7, y distinto a las diferencias leves observadas en RGB.

### **3.4. Identificación de bandas espectrales relevantes**

Para profundizar en la diferenciación entre los genotipos 1, 2 y 3, se generó un gráfico de carga de PC1 y PC2 (Figura 3.1E), que permitió identificar las bandas espectrales más importantes. Los picos y valles en las longitudes de onda 428, 468, 478, 601, 680, 931 y 962 nm sugieren diferencias químicas entre los grupos de genotipos. Las longitudes de onda mencionadas corresponden a regiones

específicas del espectro electromagnético, relacionadas con la interacción de las sustancias químicas con la luz. Los valores de 478, 680 y 931 nm abarcan desde el rango visible hasta el infrarrojo cercano. El pico en 478 nm (azul) está asociado con pigmentos específicos, mientras que el de 680 nm (rojo) está relacionado con la clorofila (Huang *et al.*, 2015). Por otro lado, el de 931 nm (infrarrojo cercano) refleja la presencia de agua (Huang *et al.*, 2024).

Por su parte, los valles en 428, 468, 601 y 962 nm presentan características complementarias. Los valores en 428 y 468 nm (región azul) están asociados con la absorción de luz por carotenoides, pigmentos clave en la fotosíntesis y la fotoprotección. El valle en 601 nm (rojo-anaranjado) también está vinculado a los carotenoides y ciertas formas de clorofila (Huang *et al.*, 2015). Finalmente, el valle en 962 nm (infrarrojo cercano) se relaciona con las vibraciones de los enlaces de hidroxilo en moléculas de agua (Huang *et al.*, 2024). Estas diferencias en los picos y valles del espectro reflejan la diversidad química de los genotipos, lo que permite su análisis mediante imágenes hiperespectrales.

Para profundizar en la diferenciación entre los genotipos 4, 5, 6 y 7, se analizaron las cargas espectrales de los componentes principales PC1 y PC2. Los valores más importantes de PC1 (Figura 3.2E), hacia los extremos positivos incluyeron 432, 509, 647 y 923 nm, mientras que en los valles destacaron 470, 554, 602,

748, 851 y 974 nm. En el caso de PC2, los extremos positivos correspondieron a 472, 553, 749 y 873 nm, y los negativos a 448, 513, 698 y 925 nm. Estas longitudes de onda abarcan regiones específicas del espectro electromagnético, relacionadas con la interacción de las sustancias químicas con la luz.

En PC1, los picos en 432 nm (azul), 509 nm (verdeazulado) y 647 nm (rojo) están asociados con la absorción de luz por carotenoides y clorofila, componentes esenciales en la fotosíntesis. El pico a 432 nm corresponde a los carotenoides, responsables de la protección contra el estrés lumínico, mientras que los valores a 509 y 647 nm reflejan la actividad de clorofila *a* y *b* (Huang *et al.*, 2015). Por otro lado, el pico en 923 nm (infrarrojo cercano) indica la presencia de agua en los tejidos vegetales, un factor crítico en la fisiología celular (Huang *et al.*, 2024). Los extremos negativos de PC1 (470, 554, 602 nm) sugieren interacciones con carotenoides y flavonoides, mientras que los picos en 748, 851 y 974 nm (infrarrojo cercano) apuntan a la absorción de agua y lípidos, evidenciando variaciones en la composición orgánica de las masas embriogénicas.

En PC2, los picos en 472 nm (azul), 553 nm (verde) y 749 nm (infrarrojo cercano) están relacionados con pigmentos fotosintéticos como carotenoides y clorofila, así como con la presencia de agua en los tejidos. El pico en 873 nm (infrarrojo cercano) refuerza la identificación de compuestos estructurales y agua. Por su

parte, los extremos negativos en 448 nm (azul) y 513 nm (verde) se vinculan a flavonoides y compuestos fenólicos, mientras que el pico en 698 nm (rojo) está asociado con la clorofila en estados foto oxidativos. Finalmente, 925 nm (infrarrojo cercano) refleja variaciones en el contenido de agua y la organización molecular de los tejidos (Huang *et al.*, 2024).

### **3.5. Generalización del diagrama de dispersión simple de dos variables (Biplot)**

En el biplot de PC1 y PC2 de genotipos 1, 2 y 3 (Figura 3.1F), se observaron dos grupos claramente diferenciados. Al marcar los píxeles clave (triángulo rojo), se evidenció que los genotipos 1 y 2 forman un grupo con espectros similares, mientras que el genotipo 3 presenta un perfil espectral distinto reafirmando una diferenciación entre los genotipos 1, 2 versus el genotipo 3.

En el biplot de PC1 y PC2 (Figura 3.2F), no se observó una agrupación clara y definida entre los genotipos 4, 5, 6 y 7. Sin embargo, en términos generales, se identificaron dos tendencias de agrupación: una que cruza por la parte superior del gráfico y otra por la inferior, ambas compartiendo componentes de PC1 y PC2. Esta distribución sugiere que los genotipos podrían estar expuestos de

distinta forma o con diferente intensidad a los factores representados por estos componentes principales, lo que permite una separación parcial en dos grupos. El primer grupo está conformado por los genotipos 4 y 6, mientras que el segundo grupo incluye a los genotipos 5 y 7. Esta separación, aunque no es completamente definida, reafirma las diferencias espectrales y bioquímicas previamente identificadas en los análisis de componentes principales. La tendencia observada en el biplot respalda la idea de que los genotipos 4 y 6 comparten características espectrales y distintas a las de los genotipos 5 y 7, lo que podría estar relacionado con variaciones en su composición química, estructura celular o estado de desarrollo.

### **3.6. Resultados de análisis PLS-DA**

En el análisis PLS-DA aplicado a los genotipos 1, 2 y 3 (Figura 3.3A), se definieron tres grupos principales de píxeles, correspondientes a cada genotipo, cuyos datos hiperespectrales fueron utilizados para entrenar y validar el modelo. Las métricas de desempeño del modelo PLS-DA demostraron una alta eficacia, con una especificidad (Spec) de 1.00, una sensibilidad (Sens) de 0.99, una precisión (Prec) de 0.99, una tasa de error de 0.01 y una exactitud general (Accuracy) de 1.00. Además, se registró un bajo porcentaje de píxeles no

asignados (0.02), lo que confirma la robustez y fiabilidad del modelo en la clasificación de los genotipos analizados.

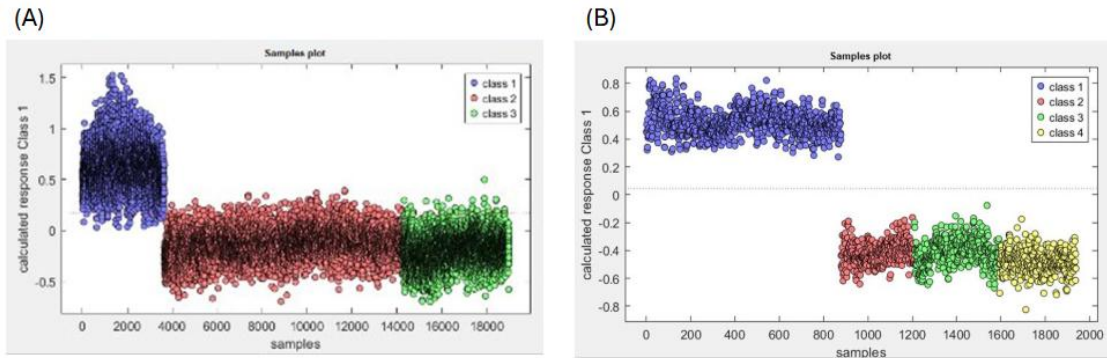


Figura 3. 3: Resultados de clasificación PLS-DA para las muestras de masas embriogénicas de *Pinus radiata* en etapa de multiplicación, diferenciadas por el método de propagación. (A) Clasificación PLS-DA de los genotipos 1, 2 y 3, multiplicados mediante el método de disgregación. Se establecen tres clases: Class 1 (genotipo 1), Class 2 (genotipo 2) y Class 3 (genotipo 3). (B) Clasificación PLS-DA de los genotipos 4, 5, 6 y 7, multiplicados mediante el método de aglomerado. Se definen cuatro clases: Class 1 (genotipo 4), Class 2 (genotipo 5), Class 3 (genotipo 6) y Class 4 (genotipo 7).

De manera similar, el análisis PLS-DA se implementó para evaluar la capacidad de discriminación entre los genotipos 4, 5, 6 y 7 (Figura 3.3B), identificando cada genotipo mediante zonas específicas dentro de las masas embriogénicas y clasificando los píxeles correspondientes mediante el modelo PLS-DA. Las métricas de desempeño del modelo PLS-DA en este caso mostraron una eficacia aún óptima, con una especificidad (Spec) de 1.00, una sensibilidad (Sens) de 1.00, una precisión (Prec) de 1.00, una tasa de error de 0.00 y una exactitud general (Accuracy) de 1.00, lo que refuerza la alta precisión y fiabilidad del

modelo en la discriminación de ambos grupos de genotipos analizados. Este enfoque integrado permite visualizar la consistencia y efectividad del modelo PLS-DA en la clasificación y discriminación de los genotipos, tanto en el primer grupo (1, 2 y 3) como en el segundo (4, 5, 6 y 7), destacando su aplicabilidad en estudios de diferenciación genotípica.

### **3.7 Análisis incluyendo todos los genotipos**

#### **3.7.1. Análisis Multivariado Mediante Componentes Principales**

El análisis de componentes principales (PCA) aplicado a las imágenes hiperespectrales de los callos embriogénicos de *Pinus radiata* preprocesados con segunda derivada reveló dos componentes principales (PC1 y PC2) que, en conjunto, explican el 82,90% de la varianza total. El PC1 representa la principal fuente de variabilidad en la composición bioquímica de las masas embriogénicas (Figura 3.4A). En este análisis, se observaron diferencias claras entre los genotipos 1, 2 y 3 frente a los genotipos 4, 5, 6 y 7, lo cual sugiere una discriminación significativa en términos de su firma espectral. Estas diferencias podrían estar relacionadas con factores como el fondo de las imágenes (negro o blanco) o la capacidad madurativa de los genotipos.

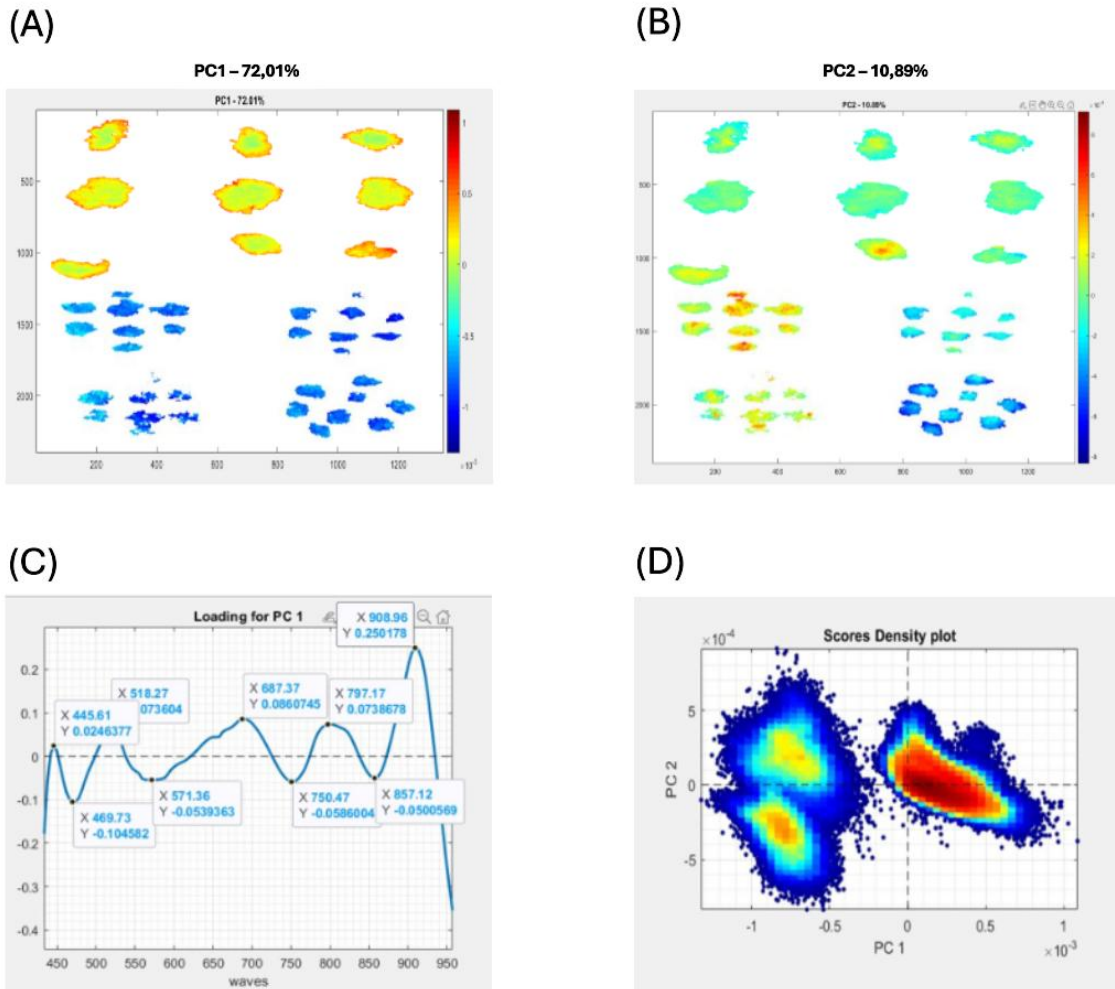


Figura 3. 4: Análisis de componentes principales (PCA) aplicado a masas embriogénicas de *Pinus radiata*. (A) PC1 de los genotipos 1, 2, 3, 4, 5, 6 y 7, calculado con segunda derivada (ventana de 35 puntos), explicando el 72,01% de la varianza total. Los genotipos 1, 2 y 3 corresponden a la etapa de maduración, mientras que los genotipos 4, 5, 6 y 7 se encuentran en la etapa de multiplicación. (B) PC2 de los mismos genotipos, con segunda derivada (ventana de 35 puntos), explicando el 10,89% de la varianza total, manteniendo la diferenciación entre etapas. (C) Gráfico de cargas de PC1, resaltando las bandas espectrales más relevantes en la variabilidad de las muestras. (D) Biplot de PC1 genotipos, ya que los genotipos 1, 2 y 3 generaron embriones, mientras que los genotipos 4, 5, 6 y 7 no lo hicieron versus PC2, visualizando los patrones de distribución de los genotipos en función de sus espectros, permitiendo identificar tendencias y agrupamientos.

Por otro lado, PC2, que explica el 10,89% de la varianza (Figura 3.4B), mostró una distribución de colores que sugiere distintas agrupaciones entre los genotipos. Sin embargo, estas agrupaciones no presentan una identificación clara, lo que indica que PC2 tiene una capacidad discriminativa limitada.

### **3.7.2. Identificación de bandas espectrales relevantes**

Para profundizar en la diferenciación entre los genotipos 1, 2, 3, 4, 5, 6 y 7, se generó un gráfico de carga de PC1 (Figura 3.4C), que permitió identificar las bandas espectrales más importantes. Los picos y valles en las longitudes de onda 446, 518, 687, 797, 909, 470, 572, 750 y 857 nm sugieren diferencias químicas significativas entre los grupos de genotipos. Estas longitudes de onda corresponden a regiones específicas del espectro electromagnético, relacionadas con la interacción de las sustancias químicas con la luz.

Los picos en 446 nm (azul), 518 nm (verde) y 687 nm (rojo) están asociados con la absorción de luz por pigmentos fotosintéticos. El pico a 446 nm se relaciona con carotenoides, responsables de la fotoprotección y la captación de luz en el rango azul (Huang *et al.*, 2015). Por su parte, el pico en 518 nm está vinculado a la clorofila *b*, mientras que el pico en 687 nm refleja la presencia de clorofila *a*, ambos esenciales en la fotosíntesis (Huang *et al.*, 2015). En el rango del infrarrojo

cercano, los picos en 797 nm y 909 nm están asociados con la absorción de agua y la estructura molecular de los tejidos vegetales, respectivamente (Huang *et al.*, 2024). Por otro lado, los valles en 470 nm (azul), 572 nm (verde-amarillo), 750 nm (infrarrojo cercano) y 857 nm (infrarrojo cercano) presentan características complementarias. El valle en 470 nm está relacionado con la absorción de carotenoides y ciertos flavonoides, mientras que el valle en 572 nm sugiere interacciones con pigmentos fotosintéticos en estado de transición. Los valles en 750 nm y 857 nm reflejan variaciones en el contenido de agua y la organización molecular de los tejidos, lo que podría indicar diferencias en la hidratación o la composición estructural de las masas embriogénicas (Huang *et al.*, 2024).

### **3.7.3. Generalización del diagrama de dispersión simple de dos variables (Biplot)**

El biplot generado a partir del análisis de componentes principales (PCA), que incluyó todos los genotipos (1, 2, 3, 4, 5, 6 y 7) (Figura 3.4D), mostró una clara diferenciación en tres grupos principales. En la región superior izquierda del biplot se ubicaron los genotipos 4 y 5, formando el primer grupo. En la región inferior izquierda se agruparon los genotipos 6 y 7, constituyendo el segundo grupo. Por último, en la región derecha del biplot se situaron los genotipos 1, 2 y 3, conformando el tercer grupo.

En el PC1, que explica la mayor parte de la varianza, los genotipos 4, 5, 6 y 7 mostraron una agrupación hacia la derecha, separándose claramente de los genotipos 1, 2 y 3. Esta separación sugiere diferencias significativas en las características espectrales entre estos grupos, lo que podría estar relacionado con variaciones en su composición bioquímica o fisiológica. Por otro lado, el PC2, que explica una menor proporción de la varianza, pero aún significativa, permitió distinguir subgrupos dentro de los genotipos 4, 5, 6 y 7. Los genotipos 4 y 5 formaron un grupo diferenciado, mientras que los genotipos 6 y 7 se agruparon de manera distinta. Esta separación en el PC2 no concuerda con los resultados previos para los genotipos 4, 5, 6 y 7.

## **3.8 Resultados del análisis FTIR**

### **3.8.1 Perfiles espectrales brutos**

Los perfiles espectrales brutos (Figura 3.5A) aplicado a los datos de FTIR no mostró patrones claros de agrupación. Esto fue consistente tanto al analizar los datos según genotipo como al diferenciarlos por su capacidad generativa (masas productoras de embriones versus no productoras). En las gráficas obtenidas, se observa una dispersión sin una separación evidente entre los grupos que generan embriones (rojo) con los que no generan embriones (verde), lo que indica que las variaciones detectadas por FTIR no son suficientes para discriminar estas características en las condiciones evaluadas.

### **3.8.2 Identificación de bandas espectrales relevantes**

Para profundizar en la diferenciación entre los genotipos 1, 2, 3, 4, 5, 6 y 7, se generó un gráfico de carga de PC1 (Figura 3.5C), que permitió identificar las bandas espectrales más importantes. Los picos y valles en las longitudes de onda 446, 518, 687, 797, 909, 470, 572, 750 y 857 nm sugieren diferencias químicas significativas entre los grupos de genotipos. No obstante, no se muestra una diferenciación clara entre masas embriogénicas y no embriogénicas.

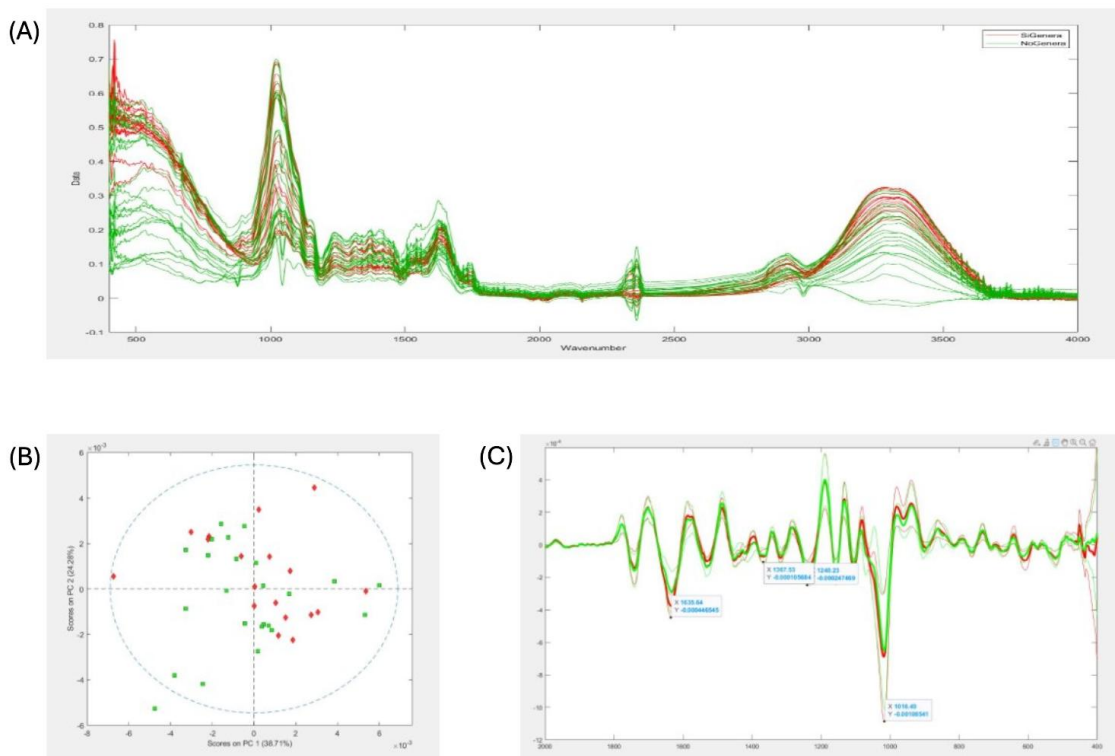


Figura 3.5. Análisis espectroscópico de extractos de metanol obtenidos de masas embriogénicas de *Pinus radiata*. (A) Gráfico de absorbancia bruta de las muestras, donde las masas embriogénicas capaces de generar embriones se representan en rojo, y aquellas que no generan embriones en verde. (B) Biplot de PC1 (38,71% de varianza explicada) vs. PC2 (24,28% de varianza explicada), mostrando patrones de diferenciación entre las muestras que generan embriones (rojo) y las que no (verde). (C) Gráfico de absorbancia con segunda derivada (ventana de 35 puntos), resaltando las bandas espectrales más relevantes en los extractos de metanol de masas embriogénicas de *Pinus radiata*.

### 3.9 Contabilización de embriones

Cinco meses después de la obtención de datos espectroscópicos en fase de multiplicación, se realizó la contabilización de embriones al final de la fase de maduración. Después de este proceso, se observaron, masas que generan embriones (genotipos 1, 2 y 3) y masas que no generan embriones (genotipos 4, 5, 6 y 7). En rendimiento de embriones normales para los genotipos 1, 2 y 3 no presentó diferencias significativas (Figura 3.6A). En otras palabras, la variabilidad observada en este caso podría deberse al azar o a factores no controlados durante el experimento, en lugar de reflejar una diferencia real en la capacidad embriogénica de los genotipos. Por otro lado, los genotipos 1, 2 y 3 si se diferenciaron en la producción de embriones anormales (Figura 3.6B).

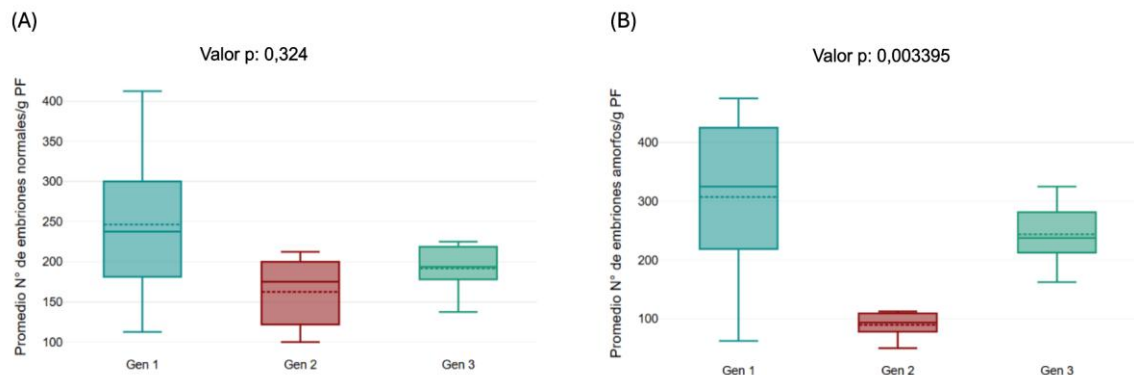


Figura 3.6: Diagramas de caja que muestran el promedio del número de embriones por gramo de peso fresco en masas embriogénicas de *Pinus radiata* en etapa de maduración para los genotipos 1, 2 y 3. (A) Número de embriones normales/g de peso fresco. (B) Número de embriones amorfos por gramo de peso fresco.

## IV. DISCUSIÓN

Los resultados de este estudio rechazan la hipótesis nula ( $H_0$ ) y validan la hipótesis alternativa ( $H_1$ ), demostrando que las técnicas espectroscópicas basadas en imágenes, en particular el análisis hiperespectral, permiten predecir la capacidad madurativa de los cultivos embriogénicos somáticos de *Pinus radiata* durante la fase de multiplicación. La combinación de análisis multivariado (PCA y PLS-DA) con la identificación de bandas espectrales clave demostró ser una estrategia efectiva para discriminar genotipos con potencial contrastante, representando un avance significativo en la optimización de programas de propagación clonal en especies forestales.

El análisis de componentes principales (PCA) reveló una clara separación entre genotipos productores (1-3) y no productores (4-7), explicando hasta el 82,9% de la varianza total (PC1 + PC2). Esta diferenciación se atribuye a variaciones en la composición bioquímica de las masas embriogénicas, particularmente en regiones espectrales asociadas a carbohidratos (931 nm) y clorofila (680 nm). Estudios previos en *P. radiata* han destacado la importancia de estas moléculas durante la maduración embrionaria, donde la acumulación de sacarosa y la disponibilidad de nitrógeno son factores críticos para la formación de embriones

(García-Mendiguren et al., 2016; Lineros et al., 2018). Además, la identificación de una banda en el rango visible (478 nm), vinculada a carotenoides, coincide con hallazgos en *Solanum tuberosum*, donde estos pigmentos cumplen un papel antioxidante durante el desarrollo de tejidos embriogénicos (Su et al., 2021).

La capacidad del PCA para agrupar genotipos según su firma espectral resalta su utilidad en la caracterización temprana de material vegetal, como se ha aplicado en *Coffea arabica* para la selección de líneas embriogénicas con alta productividad (Nic-Can et al., 2016). Sin embargo, a diferencia de estudios en cultivos agrícolas, este trabajo demuestra por primera vez en *P. radiata* que la variabilidad espectral durante la fase de multiplicación puede ser un predictor confiable de la capacidad madurativa, incluso antes del inicio de la maduración.

El modelo PLS-DA mostró un desempeño excepcional, con una exactitud del 100%, superando al PCA en la clasificación de genotipos. Esta ventaja se debe a su naturaleza supervisada, que permite correlacionar directamente las variables espectrales con la capacidad madurativa, filtrando el ruido y resaltando bandas críticas. Por ejemplo, las cargas espectrales en 931 nm y 680 nm fueron identificadas como las más influyentes, en concordancia con estudios en *Zea mays*, donde el PLS-DA permitió predecir la concentración de nitrógeno foliar con alta precisión (Miao et al., 2023). La efectividad del PLS-DA también ha sido reportada en *Daucus carota* para el monitoreo de la calidad de tejidos en biorreactores automatizados (Polder et al., 2024). En este estudio, su aplicación

no solo redujo el tiempo de análisis, sino que también evitó el uso de métodos destructivos, un aspecto clave para preservar la viabilidad de las masas embriogénicas durante la multiplicación (Akhtar, 2012).

Las bandas espectrales identificadas (931 nm, 680 nm y 478 nm) reflejan procesos metabólicos esenciales para la maduración embrionaria:

- **931 nm:** Asociada a enlaces O-H (agua) y C-O (carbohidratos), esta banda sugiere que la acumulación de sacarosa en el medio de cultivo ( $30 \text{ g L}^{-1}$ ) es un factor determinante para la formación de embriones. Estudios en *Picea abies* respaldan esta correlación, donde altas concentraciones de sacarosa favorecieron la conversión de embriones somáticos en plántulas viables (Businge et al., 2013).
- **680 nm:** Relacionada con clorofila a, esta banda actúa como un indicador de la disponibilidad de nitrógeno, un nutriente esencial para la síntesis de proteínas y ácidos nucleicos durante la diferenciación celular (Miao et al., 2023).
- **478 nm:** Asociada a carotenoides, estos compuestos podrían desempeñar un papel clave en la mitigación del estrés oxidativo en tejidos embriogénicos, un desafío común en cultivos *in vitro* de coníferas (Junmeng et al., 2022).

A pesar de los resultados prometedores, este estudio presentó algunas limitaciones. Uno de los más importantes es que los métodos de multiplicación utilizados no fueron los mismos para todos los genotipos. Además los fondos usados para la captura de las imágenes también fueron distintos. Se formó un grupo de genotipos cultivados en multiplicación por método de disgregación y adquisición de imágenes con fondo negro (genotipos 1, 2 y 3). El otro grupo de genotipos fueron cultivados en multiplicación por el método de aglomerados y adquisición de imágenes con fondo blanco. Se requieren experimentos y análisis adicionales para determinar cuánto pesan estos factores en los resultados finales. Está previamente reportado que el uso de fondos blanco y negro en la adquisición de imágenes hiperespectrales, podría haber introducido sesgos en la reflectancia, un problema previamente reportado en estudios de *Zea mays* (Junmeng et al., 2022).

La espectroscopia FT-IR no logró discriminar con eficacia los grupos, posiblemente debido a la heterogeneidad de las muestras y al proceso de molido más extracción de las muestras en metanol, un problema previamente documentado en análisis espectrales de tejidos vegetales no homogéneos (Cozzolino, 2004).

En futuros estudios la integración de algoritmos de *machine learning* podría mejorar la selección de bandas espectrales, como ha sido aplicado en *Triticum aestivum* para la predicción del rendimiento de cultivos (Miao et al., 2023). La combinación de técnicas espectroscópicas con herramientas avanzadas de análisis de datos podría fortalecer la implementación de estrategias no destructivas en la biotecnología forestal, mejorando la eficiencia en la selección de genotipos con alto potencial de propagación.

## V. CONCLUSIONES

Este trabajo confirma que las imágenes hiperespectrales son una herramienta poderosa para la determinación temprana de respuestas embriogénicas en *Pinus radiata*, lo que podría ayudar a reducir costos y tiempo en programas de propagación clonal. La identificación de bandas espectrales clave (931 nm, 680 nm, 478 nm) no solo valida la hipótesis alternativa, sino que también establece un precedente para el uso de espectroscopía en biotecnología forestal. Futuras investigaciones deberán enfocarse en superar limitaciones técnicas y escalar este enfoque para su implementación industrial, contribuyendo a la sostenibilidad de uno de los sectores económicos más relevantes de Chile.

## GLOSARIO

**Ácido Abscísico (ABA):** Fitohormona clave en la regulación de la maduración de embriones somáticos, utilizada en medios de cultivo para inducir estrés abiótico.

**Accuracy (Exactitud):** Métrica estadística que mide la proporción de predicciones correctas en un modelo (ej.: PLS-DA).

**Biplot:** Gráfico que combina puntuaciones y cargas de componentes principales (PCA) para visualizar relaciones entre variables y muestras.

**BAP (6-bencilaminopurina):** Citocinina sintética utilizada en medios de cultivo para estimular la proliferación celular.

**Capacidad Madurativa:** Habilidad de un cultivo embriogénico para generar embriones somáticos viables durante la fase de maduración.

**Chemometrics (Quimiometría):** Uso de métodos estadísticos y matemáticos para analizar datos químicos (ej.: análisis espectral).

**Derivada espectral:** Preprocesamiento matemático aplicado a espectros para reducir ruido y resaltar bandas relevantes (ej.: segunda derivada con ventana de 25 puntos).

**Embriogénesis Somática:** Proceso por el cual células somáticas se diferencian en embriones, permitiendo la propagación clonal de plantas.

**Espectroscopía FT-IR:** Técnica que mide la absorción de luz infrarroja para identificar grupos funcionales en moléculas (ej.: enlaces C=O, O-H).

**Fitagel®:** Agente gelificante utilizado en medios de cultivo para proporcionar soporte físico a tejidos vegetales.

**Genotipo:** Variante genética específica de un organismo (en tu estudio, siete genotipos de *Pinus radiata*).

**Hiperespectral (Imagen):** Técnica que captura información espectral en cientos de bandas de longitud de onda (ej: 400-1000 nm) para analizar composición química en tejidos.

**Infrarrojo Cercano (NIR):** Región del espectro electromagnético (700-2500 nm) utilizada para detectar compuestos orgánicos (ej: agua, carbohidratos).

**Litvay Modificado (mLM):** Medio de cultivo semisólido utilizado para la proliferación y maduración de masas embriogénicas.

**Metabolitos Secundarios:** Compuestos orgánicos no esenciales para el crecimiento, pero importantes para la defensa o adaptación (ej: flavonoides, alcaloides).

**PCA (Análisis de Componentes Principales):** Técnica estadística para reducir la dimensionalidad de datos, identificando patrones espectrales dominantes.

**PLS-DA (Partial Least Squares Discriminant Analysis):** Modelo de clasificación que relaciona variables espectrales con categorías (ej: genotipos).

**RMSE (Raíz del Error Cuadrático Medio):** Métrica que cuantifica la diferencia entre valores predichos y observados en un modelo.

**Sensibilidad (Sens):** Capacidad de un modelo para identificar correctamente positivos (ej: genotipos productores de embriones).

**Specificity (Especificidad):** Capacidad de un modelo para identificar correctamente negativos (ej: genotipos no productores).

**Validación Cruzada:** Método para evaluar la generalización de un modelo, dividiendo datos en subconjuntos de entrenamiento y prueba.

**VIS (Espectro Visible):** Región del espectro electromagnético (400-700 nm) asociada a pigmentos como clorofila y carotenoides.

## VI. BIBLIOGRAFÍA

Akhtar, N. (2012). *Somatic Embryogenesis for Efficient Micropropagation of Guava (Psidium guajava L.)*. In: Lambardi, M., Ozudogru, E., Jain, S. (eds) *Protocols for Micropropagation of Selected Economically-Important Horticultural Plants*. Methods in Molecular Biology, vol 994. Humana Press, Totowa, NJ. [https://doi.org/10.1007/978-1-62703-074-8\\_12](https://doi.org/10.1007/978-1-62703-074-8_12)

Aquea, F., Poupin, M.J., Matus, J.T. et al. (2008). *Synthetic seed production from somatic embryos of Pinus radiata*. *Biotechnology Letters*, 30, 1847–1852. <https://doi.org/10.1007/s10529-008-9754-x>

Benelli, A., Cevoli, C., Ragni, L., Fabbri, A. (2021). *In-field and non-destructive monitoring of grapes maturity by hyperspectral imaging*. *Biosystems Engineering*, 207, 59-67. <https://doi.org/10.1016/j.biosystemseng.2021.04.006>

Businge, E., Brackmann, K., Moritz, T., Egertsdotter, U. (2012). *Metabolite profiling reveals clear metabolic changes during somatic embryo development of Norway spruce (Picea abies)*. *Tree Physiology*, 32(2), 232–244. <https://doi.org/10.1093/treephys/tpr142>

Cozzolino, D., Moron, A. (2004). *Exploring the use of near infrared reflectance spectroscopy (NIRS) to predict trace minerals in legumes*. *Animal Feed Science and Technology*, 111, 161-173. <https://doi.org/10.1016/j.anifeedsci.2003.08.001>

García-Mendiguren, O., Montalbán, I.A., Goicoa, T. et al. (2016). *Environmental conditions at the initial stages of Pinus radiata somatic embryogenesis affect the production of somatic embryos*. *Trees*, 30, 949–958. <https://doi.org/10.1007/s00468-015-1336-7>

Huang, B., Li, S., Long, T., Bai, S., Zhao, J., Xu, H., Long, Y. (2024). *Research on predicting photosynthetic pigments in tomato seedling leaves based on near-infrared hyperspectral imaging and machine learning*. *Microchemical Journal*, 204, 111076. <https://doi.org/10.1016/j.microc.2024.111076>

Huang, J., Wei, C., Zhang, Y., Blackburn, G.A., Wang, X., et al. (2015). *Meta-Analysis of the Detection of Plant Pigment Concentrations Using*

*Hyperspectral Remotely Sensed Data*. PLOS ONE, 10(9), e0137029. <https://doi.org/10.1371/journal.pone.0137029>

Li, J., Ren, J., Cui, R., Yu, K., Zhao, Y. (2022). *Optical imaging spectroscopy coupled with machine learning for detecting heavy metal of plants: A review*. *Frontiers in Plant Science*. <https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2022.1007991>

Linerós, Y., Balocchi, C., Muñoz, X. et al. (2018). *Cryopreservation of Pinus radiata embryogenic tissue: Effects of cryoprotective pretreatments on maturation ability*. *Plant Cell, Tissue and Organ Culture*, 135, 357–366. <https://doi.org/10.1007/s11240-018-1469-7>

Linerós, Y., Rojas, M., Hernández, M., Ríos, D., Muñoz, X., Hasbun, R. (2023). *Effects of Long-Term Subculture on Maturation Ability and Plant Conversion in Pinus radiata: Using FT-IR Spectroscopy to Determine Biomarkers of Embryogenic Tissue Aging*. *Forests*, 14, 1446. <https://doi.org/10.3390/f1407144>

Miao, X., Miao, Y., Liu, Y., Tao, S., Zheng, H., Wang, J., Wang, W., Tang, Q. (2023). *Measurement of nitrogen content in rice plant using near infrared spectroscopy combined with different PLS algorithms*. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 284, 121733. <https://doi.org/10.1016/j.saa.2022.121733>

Nic-Can, G.I., Loyola-Vargas, V.M. (2016). *The Role of the Auxins During Somatic Embryogenesis*. In: Loyola-Vargas, V., Ochoa-Alejo, N. (eds) *Somatic Embryogenesis: Fundamental Aspects and Applications*. Springer, Cham. [https://doi.org/10.1007/978-3-319-33705-0\\_10](https://doi.org/10.1007/978-3-319-33705-0_10)

Poblete Hernández, P., Gysling Caselli, J., Álvarez González, V., et al. (2023). *Anuario forestal 2023*. INFOR. <https://doi.org/10.52904/20.500.12220/32652>

Polder, J., Dieleman, A., Hageraats, S., Meinen, E. (2024). *Imaging spectroscopy for monitoring the crop status of tomato plants*. *Computers and Electronics in Agriculture*, 216. <https://doi.org/10.1016/j.compag.2023.108504>

Prananto, J.A., Minasny, B., Weaver, T. (2020). *Near infrared (NIR) spectroscopy as a rapid and cost-effective method for nutrient analysis of plant leaf tissues*. In: Sparks, D.L. (Ed.), *Advances in Agronomy*, 164. Academic Press. <https://doi.org/10.1016/bs.agron.2020.06.001>

Su, W-H., Xue, H. (2021). *Imaging Spectroscopy and Machine Learning for Intelligent Determination of Potato and Sweet Potato Quality*. *Foods*, 10(9), 2146. <https://doi.org/10.3390/foods10092146>