



UNIVERSIDAD DE CONCEPCIÓN
Dirección De Postgrado
Facultad De Ciencias Forestales
Programa de Magíster en Ciencias Forestales

**AUTENTICACIÓN DE ORIGEN FILOGEOGRÁFICO DE PLÁNTULAS DE *Araucaria*
araucana MEDIANTE LA APLICACIÓN DE TÉCNICAS DE ESPECTROSCOPIA EN
DISTINTOS RANGOS DEL INFRARROJO Y MÉTODOS QUIMIOMÉTRICOS**

Macarena Estefanía Rojas Rioseco
CONCEPCIÓN- CHILE
2021

Profesor Guía: Dr. Rodrigo Hasbún Z.
Depto. de Silvicultura,
Facultad de Ciencias Forestales
Universidad de Concepción

AUTENTICACIÓN DE ORIGEN FILOGEOGRÁFICO DE PLÁNTULAS DE *Araucaria araucana* MEDIANTE LA APLICACIÓN DE TÉCNICAS DE ESPECTROSCOPIA EN DISTINTOS RANGOS DEL INFRARROJO Y MÉTODOS QUIMIOMÉTRICOS

Comisión Evaluadora:

Rodrigo Hasbún Zaror (Profesor guía)
Ingeniero Forestal, Dr.

Rosario Castillo Felices (Profesor co-guía)
Bach. Ciencias Químicas, Dra.

Martha Hernández T. (Comisión evaluación)
Ingeniero Químico, Dra.



Directora de Postgrado:

Dra. Darcy Ríos L.
Profesora de Biología y Química, Dra.

Decano Facultad de Ciencias Forestales:

Dr. Manuel Sánchez O.
Ingeniero Forestal. Dr.

DEDICATORIA

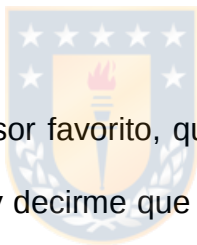
*A cada uno qué me apoyo y soportó, a los
qué siempre estuvieron a mi lado y a los qué
me cuidaron desde el cielo.*

A mí.



AGRADECIMIENTOS

Agradezco enormemente a mi familia por todo el amor y educación que me han dado, quienes día a día me motivan a querer ser la mejor persona posible. Gracias por la paciencia frente a mi mal genio y mis malas caras. A Eduardo González, pilar muy importante en mi vida, quien muchas veces se reía de mi al escucharme decir “mañana tengo que trabajar”, pues tuve que hacerlo. A mis angelitos, les agradezco por escucharme y cumplirme lo que les pido de corazón.



A Rodrigo Hasbún, mi profesor favorito, quien desde el 2015 está ahí para mí, para ayudarme, corregirme y decirme que escribo mal (espero ir mejorando día a día). A mi profesora Rosario Castillo, por su muy buena disposición a brindarme todo su conocimiento y apoyo e integrarme a su tan lindo grupo de laboratorio. Muchas gracias a ambos por creer en mí.

A mis compañeros, Silvia Correa, Inelvis Castro, Luis Felipe Fajardo y Néstor Mora. Sin dudar lo que más valoro de haber tomado este camino en mi vida es haber tenido la posibilidad de conocer personas tan maravillosas, a las que quiero y estimo. A Mario Sanhueza, por haberme enseñado y ayudado con mucha paciencia, sin esperar nada a cambio. A Jorge González por todo su apoyo y por cada miércoles o fin de semana de trabajo.

TABLA DE CONTENIDOS

CAPÍTULOS	PÁGINA
ÍNDICE DE TABLAS	vii
ÍNDICE DE ILUSTRACIONES	viii
RESUMEN	xii
ABSTRACT	xiv
I INTRODUCCIÓN	1
II INVESTIGACIÓN	12
III METODOLOGÍA	14
3.1 Obtención material de estudio	14
3.2 Muestreo de plántulas	15
3.3 Procesamiento de plántulas	16
3.4 Adquisición de espectros	16
3.5 Métodos quimiométricos	19
3.5.1 Pretratamiento de datos	19
3.5.2 Análisis de datos	20
A. Análisis exploratorio	21

B. Métodos Supervisados de Reconocimiento de Patrones (SPRM).....	21
3.6 Validación modelos de clasificación	22
3.7 Interpretación de datos	23
IV RESULTADOS Y DISCUSIÓN	25
1. Rango VIS-NIR (345 - 1037 nm)	25
2. Rango NIR con Transformada de Fourier FT-NIR (12489 - 3996 cm ⁻¹).....	35
3. Rango NIR (1595 - 2396 nm) equipo portátil	43
4. Rango FTIR (3999 - 650 cm ⁻¹).....	51
5. Rango espectral con mayor poder de discriminación	64
6. Discriminación de plántulas de <i>A. araucana</i> según origen filogeográfico ..	71
V CONCLUSIÓN.....	80
VI REFERENCIAS	83
ANEXO 1: TABLA RESUMEN DE LOS RESULTADOS DEL CÁLCULO DE MODELOS K-NN	99
ANEXO 2: TABLA RESUMEN DE LOS RESULTADOS DEL CÁLCULO DE MODELOS SIMCA.....	99

ÍNDICE DE TABLAS

TABLA N°	PÁGINA
En el texto	
Tabla 1. Información sobre región espectral, rango de cobertura (nm y cm^{-1}), resolución y número de scans para cada espectrofotómetro utilizado.....	17
Tabla 2. Preprocesados y transformaciones aplicadas a las cuatro bases espectrales para análisis exploratorios (PCA) y SPRM (K-NN y SIMCA). (*) Preprocesados. SNV: Variación normal estándar. MC: Centrado en la media. AU: Autoescalado. () Indica la ventana de puntos.	20
Tabla 3. Componentes principales calculados para matriz espectral VIS-NIR.....	28
Tabla 4. Componentes principales calculados para matriz espectral FT-NIR.	38
Tabla 5. Componentes principales calculados para matriz espectral NIR.....	45
Tabla 6. Componentes principales calculados para matriz espectral FTIR.	54
Tabla 7. Asignación de bandas que presentan diferencias encontradas en la caracterización espectroscópica y análisis multivariante de las muestras. sym: simétrico, asym: asimétrico, str: estiramiento (stretching), ben: flexión (bending).	63
Tabla 8. Evaluación de modelos supervisados desarrollados en con las 4 matrices espectrales analizadas. Sn: sensibilidad, Sp: Especificidad, CA: precisión de clasificación y ER: tasa de error.	69
Tabla 9. Evaluación de robustez de modelos K-NN y SIMCA en rango VIS-NIR con equipo portátil y FTIR, según %CA.	70

ÍNDICE DE ILUSTRACIONES

TABLA N°	PÁGINA
En el texto	
Figura 1. Esquema de sistema de certificación con la aplicación de herramientas de trazabilidad para el sector forestal. (Modificado de Alía 2005).....	3
Figura 2. Rango de distribución natural de <i>Araucaria araucana</i> en Chile y Argentina (Recuperado de Agesen,1998).....	9
Figura 3. Distribución geográfica de las dos Unidades Evolutivamente Significativas (UES) de <i>A. araucana</i> en Chile, detectadas por análisis genéticos. (Recuperado y modificado de Martín 2014)	11
Figura 4. Localidades muestreadas de semillas de <i>A. araucana</i> en su distribución natural en Chile, periodo de semillación 2018. PNN: Parque Nacional Nahuelbuta; VL: Villa las Araucarias; RNR: Reserva Nacional Ralco; RNN: Reserva Nacional las Nalcas; LON: Lonquimay; RNM: Reserva Nacional Malalcahuello; PNC: Parque Nacional Conguillio; ICA: Icalma; MAR: Marimenuco; CRU: Cruzaco; PNH: Parque Nacional Huerquehue; PNV: Parque Nacional Villarrica. Imagen obtenida de Google Earth.	15
Figura 5. Espectrofotómetros utilizados para análisis de muestras. Equipo Oceans Optics para rango VIS-NIR (345 - 1037 nm), equipo MPA-Bruker para rango NIR con transformada de Fourier (800 - 2500 nm), equipo portátil MicroPhazir para rango NIR (1595 - 2396 nm) y equipo Agilent Cary 630 para rango IR medio con transformada de Fourier (2.500 - 15384 nm).....	18
Figura 6. Proceso de generación pastillas de KBr para análisis en IR medio. A: molienda de KBr y muestra en mortero de ágata bajo luz infrarroja. B: presado a alta presión, con prensa manual. C: pastilla de KBr + muestra en su soporte previo a ser analizada. D: pastilla en su soporte, posicionada en el equipo para ser analizada por transmitancia.	18

- Figura 7.** Espectros VIS-NIR sin extremos ruidosos, en rango 475-922 nm de plántulas de *A. araucana*. (A) % reflectancia y (B) pseudo absorbancia ($\text{Log}(1/R)$). 26
- Figura 8.** Espectros en región VIS-NIR pre-tratados para (A) PCA con 1^{ra} derivada con ventana de 25 puntos y (B) K-NN y SIMCA con 1^{ra} derivada con ventana de 15 puntos. Ambos fueron centrados en la media. 27
- Figura 9.** Loadings de los primeros 3 PCs de PCA en rango VIS-NIR. Espectro verde, azul y rojo corresponden a loadings de PC 1, 2 y 3 respectivamente. Se marca dentro del rectángulo punteado zona espectral poco influyente. 28
- Figura 10.** Gráfico de puntuaciones entre los primeros 3 PCs y sus respectivos gráficos de cargas, obtenidos con matriz VIS-NIR. Cuadros azules corresponden a muestras de la Costa y cuadros naranjos a muestras de los Andes..... 29
- Figura 11.** Gráfico de cargas de PC1 de matriz VIS-NIR obtenido de plántulas de *A. araucana*..... 31
- Figura 12.** Gráfico de Coomans de modelo SIMCA en rango VIS-NIR, umbrales para un 95% de confianza. (A) set de entrenamiento y (B) set de validación. Cuadros azules corresponden a muestras de la Cordillera de la Costa (CS1) y cuadros naranjos a la Cordillera de los Andes (CS2). 32
- Figura 13.** Gráfico de poder de discriminación de variables de modelo SIMCA en rango VIS-NIR para discriminar entre orígenes filogeográficos de plántulas de *A. araucana*. Arbitrariamente solo se señalan longitudes como poder de discriminación mayor a 2.5. 33
- Figura 14.** Espectros brutos de absorbancia en unidades de Kubelka-Munk en rango NIR con Transformada de Fourier, de plántulas de *A. araucana*. 36
- Figura 15.** Espectros FT-NIR de plántulas de *A. araucana* pre-tratados para (A) PCA con MSC, (B) K-NN con 2^{da} derivada de 19 puntos y (C) SIMCA con 2^{da} derivada de 9 puntos. Todos fueron centrados en la media. 37
- Figura 16.** Gráfico de puntuaciones entre los primeros 3 PCs y sus respectivos gráficos de cargas, obtenidos con matriz FT-NIR. Cuadros azules corresponden a CS1 y cuadros naranjos a CS2..... 39
- Figura 17.** Gráfico de cargas de PC2 en rango FT-NIR. Bandas por sobre 0 son influyentes para CS2 y por debajo de 0 para CS1. 40

- Figura 18.** Gráfico de Coomans de modelo SIMCA NIR portátil, umbrales para un 95% de confianza. (A) set de entrenamiento y (B) set de validación. Cuadros azules corresponden a muestras de la Cordillera de la Costa (CS1) y cuadros naranjos a la Cordillera de los Andes (CS2). 42
- Figura 19.** Poder de discriminación de variables en modelo SIMCA en rango FT-NIR. Solo se señalan longitudes con mayor poder de discriminación..... 43
- Figura 20.** Espectros brutos de absorbancia en rango NIR con equipo portátil de plántulas de *A. araucana*. 44
- Figura 21.** Espectros de absorbancia en rango NIR con equipo portátil, de plántulas de *A. araucana* pre-tratados para (A) PCA con MSC, (B) K-NN con 2^{da} derivada de 25 puntos y (C) SIMCA con 2^{da} derivada de 5 puntos. Todos fueron centrados en la media..... 45
- Figura 22.** Gráfico de puntuaciones entre los primeros 3 PCs y sus respectivos gráficos de cargas, obtenidos con matriz NIR portátil. Cuadros azules corresponden a muestras de la costa y cuadros naranjos a muestras de los andes..... 46
- Figura 23.** Gráfico de cargas de PC2 para información adquirida en rango NIR con equipo portátil. Bandas por sobre 0 son influyentes para CS1 y por debajo de 0 para CS2..... 48
- Figura 24.** Gráfico de Coomans de modelo SIMCA NIR portátil, umbrales para un 95% de confianza. (A) set de entrenamiento y (B) set de validación. Cuadros azules corresponden a muestras de la Cordillera de la Costa (CS1) y cuadros naranjos a la Cordillera de los Andes (CS2). 49
- Figura 25.** Poder de discriminación de variables en modelo SIMCA en rango NIR portátil. Solo se señalan longitudes con valor de discriminación mayor a 3. 50
- Figura 26.** Espectros de absorbancia FTIR plántulas de *A. araucana*, se señala en recuadro punteado rango de 2600 a 2000 cm⁻¹ excluido..... 52
- Figura 27.** Espectros en región FTIR pre-tratados para (A) PCA con corrección de línea base, normalizado, (B) K-NN con 2^{da} derivada de 45 puntos y (C) SIMCA con 2^{da} derivada de 25 puntos. Todos fueron centrados en la media. 53

- Figura 28.** Gráfico de puntuaciones entre los primeros 3 PCs y sus respectivos gráficos de cargas. Cuadros azules y naranja corresponden a muestras de la costa y andes respectivamente..... 54
- Figura 29.** Gráfico de puntuaciones entre los primeros 3 PCs y sus respectivos gráficos de cargas. Cuadros azules y naranja corresponden a muestras de la costa y andes respectivamente..... 55
- Figura 30.** Gráfico de cargas de PC2 (FTIR). (A) Gráfico rango 4000 - 650 cm^{-1} . (B) Gráfico rango 4000 - 3380 cm^{-1} . (C) Gráfico rango 2000 - 650 cm^{-1} 57
- Figura 31.** Segunda derivada del promedio de espectros FTIR de cada clase. (A) Rango 4000 - 3400 cm^{-1} y (B) Rango 2000 - 600 cm^{-1} . Prom CS1 y Prom CS2: Promedio espectros clase 1 y clase 2, respectivamente. – SD: Promedio menos desviación estándar. + SD: Promedio más desviación estándar. 59
- Figura 32.** Gráfico de Coomans de modelo SIMCA FTIR, umbrales para un 95% de confianza. (A) set de entrenamiento y (B) set de validación. Cuadros azules corresponden a muestras de la Cordillera de la Costa (CS1) y cuadros naranjos a la Cordillera de los Andes (CS2). 60
- Figura 33.** Poder de discriminación de variables (longitudes de onda, cm^{-1}) en modelo SIMCA en rango FTIR. Solo se señalan longitudes con mayor poder de discriminación. (A) Rango completo, (B) 4000 - 2600 cm^{-1} y (C) 2000 - 657 cm^{-1} 62

RESUMEN

El origen de plántulas es un factor importante para el éxito de programas de restauración, de ser inadecuado puede tener impactos negativos sobre la biodiversidad, por lo que la trazabilidad del material de propagación es clave. Una técnica que permite autenticar el origen filogeográfico es la espectroscopia infrarroja, herramienta rápida, precisa y de bajo costo. Una especie que requiere de programas de restauración y, por ende, de trazabilidad es *Araucaria araucana*, cuyos estudios filogeográficos sugieren diferencias significativas entre poblaciones andinas y costeras. El objetivo de este estudio fue discriminar el origen filogeográfico de plántulas de *A. araucana* mediante técnicas espectroscópicas y métodos quimiométricos. Se trabajó con plántulas de *A. araucana* de ambos orígenes filogeográficos y se colectaron espectros en rango VIS-NIR, FT-NIR, NIR portátil y FTIR. Se aplicaron métodos no supervisados y supervisados (KNN y SIMCA). Los resultados indican que los mejores modelos predictivos fueron SIMCA VIS-NIR y FTIR con un 90% de precisión de clasificación. Se determinó que las muestras costeras presentan mayor contenido de pigmentos carotenoides, Chlb, carbohidratos y proteínas y menor contenido de Chla, lípidos y ácidos grasos respecto a las muestras de origen andino. Se evidencia que las diferencias reportadas a nivel genético entre ambas UES de *A.*

araucana se expresan claramente a nivel químico. Se concluye que los espectros infrarrojos obtenidos de *Araucaria araucana*, tratados con métodos quimiométricos, permiten capturar la señal filogeográfica que separa procedencias costeras de las andinas. Los modelos resultantes podrían utilizarse en futuros programas de restauración de esta especie.



ABSTRACT

The origin of seedlings is an important factor for the success of restoration programs, if it is inappropriate it can have negative impacts on biodiversity, so the traceability of the propagation material is key. One technique that allows authenticating the phylogeographic origin is infrared spectroscopy, a fast, precise and inexpensive tool. One species that requires restoration programs and, therefore, traceability is *Araucaria araucana*, whose phylogeographic studies suggest significant differences between populations and coastal areas. The objective of this study was to discriminate the phylogeographic origin of *A. araucana* seedlings using spectroscopic techniques and chemometric methods. We worked with *A. araucana* seedlings of both phylogeographic origins and spectra were collected in the VIS-NIR, FT-NIR, portable NIR and FTIR range. Unsupervised and supervised methods (KNN and SIMCA) were applied. The results indicate that the best predictive models were SIMCA VIS-NIR and FTIR with 90% classification accuracy. It was determined that the coastal samples have a higher content of carotenoid pigments, Chlb, carbohydrates and proteins and a lower content of Chla, lipids and fatty acids compared to samples of Andean origin. It is evident that the differences reported at the genetic level between both

UES of *A. araucana* are clearly expressed at the chemical level. It is concluded that the infrared spectra obtained from *Araucaria araucana*, treated with chemometric methods, allow to capture the phylogeographic signal that separates coastal and Andean provenances. The resulting models could be possible in future restoration programs of this species.



I INTRODUCCIÓN

En las últimas décadas, los ecosistemas forestales se han visto altamente amenazados, fragmentados y degradados, causando una reducción considerable de su superficie. Lo anterior, debido principalmente a las presiones ejercidas por el ser humano tales como deforestación, incendios intencionales, sobreexplotación de especies, entre otros (Fernández et al., 2010). Dichas presiones no solo impactan a nivel de superficie, estructura y dinámica de los bosques, sino que también reducen el potencial de adaptación y respuesta de distintas especies forestales, aumentando el riesgo de extinción (Bhamondez, 2016; Caruso et al., 2015; Loo, 2011; Matesanz et al., 2019; Romero-Saritama, 2018). En consecuencia, en la actualidad se ha visto incrementado el interés por proteger y recuperar los ecosistemas forestales a través del desarrollo e implementación de estrategias y programas de conservación y restauración (Donoso et al., 2004; INFOR, 2009). Esta última es considerada una tarea prioritaria y urgente, cuyo objetivo es ayudar en el proceso de restablecimiento o recuperación de un ecosistema alterado (Society for Ecological Restoration (SER) 2004). Las prácticas de restauración pueden ser pasivas, al eliminar factores que impiden la recuperación natural o activas, lo que implica intervención al ecosistema mediante la restauración del hábitat y reforestación (Bannister

et al., 2013; Vargas & Sc, 2011). En estos proyectos, la reintroducción de especies presenta un punto crítico, ya que la selección del material de propagación a utilizar, se debe hacer en conocimiento de la calidad externa y genética del propágulo (Zamorano-Elgueta et al., 2012). La primera hace alusión a características físicas, fisiológicas y fitosanitarias del material y, por otro lado, la calidad genética engloba la diversidad genética, el valor de mejora y la identidad u origen del material. El origen es un factor de gran importancia, y según McKay et al (2005) existe suficiente evidencia que argumenta que la utilización de propágulos foráneos puede resultar ser más dañino que beneficioso para el hábitat a restaurar. De hecho, la práctica más tradicional de reforestación es abogar por la utilización de propágulos locales (Broadhurst et al., 2008; Hancock & Hughes, 2014). Lo anterior, tras la premisa de que el material de origen local se encuentra mejor adaptado para dichas condiciones (climáticas y edáficas) que un genotipo foráneo. De esta forma, la utilización de genotipos locales otorga una mayor probabilidad de supervivencia, establecimiento y crecimiento de los propágulos, y por ende, una mayor probabilidad de éxito del programa de restauración (Hancock & Hughes, 2014).

Existen dos principales riesgos al utilizar material no local; por un lado, los materiales no locales tienen mayor probabilidad de sufrir una mala adaptación, lo que conllevaría a un establecimiento deficiente y por ende a la reducción del éxito de la restauración. Y por otro lado, se encuentra el riesgo de una contaminación

o pérdida de los acervos genéticos locales, lo cual podría generar efectos desestabilizadores negativos en las poblaciones locales a través del flujo genético (Broadhurst et al., 2008; Keller et al., 2000; McKay et al., 2005). Desafortunadamente, en la mayoría de los programas de restauración el material de propagación se obtiene a partir de viveros forestales, los cuales en gran parte proveen material cuya procedencia es desconocida o no trazable (Quiroz et al., 2012). Esto genera una alta incertidumbre sobre el origen del material que se está empleando, pudiendo utilizarse genotipos foráneos sin saberlo, lo que podría ser potencialmente negativo para la restauración (Alía et al., 2005). Es aquí, donde surge la necesidad de contar con una herramienta de trazabilidad que pueda ser utilizada en sistemas de certificación forestal (Figura 1). Dicha herramienta debe ser capaz de autenticar y/o certificar el origen de material de propagación. De esta forma los propágulos a utilizar para reforestar podrán ser seleccionados con una base científica, fidedigna y confiable (Farhadi et al., 2017; Quiroz et al., 2012; Tigabu et al., 2005).

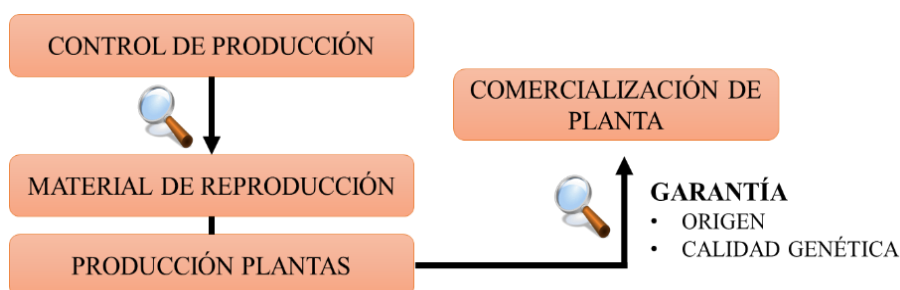


Figura 1. Esquema de sistema de certificación con la aplicación de herramientas de trazabilidad para el sector forestal. (Modificado de Alía 2005).

Hasta la fecha, herramientas visuales (microscopía y macroscopía), moleculares (basadas en ADN y proteínas), Tecnologías de la Información y Comunicación (TIC) y técnicas de análisis instrumental (espectrometría y cromatografía), han sido utilizadas para identificar, discriminar y autenticar de material vegetal (G. Campos et al., 2016; Rohman et al., 2019). No obstante, dichas técnicas son lentas, de alto costo y requieren de personal altamente capacitado para su aplicación, características que no son las deseadas para una herramienta de trazabilidad (Alía et al., 2005; Danezis et al., 2016; Danusevičius et al., 2014; Godbout et al., 2018; Pustjens et al., 2016). Por otro lado, la espectroscopia infrarroja (IR) resulta ser una herramienta bastante llamativa por sus ventajosas características, dentro de las cuales se destacan la rapidez y facilidad de uso, el requerir pequeñas cantidades de muestra con mínima o nula preparación y ser una técnica no destructiva (Dufour, 2009; Pasquini, 2003; Pustjens et al., 2016; Rohman et al., 2019; Su et al., 2017). Además, son técnicas que en su mayoría no utilizan químicos, por lo cual, se consideran amigables con el medio ambiente y presentan un relativo bajo costo de aplicación en comparación con las técnicas anteriormente mencionadas (Garrido et al., 2000; Vitale et al., 2013).

La espectroscopia IR ha sido ampliamente utilizada en distintos campos de la investigación aplicada y ha demostrado gran potencial como herramienta de autenticación de origen, tanto de productos agrícolas como forestales (Blanco & Villarroja, 2002; Loewe M. et al., 2016). Esta técnica es una modalidad de la

espectroscopia vibracional, así como la espectroscopía Raman, y ambas tienen su fundamento en la vibración de las moléculas (Monsef, 2012). El principio de la espectroscopia IR se basa en la medición de los cambios vibracionales de los enlaces de las moléculas presentes en la muestra, tras la interacción con radiación IR (Luykx & Ruth, 2008). Las longitudes de onda de dicha radiación serán absorbidas por aquellos enlaces que presenten la misma frecuencia de vibración, causando un cambio en la vibración molecular por la excitación de las moléculas a niveles energéticos superiores (M. González & Montaña, 2015; W. Rodríguez et al., 2016). Por otro lado, las demás longitudes de onda serán transmitidas o reflejadas, siendo medidas por transmitancia o reflectancia, respectivamente. Para esta última, se han desarrollado distintas modalidades que permiten medir la radiación reflejada en función de las características de la muestra a analizar: reflectancia total atenuada (ATR) para análisis de superficies o películas delgadas, reflectancia especular para el análisis de superficies altamente pulidas y reflectancia difusa o DRIFTS, específicamente para el análisis de muestras pulverizadas o con superficies rugosas (Manzano-Agugliaro et al., 2018; Monsef, 2012; Stuart, 2004).

Sin importar los fenómenos o modalidades aplicadas, la espectroscopia permite cuantificar la vibración de las moléculas, entregando como resultado un espectro de absorción, un patrón o fingerprint de cada muestra analizada (Luykx & Ruth, 2008). Dichos resultados permiten realizar análisis cuantitativos y cualitativos de

las propiedades químicas y físicas de distintos materiales biológicos (Lohumi et al., 2015; Moore et al., 2014; Tsuchikawa & Kobori, 2015; Wang & Yu, 2015)

La espectroscopia IR, como su nombre lo indica se sitúa en la región infrarroja del espectro electromagnético, es decir desde los 12.800 a 10 cm^{-1} , rango que por convención se encuentra dividido en tres regiones: IR cercano (NIR), IR medio (MIR) e IR lejano (FIR) (Lerma-García et al., 2018; Rohman et al., 2019; Skoog et al., 2008). Según Lohumi et al (2015) la espectroscopia NIR y MIR son las regiones más utilizadas con fines de autenticación y trazabilidad de origen. Por su parte la espectroscopia NIR se extiende en el rango de $780 - 2.500\text{ nm}$ ($12.800 - 4.000\text{ cm}^{-1}$), cuyos espectros exhiben bandas de absorción relacionadas a combinaciones de vibraciones moleculares. En consecuencia, los espectros obtenidos se caracterizan por presentar bandas amplias y superpuestas, lo que hace más complejo su procesamiento e interpretación (Lerma-García et al., 2018; Su et al., 2017). En sus inicios esta técnica presentaba limitaciones relacionadas con la dispersión de la luz, razón por la que se desarrolló la técnica espectroscopia en el infrarrojo cercano por Transformada de Fourier (FT-NIR). Por otro lado, la espectroscopia MIR se focaliza en el rango de $2.500 - 25.000\text{ nm}$ ($4.000 - 400\text{ cm}^{-1}$), considerado la principal región de la espectroscopia vibracional. Lo anterior, debido a que los espectros entregan información sobre las vibraciones fundamentales de las moléculas excitadas, permitiendo asignar picos característicos a grupos químicos específicos (Lohumi

et al., 2015; Manley, 2014). Además, en esta región se encuentra la zona de huella digital ($1.300 - 400 \text{ cm}^{-1}$), zona en la cual pequeñas diferencias en estructura y/o constitución de moléculas resultan en variaciones importantes en los máximos de absorción. Comparando ambas técnicas, la espectroscopia en IR medio proporciona mayor especificidad que NIR (Lin et al., 2009; Su et al., 2017).

La aplicación de espectroscopia IR va de la mano con el uso de métodos quimiométricos, los que según Rohman et al (2019), constan de la aplicación de métodos matemáticos y estadísticos que permiten el tratamiento de datos químicos. La quimiometría permite procesar los espectros, eliminando información no deseada como ruido espectral, dispersión, entre otros, así como también permite aplicar métodos exploratorios, como Análisis de Componentes Principales (PCA), y de clasificación y predicción, conocidos como métodos supervisados de reconocimiento de patrones (SPRM) (PLS-DA, SIMCA, KNN, SVM). La combinación de espectroscopia IR y quimiometría han permitido captar la señal filogenética de distintos sistemas biológicos, permitiendo distinguir entre distintas especies, géneros, variedades y subvariedades (Chalmers et al., 2012; Lohumi et al., 2015; Luykx & Ruth, 2008; Rohman et al., 2019; Wang & Yu, 2015). A su vez, se han reportado investigaciones en las cuales gracias a estas técnicas se ha logrado discriminar especies forestales según su origen geográfico a partir de madera, semillas y hojas, logrando generar modelos de predicción con una

precisión media de clasificación entre el 80 - 100% (Farhadi et al., 2017; Loewe M. et al., 2016; Sandak et al., 2011; Tigabu et al., 2005). Por otra parte, existe solo una investigación en la cual estas herramientas hayan sido aplicadas para discriminar especies leñosas según su origen filogeográfico. El estudio referido fue desarrollado por Jianguo et al (2012) en el cual se analizaron hojas de *Ulmus elongata* (Ulmaceae) de diferentes regiones a través de espectroscopia FTIR. Como resultados se obtuvieron variaciones en los máximos de absorción de distintas bandas, lo que se traduce en la existencia de diferencias a nivel de composición química (celulosa, polisacáridos, ácidos grasos, lignina y proteínas, entre otros) en las hojas de esta especie atribuibles a su origen filogeográfico. De esta forma, se tiene evidencia de que la aplicación de espectroscopia en combinación con métodos quimiométricos resulta ser una herramienta bastante prometedora para fines de autenticación y trazabilidad forestal (Chalmers et al., 2012; Rohman et al., 2019).

En Chile, una especie que requiere de esfuerzos para su conservación a través de programas de restauración y por ende, herramientas de trazabilidad es *Araucaria araucana*, conífera de gran valor social, cultural y económico (Rodríguez et al., 1983). Esta conífera en Chile se distribuye en la Cordillera de los Andes entre las latitudes 37°30'-39°30' S (Drake et al., 2009) y en la Cordillera de la Costa en dos poblaciones pequeñas y aisladas comprendidas entre las latitudes 37°20'-38°40' S (Martín et al., 2014) (Figura 2).



Figura 2. Rango de distribución natural de *Araucaria araucana* en Chile y Argentina (Recuperado de Agesen,1998)

Actualmente el estado de conservación de esta conífera es amenazado, cuya población de Nahuelbuta fue declarada “En Peligro” y las poblaciones de la Cordillera de Los Andes en categoría “Vulnerable” según el Reglamento de Clasificación de Especies del Ministerio del Medio Ambiente (MMA, 2020). Por otra parte, según la lista roja de la Unión Internacional para la Conservación de la Naturaleza (UICN), las poblaciones de ambas cordilleras han sido categorizadas “En Peligro” (UICN, 2020). Para la planificación de proyectos de

restauración de *A. araucana* es imprescindible tener en cuenta que su estructura genética poblacional y patrón de distribución geográfica han sido influenciados por eventos históricos como el último máximo glacial (UMG) (Armesto et al., 1994; Ruiz et al., 2007; Villagrán, 2018; Villagrán & Armesto, 2005). En dicho evento ambos cordones montañosos en los que habita cumplieron un rol clave como refugios glaciares, siendo probablemente área de refugio para esta conífera la Cordillera de Nahuelbuta (Armesto et al., 1994; Markgraf et al., 1995; Villagrán, 2018; Villagrán & Armesto, 2005). Los estudios sobre la filogeografía de esta especie basados en marcadores moleculares y datos palinológicos han mostrado diferencias significativas e históricas entre las poblaciones andinas y costeras, poblaciones que corresponderían a dos Unidades Evolutivamente Significativas (UES) (Figura 3) (Bekessy et al., 2002; Martín et al., 2014; Ruiz et al., 2007). En este sentido, Bekessy et al (2004) sugieren que de realizarse actividades de restauración en esta especie, se debe procurar no mezclar ambas UES, utilizando únicamente germoplasma local. Lo anterior tiene por objetivo conservar el patrimonio genético local de esta conífera y disminuir las probabilidades de una mala adaptación y/o contaminación genética por proyectos de restauración.

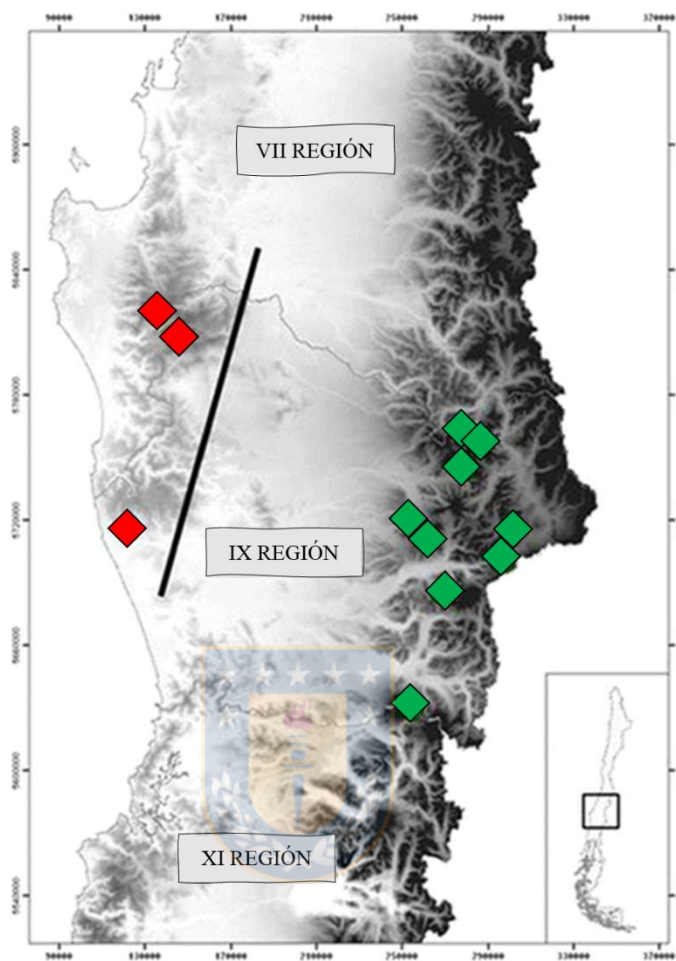


Figura 3. Distribución geográfica de las dos Unidades Evolutivamente Significativas (UES) de *A. araucana* en Chile, detectadas por análisis genéticos. (Recuperado y modificado de Martín 2014)

De esta forma, la espectroscopia IR en combinación con métodos quimiométricos podría ser una herramienta útil para facilitar la trazabilidad, en caso de que sea capaz de captar la señal filogeográfica que diferencia ambas UES de *Araucaria araucana*.

II INVESTIGACIÓN

La presente investigación busca determinar si la aplicación de espectroscopia en el infrarrojo y métodos quimiométricos son capaces de discriminar plántulas de *Araucaria araucana* procedentes de distintos orígenes filogeográficos de Chile. La hipótesis que se deriva de dicha pregunta es: “los espectros en la región infrarroja obtenidos desde plántulas de *Araucaria araucana*, tratados a través de métodos quimiométricos, permiten captar la señal filogeográfica que separa las procedencias de la Cordillera de la Costa con respecto a la Cordillera de los Andes”.



Para dar respuesta a la interrogante el objetivo general es discriminar el origen filogeográfico de plántulas de *Araucaria araucana* mediante la aplicación de técnicas de espectroscopia en distintos rangos del infrarrojo y métodos quimiométricos, para lo cual se plantearon tres objetivos específicos:

(1) Evaluar selectividad y poder de discriminación de distintos rangos de la región infrarroja en plántulas de *A. araucana* provenientes de distintos orígenes filogeográficos.

(2) Establecer y validar un modelo predictivo de clasificación del origen filogeográfico de plántulas de *A. araucana* a partir de espectros infrarrojo y el uso de métodos supervisados de reconocimiento de patrones.

(3) Identificar bandas espectrales que más aportan la clasificación de orígenes filogeográficos de plántulas de *A. araucana*, con el fin de extraer información sobre las principales diferencias químicas entre las plántulas de diferente origen.



III METODOLOGÍA

3.1 Obtención material de estudio

Las plántulas de *A. araucana* fueron facilitadas por el Instituto Forestal INFOR en el marco del proyecto de Sistema Integrado de Monitoreo de Ecosistemas Forestales Nativos (SIMEF). Se muestrearon semillas de 12 poblaciones de la distribución natural de la especie (Figura 4), considerando la estructura genética reportada para esta conífera en Martín et al (2014). Una vez colectadas las semillas, fueron viverizadas bajo condiciones ambientales semi-controladas en dependencias del vivero Carlos Douglas ubicado en la comuna de Yumbel, Región del Biobío, perteneciente a la empresa forestal CMPC. Como pretratamiento para estimular la germinación, se realizó un corte en la semilla y se dejaron en remojo por 48 horas en agua destilada, bajo cámara de frío a 4°C. Los cuidados posteriores fueron realizados según los estándares de calidad de la empresa.

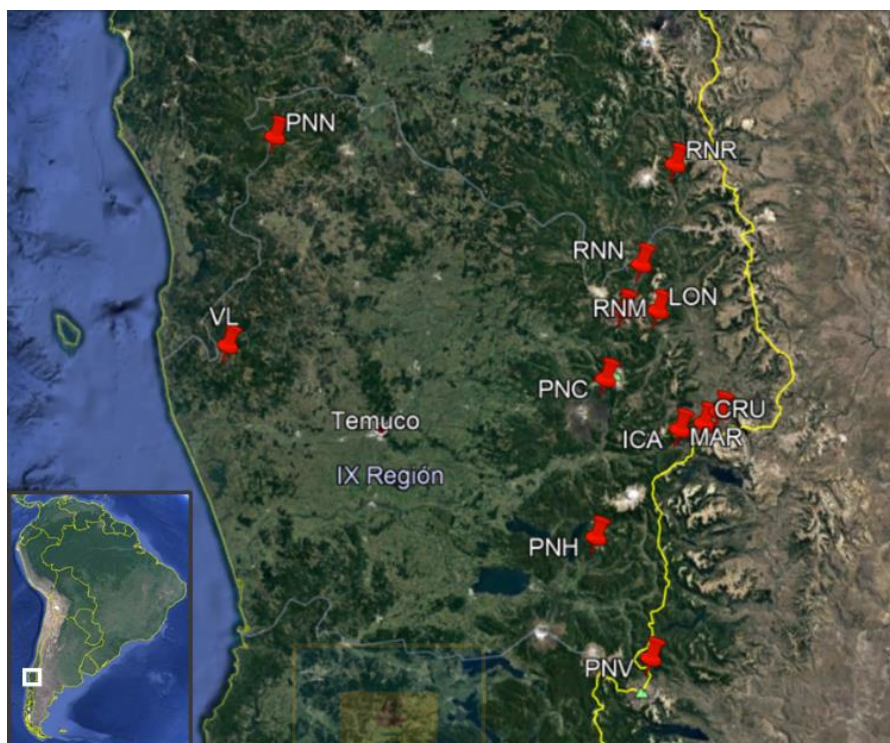


Figura 4. Localidades muestreadas de semillas de *A. araucana* en su distribución natural en Chile, periodo de semillación 2018. PNN: Parque Nacional Nahuelbuta; VL: Villa las Araucarias; RNR: Reserva Nacional Ralco; RNN: Reserva Nacional las Nalcas; LON: Lonquimay; RNM: Reserva Nacional Malalcahuello; PNC: Parque Nacional Conguillio; ICA: Icalma; MAR: Marimenuco; CRU: Cruzaco; PNH: Parque Nacional Huerquehue; PNV: Parque Nacional Villarrica. Imagen obtenida de Google Earth.

3.2 Muestreo de plántulas

El muestreo de plántulas se realizó a los 11 meses desde la siembra. Se seleccionaron nueve plántulas por sitio de muestreo, muestreando un total de 108 plántulas. Las plántulas seleccionadas se dispusieron en bolsas de papel debidamente rotuladas. Posterior a ello, las muestras fueron trasladadas directamente a dependencias de la Universidad de Concepción.

3.3 Procesamiento de plántulas

Las plántulas fueron llevadas directamente a estufa a 40°C por siete días hasta peso constante. Posteriormente se llevó a cabo el proceso de pulverización. Para la facilitación de dicho proceso el material fue recortado con una tijera, seleccionando los 2/3 superiores de las plántulas. El proceso de molienda o pulverización se realizó con un molino de bolas mezclador Retsch MM400 a una frecuencia de 30.0 1/s por un tiempo de 3 minutos. Una vez pulverizado el material fue almacenado en frascos negros en oscuridad hasta la colecta de espectros.



3.4 Adquisición de espectros

Las muestras ya procesadas fueron analizadas en distintas regiones del espectro electromagnético: visible, IR cercano e IR medio. Para lo anterior se utilizaron cuatro equipos (Figura 5), los cuales se especifican en la Tabla 1 junto con su respectiva cobertura espectral, resolución y scans. Para el análisis en el visible e infrarrojo cercano, se trabajó bajo el módulo de reflectancia difusa, por lo cual para coleccionar los espectros las muestras fueron directamente dispuestas en los viales respectivos para cada equipo. Para la obtención de los espectros de absorción se trabajó con la función de $A = \log(1/R)$, a excepción de los espectros FT-NIR que fueron derivados con la función de Kubelka-Munk (Pasquini, 2003).

Por otro lado, para el análisis en IR medio se colectó información a través del fenómeno de transmitancia, siendo necesario hacer pastillas de bromuro de potasio (KBr) para cada muestra. Las muestras fueron mezcladas homogéneamente entre un 5 a 10 p/p % en KBr en un mortero de ágata, posterior a ello, dicha mezcla fue prensada a alta presión generándose una pastilla translúcida apta para ser escaneada en dicho rango del espectro electromagnético (Sherman, 1997) (Figura 6). Se tomaron espectros por duplicado, y considerando que en la macrozona costa existen menos localidades, se tomaron mayor cantidad de espectros por muestra (por cada muestra de la costa se colectaron 5 espectros por duplicado) para compensar la brecha en el tamaño muestral.



Tabla 1. Información sobre región espectral, rango de cobertura (nm y cm^{-1}), resolución y número de scans para cada espectrofotómetro utilizado.

Equipo	Región	Rango de cobertura (nm) (cm^{-1})		Resolución	N° Escaneos
Oceans Optics	VIS-NIR	400-1037	28985-9643	0.2 nm	5
MPA-Bruker	FT-NIR	800-2500	12489-3996	16 cm^{-1}	32
MicroPhazir	NIR	1595-2396	7720-4173	8 nm	5
Agilent Cary 630	FTIR	2500-15384	3999-650	4 cm^{-1}	32

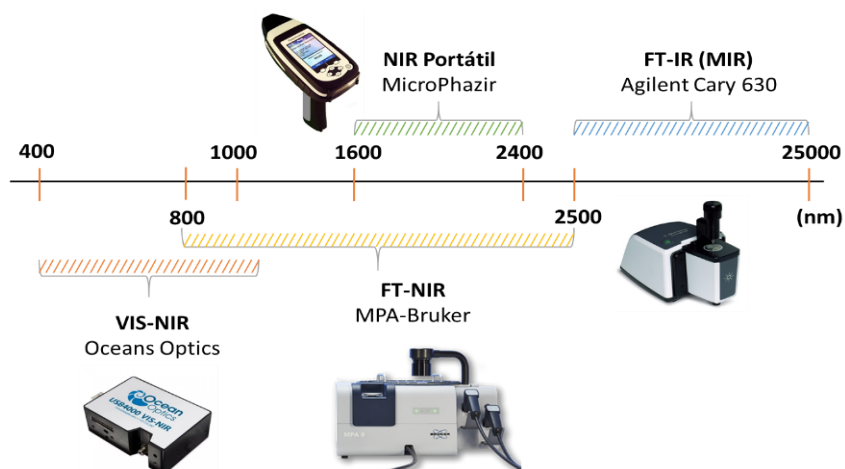


Figura 5. Espectrofotómetros utilizados para análisis de muestras. Equipo Oceans Optics para rango VIS-NIR (345 - 1037 nm), equipo MPA-Bruker para rango NIR con transformada de Fourier (800 - 2500 nm), equipo portátil MicroPhazir para rango NIR (1595 - 2396 nm) y equipo Agilent Cary 630 para rango IR medio con transformada de Fourier (2.500 - 15384 nm).

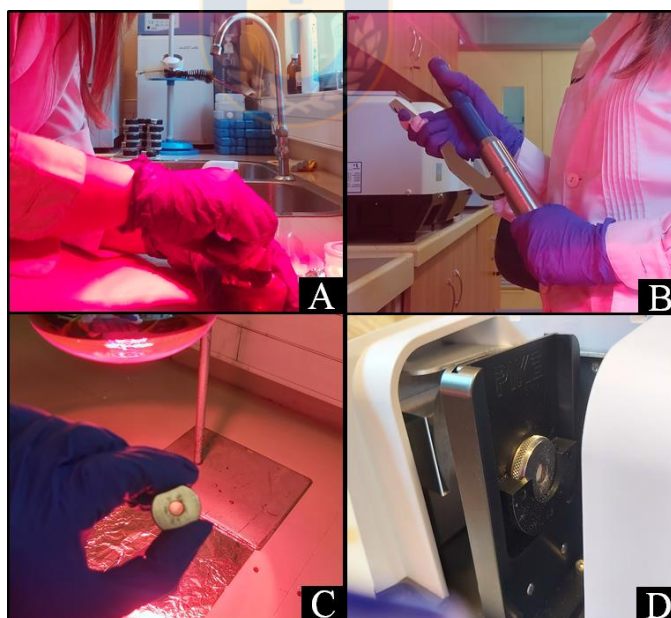


Figura 6. Proceso de generación pastillas de KBr para análisis en IR medio. A: molienda de KBr y muestra en mortero de ágata bajo luz infrarroja. B: presado a alta presión, con prensa manual. C: pastilla de KBr + muestra en su soporte previo a ser analizada. D: pastilla en su soporte, posicionada en el equipo para ser analizada por transmitancia.

3.5 Métodos quimiométricos

3.5.1 Pretratamiento de datos

El pretratamiento espectral permite mejorar el posterior desarrollo de modelos matemáticos y el análisis de datos, tras la eliminación de información no deseada sin perder información importante a través de transformaciones y preprocesados (Lerma-García et al., 2018; Manley, 2014). Las características de las muestras tales como el tamaño irregular de las partículas, pueden tener efectos sobre las propiedades de los espectros (ruido, dispersión, desplazamiento de línea base, entre otros) los cuales dificultan el posterior análisis. En función de lo anterior, es que se aplican transformaciones matemáticas con el fin de reducir su efecto en los espectros, siendo las más comúnmente aplicadas las derivadas espectrales, la corrección de línea base, suavizado y aquellas cuya finalidad es compensar la dispersión como MSC (Multiplicative Scatter Correction) y SNV (Standard Normal Variate). Por otro lado, también se aplican pre-procesados, cuya finalidad es acentuar las diferencias entre espectros muy similares, siendo los más comunes el centrado en la media y el autoescalado (Foley et al., 1998; Williams & Norris, 2001). En este caso las cuatro bases espectrales fueron tratadas con variadas combinaciones de transformaciones y preprocesado mediante el Software Pirouette 4.5 de Infometrix, de las cuales para efectos de la investigación solo se presentan los resultados de los mejores tratamientos en función de las

características espectrales y como respondían los modelos. Los respectivos pre-tratamientos para cada rango y modelo se presentan en la Tabla 2.

Tabla 2. Preprocesados y transformaciones aplicadas a las cuatro bases espectrales para análisis exploratorios (PCA) y SPRM (K-NN y SIMCA). (*) Pre-procesados. SNV: Variación normal estándar. MC: Centrado en la media. AU: Autoescalado. () Indica la ventana de puntos.

Base	PCA	K-NN	SIMCA
VIS-NIR	1 ^{ra} derivada (25) MC*	1 ^{ra} derivada (15) MC*	1 ^{ra} derivada (15) MC*
FT-NIR	MSC MC*	2 ^{da} derivada (19) MC*	2 ^{da} derivada (9) MC*
NIR (port)	MSC MC*	2 ^{da} derivada (25) MC*	2 ^{da} derivada (5) MC*
FTIR	Corrección línea base Normalización MC*	2 ^{da} derivada (45) MC*	Normalización 2 ^{da} derivada (25) MC*

3.5.2 Análisis de datos

Mediante el Software Piruette 4.5 de Infometrix los datos espectrales fueron analizados de forma exploratoria mediante un Análisis de Componentes Principales (PCA). Por otro lado, también se aplicaron Métodos Supervisados de Reconocimiento de Patrones (SPRM) como K-NN (los vecinos más cercanos) y SIMCA (Independent Modeling of Class Analogy).

A. Análisis exploratorio

El PCA es un método exploratorio que permite la reducción de dimensionalidad a través de la transformación de variables no correlacionadas entre sí en una combinación lineal (componente principal: PCs). La generación de PCA permitió identificar muestras con valores atípicos (outliers), las cuales serán eliminadas para el posterior desarrollo de los modelos. Por otro lado el PCA también permitirá observar el comportamiento de las muestras en los distintos componentes, permitiendo apreciar agrupaciones por similitudes (Lucio-Gutiérrez, 2012; McGarigal et al., 2000).



B. Métodos Supervisados de Reconocimiento de Patrones (SPRM)

Se generaron modelos de clasificación de los vecinos más cercanos con las bases espectrales obtenidas. K-NN es un método que permite clasificar muestras según su similaridad con otras, sin requerir coincidencia exacta de patrones para su agrupación. Por consiguiente, las muestras más parecidas se encuentran más próximas y las más disímiles más alejadas (IBM Knowledge Center, 2021). Como se mencionó anteriormente también se trabajó con modelos de clasificación SIMCA, método basado en la generación de un PCA para cada clase en forma independiente. De esta forma se construye un “espacio para cada clase”, cuyos límites determinan la pertenencia o no pertenencia de una muestra a dicha clase,

con un cierto nivel de confianza (Lucio-Gutiérrez, 2012). Para la generación de ambos modelos de clasificación se trabajó con el 70% de las muestras, correspondiente al set de entrenamiento de los modelos y estipularon dos clases, considerando cada cordillera como una clase. Para los rangos que resulten más discriminantes (mejores resultados de clasificación) se evaluará la robustez de los respectivos modelos, mediante 2 nuevas redistribuciones de las muestras entre los sets de entrenamientos y validación. De esta forma se tendrá resultados de clasificación de los modelos con 3 sets distintos.

3.6 Validación modelos de clasificación



La validación de los modelos permite evaluar el desempeño de estos, estimando la incertidumbre de las futuras predicciones usando muestras nuevas. Para la validación se utilizó el 30% del total de las muestras, las cuales no fueron utilizadas en el set de entrenamiento. Se validó mediante el porcentaje de muestras que resulten clasificadas correcta o incorrectamente, con un 95% de confianza (Massart et al., 1998). Para la evaluación del rendimiento del modelo de clasificación generado se evaluará la sensibilidad (S_n), especificidad (S_p), tasa de error de clasificación (ER) y precisión de clasificación (CA). Por su parte, la S_n es aquel parámetro que describe la capacidad que tiene el modelo de reconocer correctamente la pertenencia de muestras a una clase determinada. Por el contrario, la S_p describe la capacidad que presenta el modelo para

rechazar muestras no pertenecientes a esa clase. Los valores de ambos parámetros van de 0 a 1. Los parámetros serán calculados con las ecuaciones descritas por Ballabio & Consonni (2013), donde TP es verdadero positivo, TN verdadero negativo FP falso positivo y FN falso negativo. Otro parámetro a evaluar en los mejores modelos es la robustez, mediante 3 variaciones del set de entrenamiento y validación, analizando la variación en los %CA.

$$Sn = \frac{TP}{TP + FN}$$

$$Sp = \frac{TN}{FP + TN}$$

$$ER = 1 - \frac{(Sn + Sp)}{2}$$

$$CA = \frac{(TP + TN)}{(TP + FP + TN + FN)}$$

3.7 Interpretación de datos

Para la interpretación de datos, se identificarán las bandas que más aportan al modelo para la discriminación de orígenes de filogeográficos de plántulas de *A. araucana*. Para ello se analizarán los loadings de componentes (PCA) y el poder de discriminación (SIMCA). Posterior a ello, se realizará búsqueda en bibliografía sobre posibles compuestos que pueden estar relacionados con absorbancias en las longitudes de ondas con mayor poder de discriminación. Por otro lado, en IR

medio para detectar diferencias en los espectros de ambas clases, se calculó el promedio de la segunda derivada y suavizado de 35 puntos (algoritmo de Savitzky-Golay) de cada clase y la respectiva desviación estándar para cada longitud de onda. De esta forma, se obtuvieron 6 espectros correspondientes a el promedio de cada clase, el promedio de cada clase menos sus respectivas desviaciones estándar por cada variable y el promedio de cada clase más sus respectivas desviaciones estándar por cada variable. Se analizó el mínimo de cada espectro, a cuyas bandas que presenten diferencias (que mantengan el comportamiento con y sin las desviaciones estándar) se realizó asignación química por búsqueda bibliográfica.



IV RESULTADOS Y DISCUSIÓN

1. Rango VIS-NIR (345 - 1037 nm)

Los espectros de reflectancia colectados presentaron extremos ruidosos los cuales fueron excluidos, trabajando finalmente en el rango 475 - 922 nm (Figura 7A). Los espectros de pseudo absorbancia ($A = \text{Log} (1/R)$) de plántulas de *A. araucana* en el rango VIS-NIR se observan en la Figura 7B, en los cuales se observan tres bandas en las longitudes 538, 616 y 672 nm. Según Reusch (2013) y Walsh et al (2020) estas absorben en la zona espectral del verde, naranja y rojo, respectivamente. El espectro obtenido de *A. araucana* presenta un comportamiento similar (patrón) a los espectros reportados para otras especies forestales como *Pinus sylvestris* L. (Danusevičius et al., 2014), *P. strobus* L. y *Quercus rubra* L. (Hallett et al., 1997). En el rango 400 – 530 nm (zona del verde) absorben los pigmentos carotenoides, específicamente en la banda de 535 nm absorben los pigmentos fotoprotectores xantófilas (Carranco et al., 2011) y en la región del rojo donde se observa la banda con mayor absorbancia corresponde a la zona de absorción de la clorofila a y b (600-700 nm), siendo las longitudes que reflejan las propiedades comunes de la vegetación verde (Hoffer, 1978).

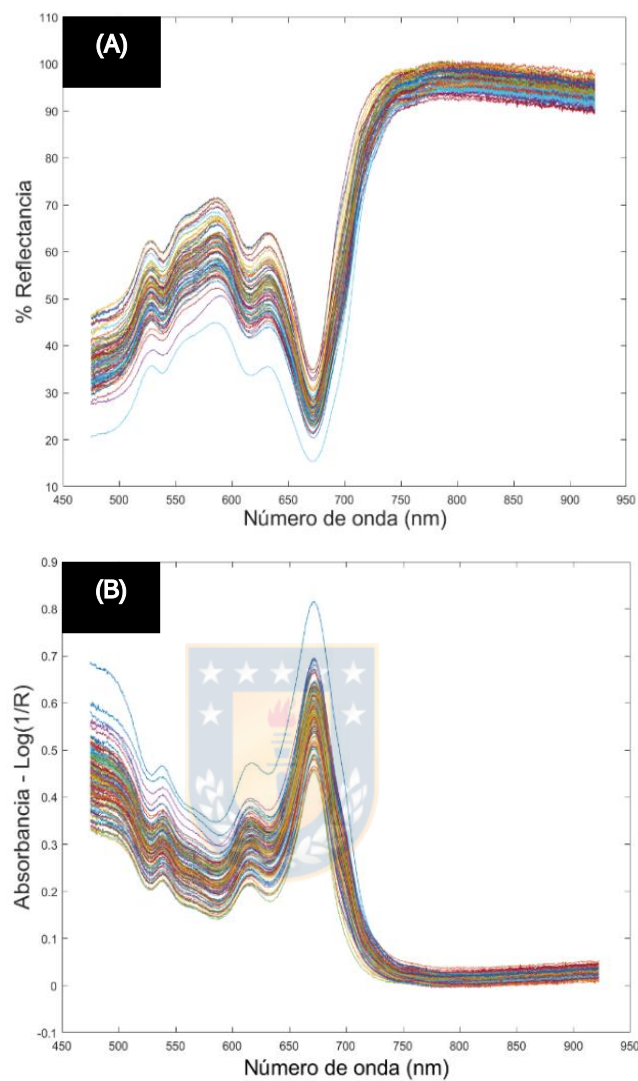


Figura 7. Espectros VIS-NIR sin extremos ruidosos, en rango 475-922 nm de plántulas de *A. araucana*. (A) % reflectancia y (B) pseudo absorbancia ($\text{Log}(1/R)$).

En la Figura 8 se presentan los espectros transformados y pre-procesados para los métodos de reconocimientos de patrones (Revisar Tabla 2). Mediante de la distancia de Mahalanobis y prueba Q se identificaron 3 outliers, los cuales fueron excluidos de los análisis.

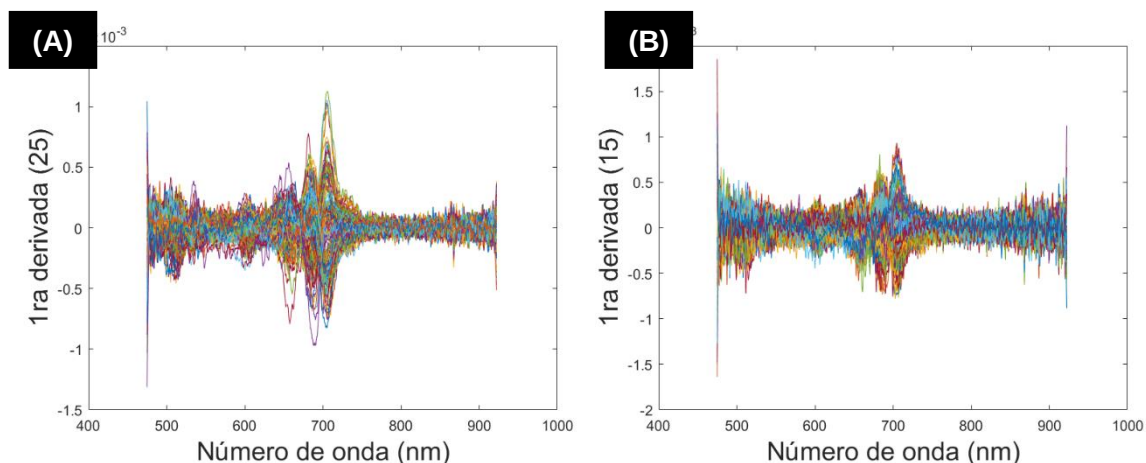


Figura 8. Espectros en región VIS-NIR pre-tratados para (A) PCA con 1^{ra} derivada con ventana de 25 puntos y (B) K-NN y SIMCA con 1^{ra} derivada con ventana de 15 puntos. Ambos fueron centrados en la media.

Tras los análisis de PCA, el gráfico de cargas de los tres primeros componentes permitió apreciar que el rango 782 - 922 nm no aporta información relevante o influyente para ninguno de estos componentes (Figura 9), por lo que el análisis exploratorio se realizó excluyendo dicho rango. Para determinar cuántos componentes analizar se observó el porcentaje de varianza explicada por cada PC y el porcentaje de varianza acumulada (Tabla 3). Se aprecia que los 3 primeros PCs logran explicar un 78.79% de la varianza y que las demás componentes aportan muy poca información, principalmente ruido.

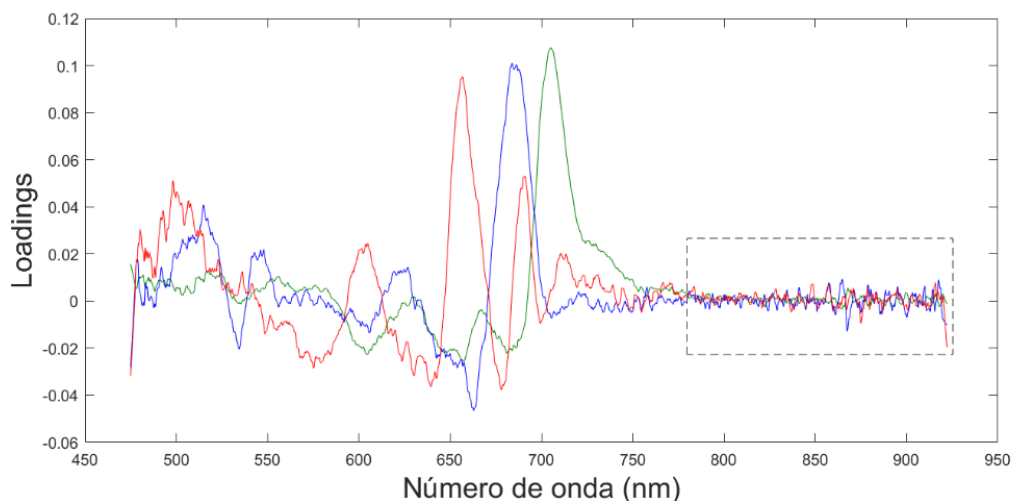


Figura 9. Loadings de los primeros 3 PCs de PCA en rango VIS-NIR. Espectro verde, azul y rojo corresponden a loadings de PC 1, 2 y 3 respectivamente. Se marca dentro del rectángulo punteado zona espectral poco influyente.



Tabla 3. Componentes principales calculados para matriz espectral VIS-NIR.

Componente principal	Varianza explicada (%)	Varianza acumulada (%)
PC1	51.63	51.63
PC2	20.45	72.08
PC3	6.72	78.79
PC4	4.22	83.02
PC5	3.51	86.53

En la Figura 10 se presentan los gráficos de puntuación (Scores) obtenidos con los tres primeros PCs y sus respectivos gráficos de cargas (Loadings). Se observa que el PC1 el cual explica un 51.6% de la varianza, es el componente que mejor explica una tendencia de separación entre ambas clases.

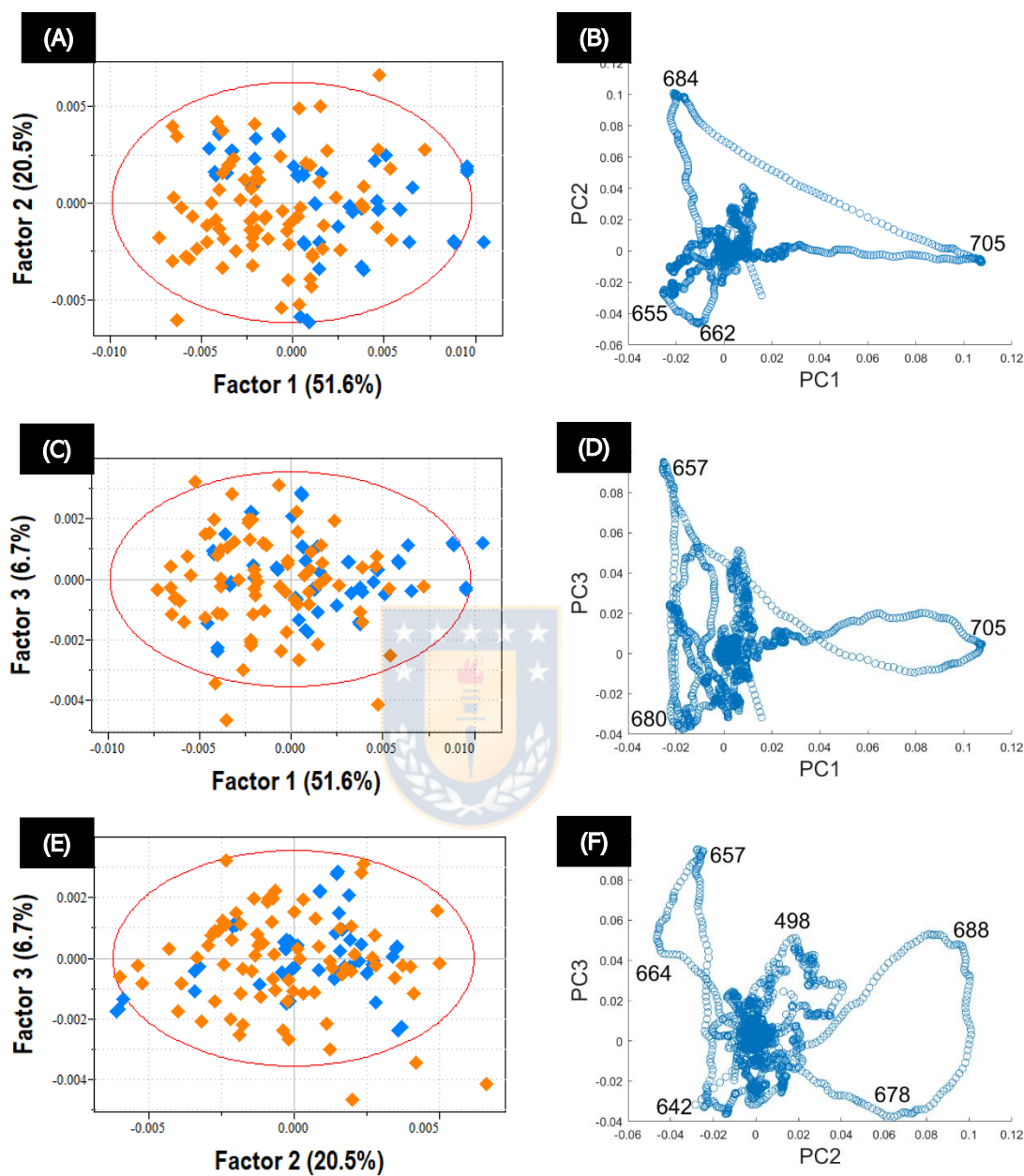


Figura 10. Gráfico de puntuaciones entre los primeros 3 PCs y sus respectivos gráficos de cargas, obtenidos con matriz VIS-NIR. Cuadros azules corresponden a muestras de la Costa y cuadros naranjos a muestras de los Andes.

El PC1 resultó ser el componente que reúne la información más útil para diferenciar muestras de *A. araucana* costeras y andinas. Al observar los gráficos de cargas del PC1 versus PC2 o PC3 (Figura 10B y D), se aprecia que la banda cercana a los 705 nm es influyente para la clase costera (CS1) y aquellas cercanas a 684, 655, 622 nm son más influyentes para la clase andina (CS2). Para visualizar e identificar más claramente las bandas más importantes para la separación según el PC1, se analizó su gráfico de cargas por separado (Figura 11). Las bandas más influyentes corresponden a las longitudes de onda 517, 555, 575, 605, 655, 684 y 705 nm, bandas asociadas a la de absorción de carotenoides y clorofilas (Azcón-Bieto et al., 2008). Se aprecia que la banda más importante para las muestras andinas es en la longitud 705 nm, seguida por las bandas bajo los 585 nm, las cuales podrían relacionarse a absorción de clorofila a (705 nm) y de clorofila a + b (<585 nm). Por otro lado, las bandas más influyentes para las muestras costeras corresponden a 605, 655 y 684 nm, relacionadas principalmente a clorofila a (Chla). De esta forma PCA indica que las plántulas de ambos orígenes difieren en la absorbancia de pigmentos fotosintéticos, específicamente en la región del verde (490-570 nm), naranja (585 - 620 nm) y principalmente del rojo (620 - 780 nm) (Reusch, 2013). Según Merzlyak et al (2003) la absorción en la región del rojo se relaciona con la absorbancia de las clorofilas. Cabe mencionar que cerca de los 650 y 678 nm la clorofila b y a respectivamente presentan su absorción característica, y es por

ello que la banda de 680 nm es muy útil para el cálculo de distintos índices relacionados al contenido de clorofila (Walsh et al., 2020).

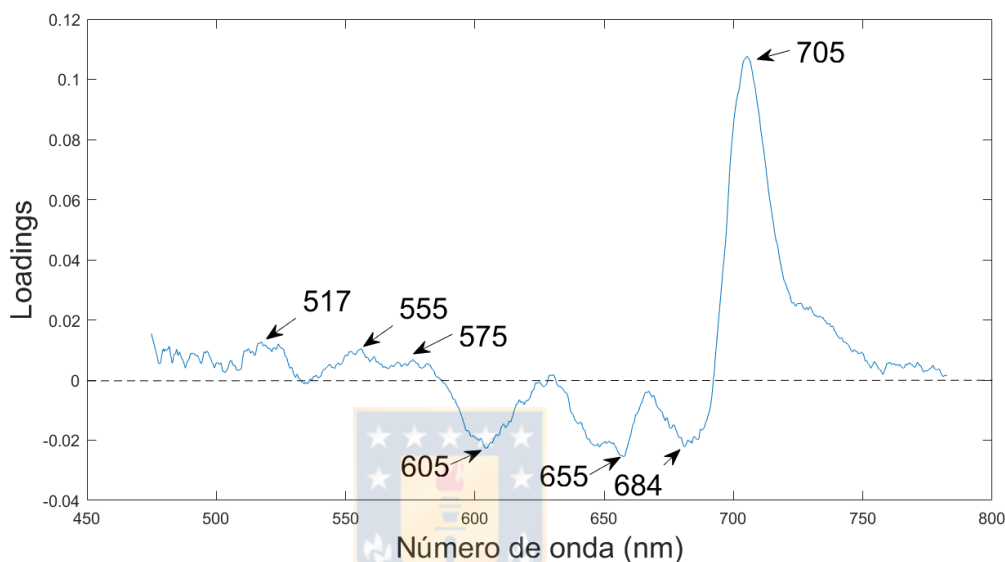


Figura 11. Gráfico de cargas de PC1 de matriz VIS-NIR obtenido de plántulas de *A. araucana*.

Para el desarrollo de los modelos supervisados se utilizaron todas las longitudes de onda (475 - 922 nm). El modelo K-NN se desarrolló con 2 vecinos ($K=2$), para el entrenamiento del modelo se obtuvieron 11 muestras mal clasificadas, obteniendo un 88.89% de precisión de clasificación (%CA). En la etapa de validación solo 4 muestras de CS2 fueron mal clasificadas, logrando un %CA del 90.5% (Para ver resultados de clasificación revisar Anexo 1). Para el modelo SIMCA, se modelaron 4 PCs para cada clase con una distancia interclase de 0.24. En el entrenamiento se logró clasificar correctamente el 98.99% de las muestras (1 muestra mal clasificada) y un 90.5% tras validación externa (4

muestras mal clasificadas) (Para ver resultados de clasificación revisar Anexo 2). Como se mencionó anteriormente las clases presentan muy poca distancia entre sí, en la Figura 12 se observan los gráficos de Coomans del entrenamiento y validación. En ambos casos la mayoría de las muestras de ambas clases se encuentran bajo ambos umbrales de pertenencia a una clase, pero a pesar de ello SIMCA logra clasificar y predecir de forma correcta con un bajo error.

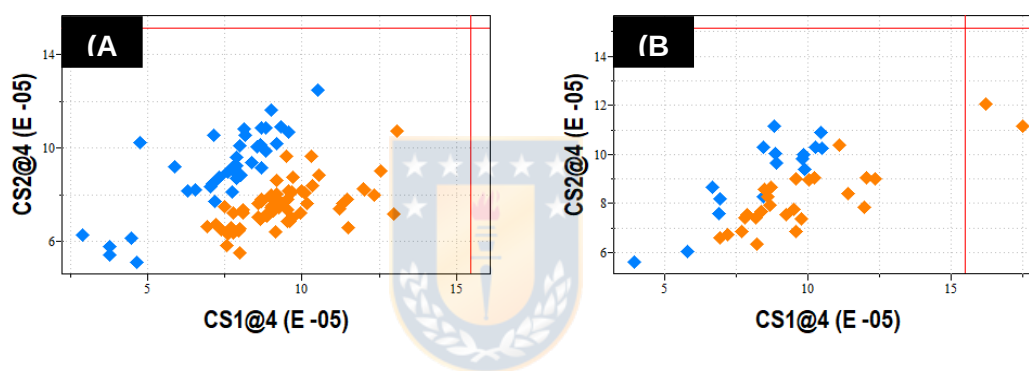


Figura 12. Gráfico de Coomans de modelo SIMCA en rango VIS-NIR, umbrales para un 95% de confianza. (A) set de entrenamiento y (B) set de validación. Cuadros azules corresponden a muestras de la Cordillera de la Costa (CS1) y cuadros naranjos a la Cordillera de los Andes (CS2).

La gráfica de poder de discriminación que entrega SIMCA permite extraer información sobre las variables más importantes para discriminar (Figura 13). A los 700 nm se observa la banda más discriminadora, lo que concuerda con lo observado en PCA cuya banda más influyente se encontraba alrededor de los 705 nm, asociado a clorofila. Según los resultados reportados por Merzlyak et al (2003) la banda en 700 nm tiene una alta correlación con el contenido de clorofila ($r^2 > 0.90$). Seguido a ella se observa que las bandas 535, 641, 657, 688, 690, 711

y 713 nm presentan altos valores de discriminación. De esta forma las longitudes más importantes o discriminantes se encuentran en la región del verde (490-570 nm) y del rojo (620 – 780 nm), siendo este último el que presenta las variables más discriminantes (641 - 700 nm). Como se mencionó anteriormente, la banda en 535 nm corresponde por literatura a la absorción de xantófilas. Las longitudes restantes corresponden a absorción de Chla, a excepción de la longitud 641 nm que corresponde al máximo de absorción de Chlb (642 nm) (Ustin et al., 2009

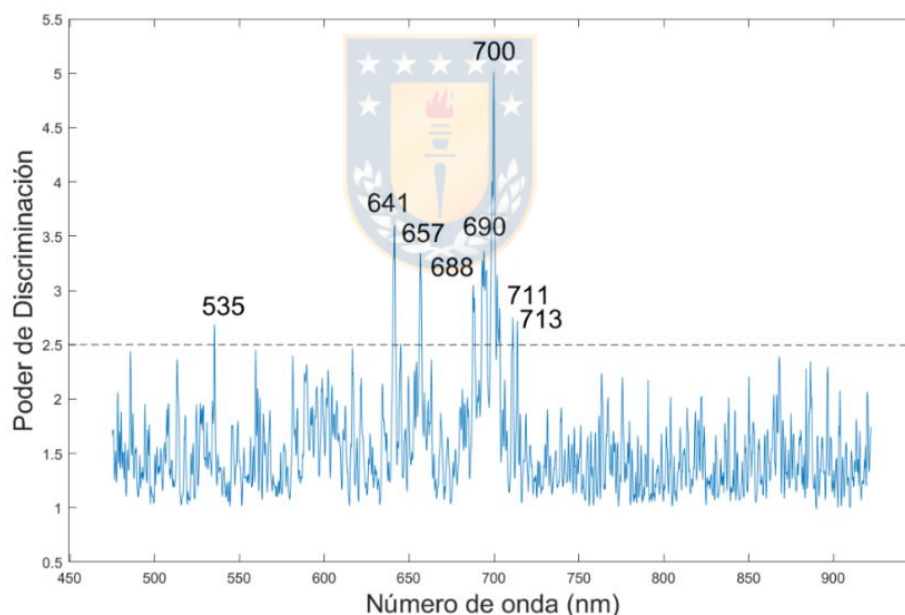


Figura 13. Gráfico de poder de discriminación de variables de modelo SIMCA en rango VIS-NIR para discriminar entre orígenes filogeográficos de plántulas de *A. araucana*. Arbitrariamente solo se señalan longitudes como poder de discriminación mayor a 2.5.

Las diferencias de absorbancia observadas en el rango VIS-NIR indican que las muestras difieren en función de su origen filogeográfico principalmente en el contenido de pigmentos verdes y xantófilas. Lo anterior, podría estar reflejando diferencias en los niveles de estrés a los que se ven expuestos ambos núcleos poblacionales de *Araucaria* (Datt, 1998; Zarco-Tejada et al., 2004). Por su parte, Li et al (2019) afirman que el color y por ende, la pigmentación de las hojas son una respuesta al estrés y a la variación ambiental, pero a su vez también a la variación genética, indicando que la variación de pigmentos y principalmente clorofilas está parcialmente controlada por genes. En especies como *Sassafras tzumu* H. (Lauraceae) se reportan que parámetros como el color de las hojas y por ende el contenido de pigmentos presentan altas estimaciones de heredabilidad. Asimismo, se reporta que para la especie *Acer rubrum* L más conocida como Arce rojo (Sapindaceae) el color de sus hojas presenta una influencia significativa de los padres (Townsend & McIntosh, 1993). Por otro lado, se tiene evidencia de que la concentración de clorofila foliar se ve principalmente influenciada por la temperatura, hallándose mayor concentración en aquellas plántulas de climas más cálidos (Azcón-Bieto et al., 2008; Danusevičius et al., 2014). Al existir diferencias en el contenido de clorofilas, se asume una variación en la tasa fotosintética de las plántulas, la cual según Premoli et al. (2012) tiene un control genético significativo.

En concreto, el procesamiento y análisis de la información espectral adquirida en el rango VIS-NIR ha permitido diferenciar y discriminar plántulas de *A. araucana* de ambos orígenes filogeográficos. Lo anterior se atribuye a diferencias en el contenido de pigmentos fotosintéticos. Los modelos K-NN y SIMCA desarrollados presentaron una precisión de clasificación del 90%, donde este último modelo permitió identificar como las bandas más discriminadoras aquellas relacionadas con la absorbancia de las clorofilas y compuestos carotenoides xantófilas.

2. Rango NIR con Transformada de Fourier FT-NIR (12489 - 3996 cm^{-1})



En los espectros FT-NIR adquiridos se aprecia que las plántulas de *A. araucana* presentan ocho bandas, específicamente en las longitudes 4258, 4335, 4690, 5176, 5701, 5793, 6326 y 6904 cm^{-1} (Figura 13). Richardson et al (2004) señala que la región del IR cercano se caracteriza por presentar bandas de tonos, sobretonos y combinaciones. Según la tabla de asignación de bandas especificada por Schwanninger et al (2011) la absorbancia en la longitud 4258 cm^{-1} se asocia a la combinación del estiramiento y deformación de C-H y el segundo sobretono de la flexión C-H₂, lo que en bibliografía se ha asociado a polisacáridos como celulosa, al igual que la absorción en 4335 cm^{-1} , tras el estiramiento C-O combinado con el estiramiento O-H o a la combinación de la flexión y estiramiento de C-H₂. Por su parte la absorción en los 4690 cm^{-1} se

atribuye proteínas, específicamente a señales de Amida I y Amida II. La absorción en 5176 cm^{-1} se asocia a estiramientos del segundo sobretono de C-O atribuido a carbohidratos. El estiramiento del primer sobretono de C-H de grupos C-H₂ se observan en las longitudes 5701 y 5793 cm^{-1} . Por último, la absorción en 6326 y 6904 cm^{-1} se atribuyen primer y segundo sobretono de estiramiento de N-H atribuible a proteínas (Tüker-Kaya, 2017). En la Figura 14 se presentan los espectros FT-NIR pretratados para la aplicación de los métodos de reconocimientos de patrones.

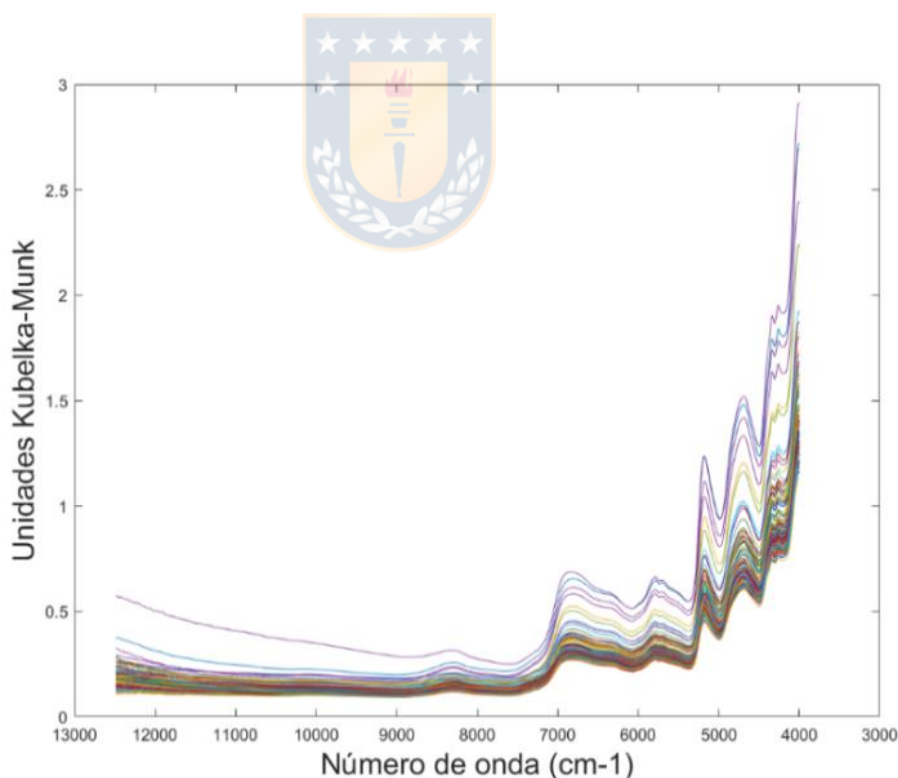


Figura 14. Espectros brutos de absorbancia en unidades de Kubelka-Munk en rango NIR con Transformada de Fourier, de plántulas de *A. araucana*.

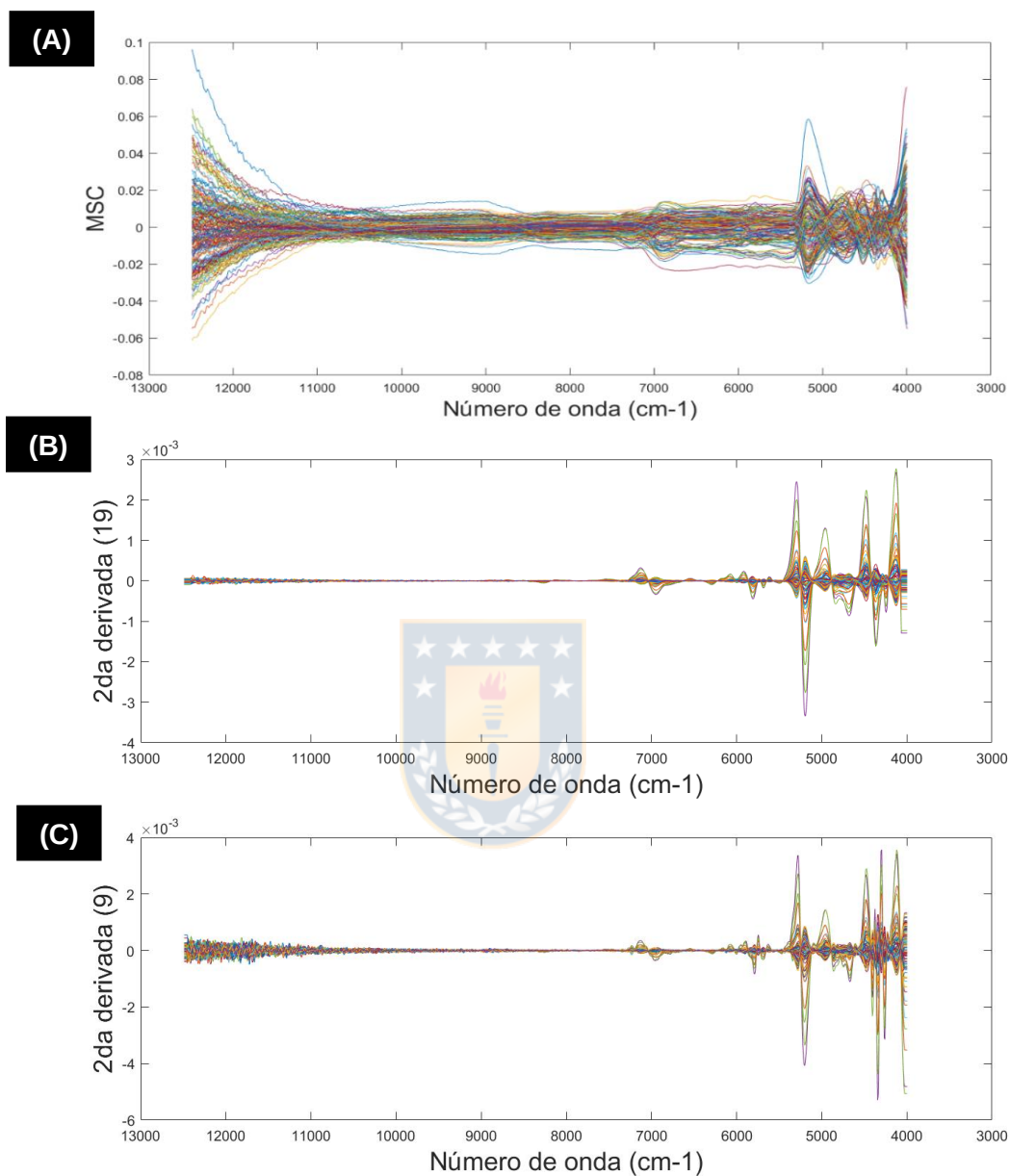


Figura 15. Espectros FT-NIR de plántulas de *A. araucana* pre-tratados para (A) PCA con MSC, (B) K-NN con 2^{da} derivada de 19 puntos y (C) SIMCA con 2^{da} derivada de 9 puntos. Todos fueron centrados en la media.

Según la distancia de Mahalanobis y gráfico de residuales se identificó 1 espectro atípico el cual se excluyó de los análisis. Se analizaron los tres primeros componentes, dado que logran explicar un 90.24% de la varianza (Tabla 4). En la Figura 14 se presentan los distintos gráficos de puntuación de los versus de los primeros 3 PCs y sus respectivos gráficos de cargas. El PC2 (25.9%) permite observar una pequeña tendencia de separación entre ambas clases (Figura 15A y E), donde bandas cercanas a 12489, 5168, 5438 cm^{-1} son las más influyentes para CS2 y 3996 cm^{-1} para CS1.

Tabla 4. Componentes principales calculados para matriz espectral FT-NIR.



Componente	Varianza explicada (%)	Varianza acumulada (%)
PC1	56.54	56.54
PC2	25.88	82.42
PC3	7.82	90.24
PC4	3.36	93.60
PC5	2.49	96.09

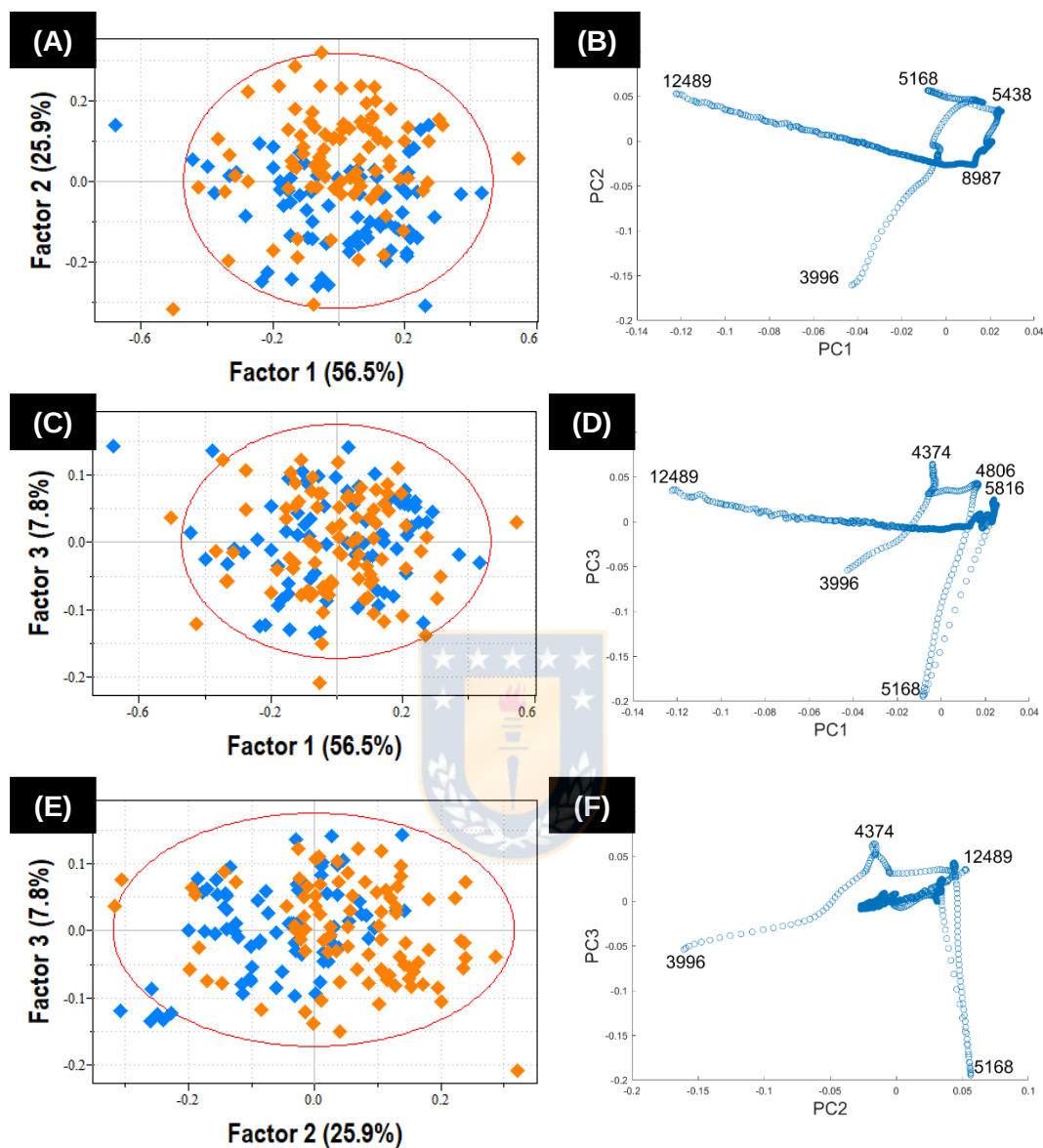


Figura 16. Gráfico de puntuaciones entre los primeros 3 PCs y sus respectivos gráficos de cargas, obtenidos con matriz FT-NIR. Cuadros azules corresponden a CS1 y cuadros naranjos a CS2.

En la Figura 17 se observa el gráfico de cargas del PC2. Se aprecia que los rangos 11610 - 7136 y 4550 - 3996 cm^{-1} son más influentes para CS1 y el rango

7136 - 4550 cm^{-1} a CS2, siendo las bandas más influyentes 5169 y 4706 cm^{-1} . Según Schwanninger et al (2011) en el rango 12800 - 7000 cm^{-1} (rango influyente para la clase costera) absorben los primeros armónicos y segundos armónicos de las vibraciones de estiramiento O-H y N-H, los primeros armónicos de las bandas de combinación C-H y segundo y tercer armónicos del estiramiento C-H y N-H. En cuanto a las longitudes influyentes para la clase andina, es decir, entre 7000 - 6000 cm^{-1} absorben sobre todo el primer sobretono de las vibraciones de estiramiento O-H atribuibles principalmente a carbohidratos y entre 6000 - 5000 cm^{-1} el primer sobretono de vibraciones de estiramiento C-H alifáticas y aromáticos y bandas de combinación O-H. Finalmente el rango 5000 - 4000 cm^{-1} , se reporta como difícil de asignar dadas las diferentes combinaciones químicas que se pueden dar entre estiramientos de enlaces C-H, C=O y/o deformación O-H.

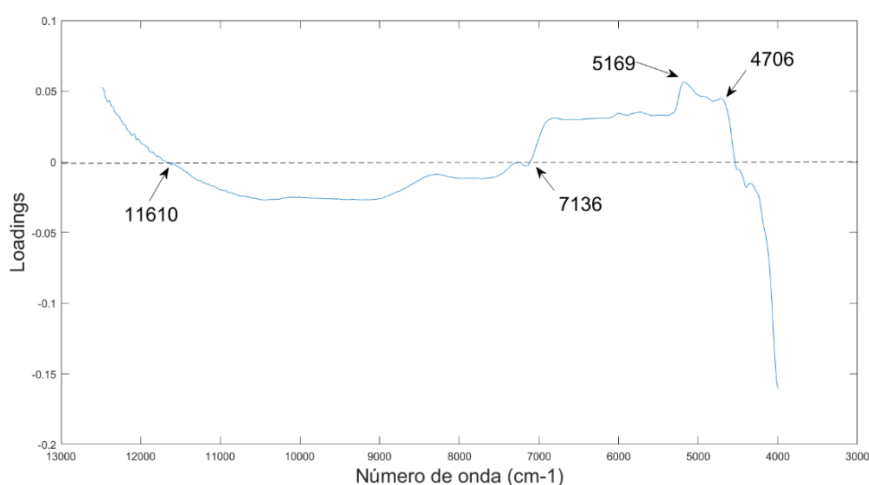


Figura 17. Gráfico de cargas de PC2 en rango FT-NIR. Bandas por sobre 0 son influyentes para CS2 y por debajo de 0 para CS1.

Continuando con los métodos supervisados, en el entrenamiento del modelo K-NN (K=2) se obtuvo un 82.8% de precisión de clasificación, y en la validación con set externo un 86.27% (Para ver resultados de clasificación revisar Anexo 1). Mientras que en SIMCA (3 PCs para cada clase con una distancia interclase de 0.28) se obtuvo un %CA del 93.97% cuya validación resultó en un 88.24% (Para ver resultados de clasificación revisar Anexo 2). En la Figura 18 se presentan los gráficos de Coomans del set de entrenamiento y de validación, observándose el mismo comportamiento que en VIS-NIR. En el gráfico de poder de discriminación de variables en SIMCA (Figura 19) se observa que las bandas más discriminadoras son 11108, 10669, 8702, 7128, 6912, 6858, 6488, 5276, 4999, 4752, 4235, 4189 cm^{-1} . Como ya se mencionó en el rango 12800 – 7000 cm^{-1} se observan bandas de primeros y segundos sobretonos y combinación de estiramiento O-H, C-H y N-H. La banda en 7128 cm^{-1} se ha atribuido a la combinación de estiramiento y deformación de C-H de carbohidratos (Tüker-Kaya, 2017). En 6912, 6858 y 6488 cm^{-1} se observa la banda correspondiente al estiramiento del primer sobretono de O-H, bandas que en literatura han sido atribuidas a grupos -OH de compuestos fenólicos y al enlace intermolecular de almidón respectivamente. La banda en 5276 cm^{-1} se asocia a vibraciones en el enlace C=O. Ali et al (2001) indican que la banda en 4998 cm^{-1} (4999 cm^{-1}) se asocia a la combinación de estiramiento simétrico de N-H y Amida II, atribuible a proteínas. La banda observada cerca de los 4764 cm^{-1} se asocia a O-H de carbohidratos, al igual que la banda 4235 cm^{-1} . Por

último, las bandas en 4189 y 4189 cm^{-1} se han asignado a grupos aromáticos (Beratto et al., 2017)

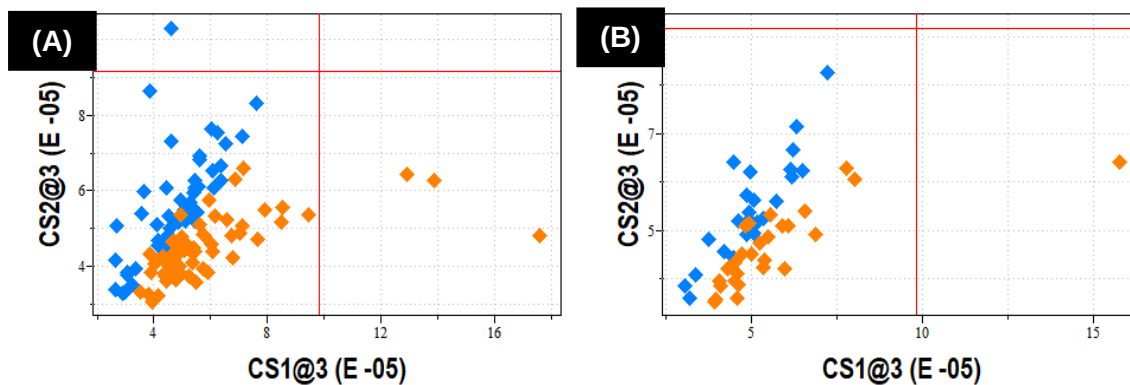


Figura 18. Gráfico de Coomans de modelo SIMCA NIR portátil, umbrales para un 95% de confianza. (A) set de entrenamiento y (B) set de validación. Cuadros azules corresponden a muestras de la Cordillera de la Costa (CS1) y cuadros naranjos a la Cordillera de los Andes (CS2).

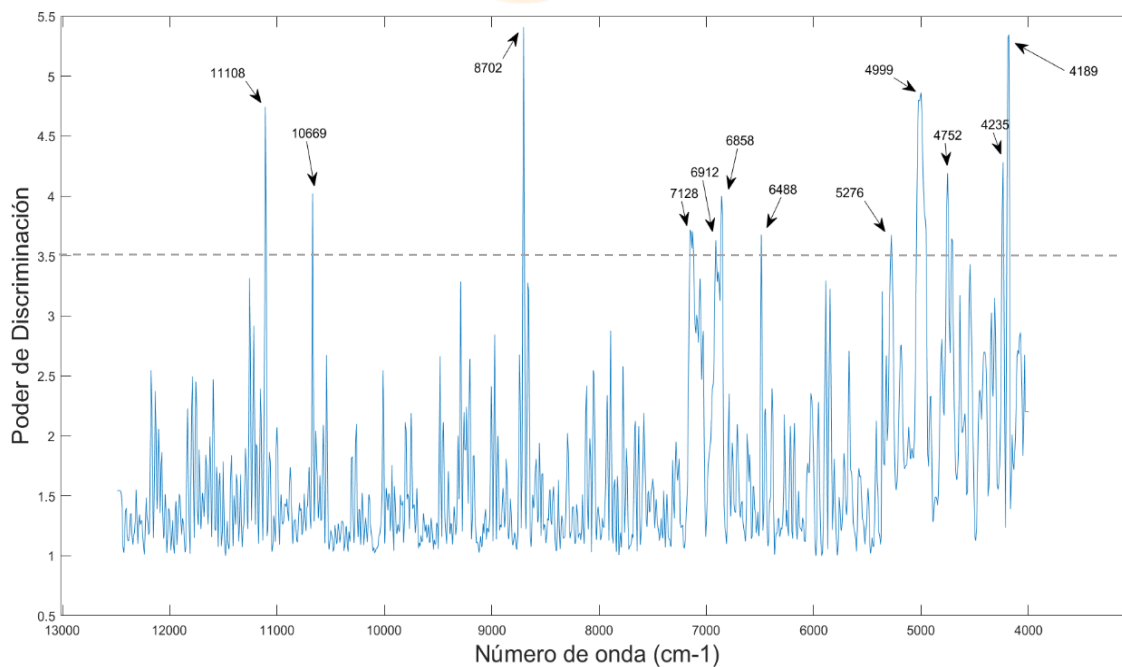


Figura 19. Poder de discriminación de variables en modelo SIMCA en rango FT-NIR. Solo se señalan longitudes con mayor poder de discriminación.

En definitiva, el análisis de la información espectral adquirida en FT-NIR, permitió diferenciar mediante PCA plántulas de *A. araucana* costeras y andinas con el PC2 (25.88% de la varianza). En cuanto a modelación K-NN y SIMCA, se logró una precisión de clasificación del 86 y 88% respectivamente. Cuyas bandas más discriminadoras se atribuyen principalmente a carbohidratos, proteínas y compuestos fenólicos.

3. Rango NIR (1595 - 2396 nm) equipo portátil

Los espectros obtenidos de las muestras de *A. araucana* adquiridos en la región NIR con el equipo portátil, presentan cuatro bandas en las longitudes de onda 1927, 2114, 2277 y 2307 nm (Figura 20). Según literatura, la banda observada en 1927 nm se atribuye estiramientos del segundo sobretono de C-O de carbohidratos (Tüker-Kaya, 2017). Asimismo, en literatura la banda en 2114 nm se ha asociado a celulosa, como resultado de la combinación de flexión y estiramientos de enlaces O-H. La banda en 2277 nm según Stenberg et al (2010) corresponde al estiramiento de C-H alifáticos y por último, la absorción en 2307 nm se ha asignado a la combinación de estiramiento N-H y C-O atribuido a proteínas (Tüker-Kaya, 2017).

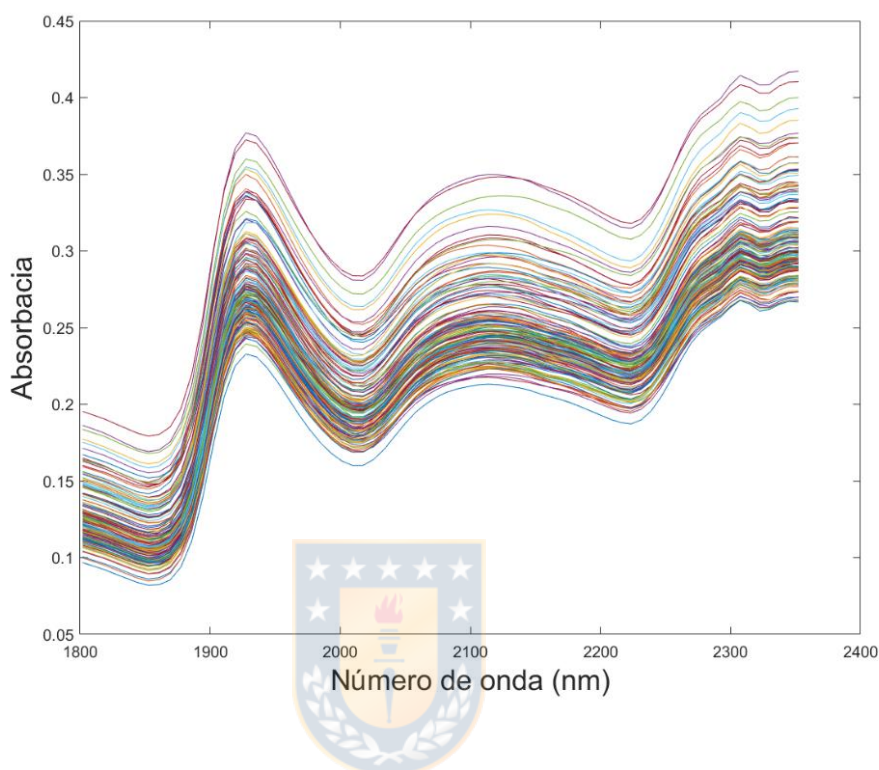


Figura 20. Espectros brutos de absorbancia en rango NIR con equipo portátil de plántulas de *A. araucana*.

En la Figura 21 se observan los espectros NIR pretratados para análisis exploratorio y métodos supervisados. En NIR no se detectaron muestras outliers. Se analizaron los tres primeros PCs los cuales explican el 91.5% de la varianza (Tabla 5). En la Figura 22 se muestran los distintos gráficos de puntuación de los versus de los primeros 3 PCs y sus respectivos gráficos de cargas, donde el PC2 (12.5% de varianza) permite observar una tendencia de separación entre ambas clases, cuyas variables más influyentes corresponden a 1902 nm para CS1 y 2145 nm para CS2 (Figura 22B y F).

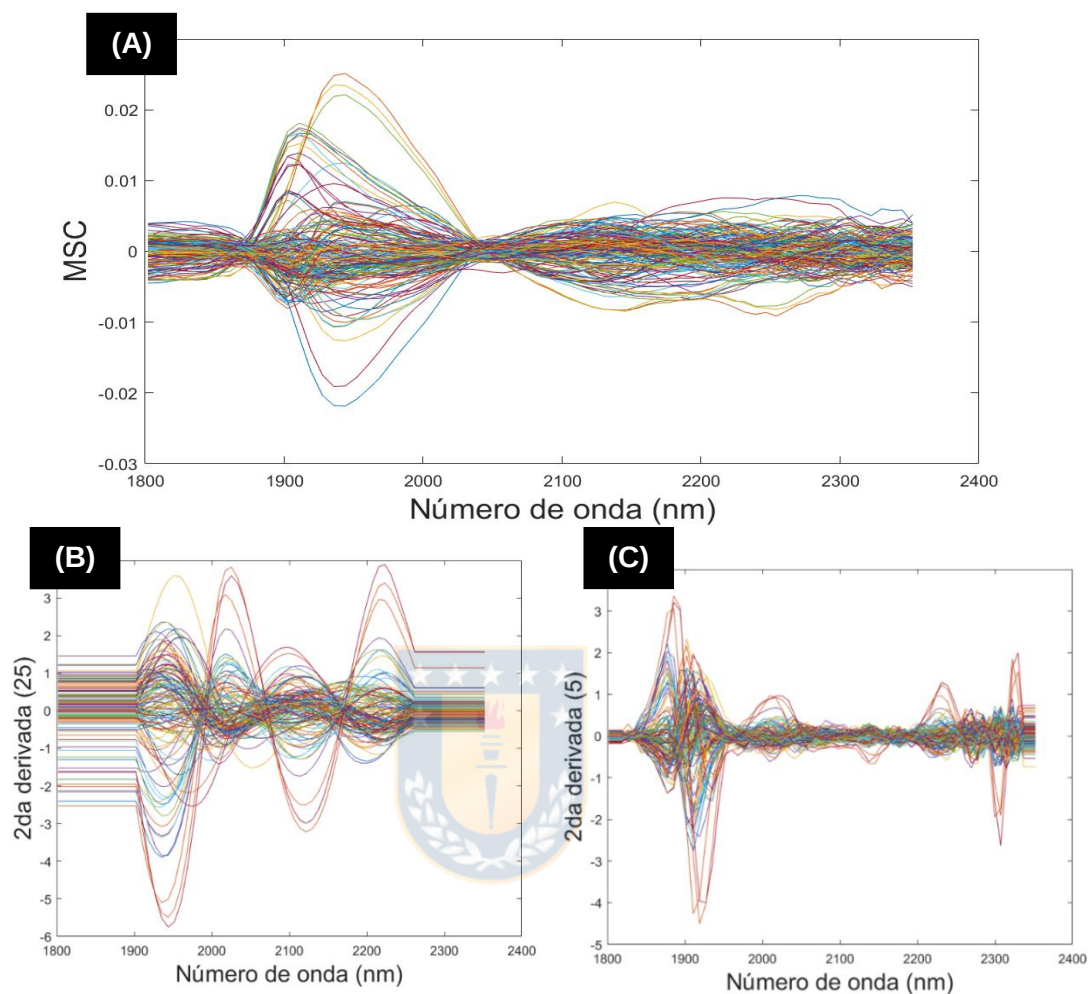


Figura 21. Espectros de absorbancia en rango NIR con equipo portátil, de plántulas de *A. araucana* pre-tratados para (A) PCA con MSC, (B) K-NN con 2^{da} derivada de 25 puntos y (C) SIMCA con 2^{da} derivada de 5 puntos. Todos fueron centrados en la media.

Tabla 5. Componentes principales calculados para matriz espectral NIR.

Componente	Varianza explicada (%)	Varianza acumulada (%)
PC1	72.28	72.28
PC2	12.52	84.79
PC3	6.72	91.52
PC4	6.00	97.52
PC5	1.00	98.53

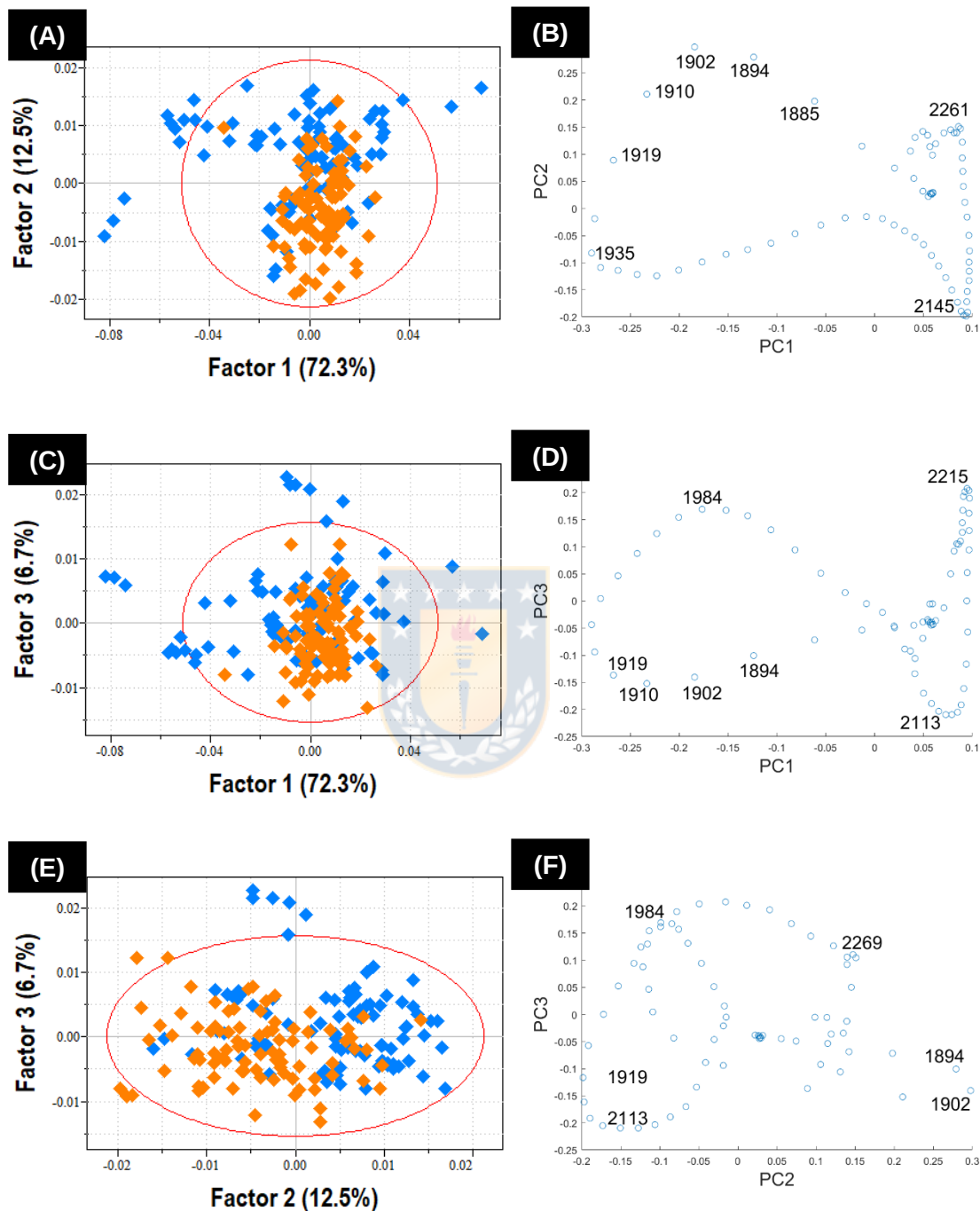


Figura 22. Gráfico de puntuaciones entre los primeros 3 PCs y sus respectivos gráficos de cargas, obtenidos con matriz NIR portátil. Cuadros azules corresponden a muestras de la costa y cuadros naranjos a muestras de los andes.

Tras el análisis del gráfico de cargas del PC2 (Figura 23) se aprecia que las bandas más influyentes para CS1 corresponden a 1902, 2269, 2292 y 2337 nm. La banda en 1902 nm corresponde al estiramiento del segundo sobretono de C=O y la banda en 2269 nm a la combinación de estiramientos O-H y C-O, ambas bandas asignadas a carbohidratos. Schwanninger et al (2011) reporta que en dichas longitudes absorben las hemicelulosas, celulosa y ligninas, al igual que la banda en 2337 nm correspondiente a la combinación del estiramiento C-H y deformación de C-H₂. La banda observada en 2292 nm se atribuye a la señal de combinación del estiramiento N-H y C-O de proteínas. Por otro lado, las bandas más influyentes para CS2, son 1968 y 2145 nm. Según bibliografía ambas bandas son asignadas a proteínas, siendo señales de combinación de estiramiento asimétrico de N-H y Amida II o al estiramiento del segundo sobretono de C=O, y a la combinación de Amida I y Amida III, respectivamente (Tüker-Kaya, 2017). Según estos resultados las plántulas de la Cordillera de la Costa presentarían mayor contenido de carbohidratos y menor contenido de proteínas en comparación con aquellas de origen Andino.

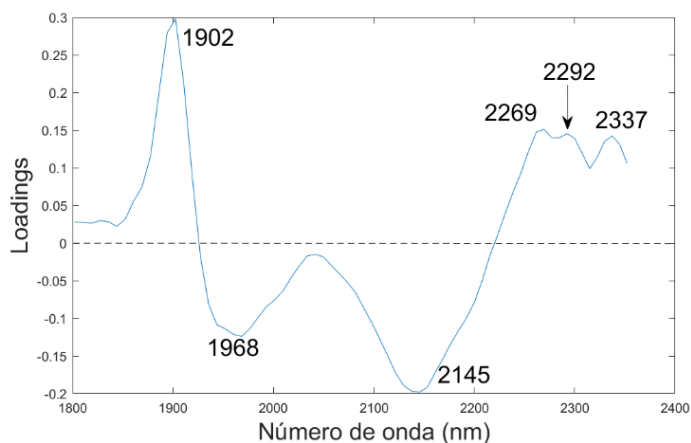


Figura 23. Gráfico de cargas de PC2 para información adquirida en rango NIR con equipo portátil. Bandas por sobre 0 son influyentes para CS1 y por debajo de 0 para CS2.

En cuanto a los modelos supervisados, K-NN (K=5) obtuvo un %CA de 90.48%, cuya validación externa resultó en un 85.2% de clasificación exitosa (Para ver resultados de clasificación revisar Anexo 1). Con la modelación SIMCA (6 y 4 PCs para CS1 y CS2, separadas por una distancia de 0.53) se obtuvo un 88.89% de clasificación exitosa en el entrenamiento y validación (Para ver resultados de clasificación revisar Anexo 2). En la Figura 24 se presentan los gráficos de Coomans del entrenamiento y validación, donde se observa nuevamente que la mayoría de las muestras se encuentran por debajo de ambos umbrales, es decir, en la zona de confusión. Al analizar el gráfico de poder de discriminación de variables del modelo SIMCA desarrollado (Figura 25) se observa que las bandas que más aportan a la discriminación de ambas clases son 1844, 1877, 1902, 1943, 1960, 2001, 2017, 2041, 2066, 2145, 2277 nm. En la longitud 1844 nm se observan la combinación de vibraciones de estiramiento de O-H y estiramiento

del segundo sobretono de C-O asociado a carbohidratos al igual que las bandas en 1902 y 1943 nm, correspondientes a señales del estiramiento del segundo sobretono de C=O (Schwanninger et al., 2011). Según Stenberg et al (2010) la banda en 1960 nm corresponde al estiramiento de C-OH de compuestos fenólicos. En cuanto a las bandas en 2001, 2066 y 2145 nm se atribuyen a proteínas, donde las bandas corresponden a la combinación del estiramiento simétrico de N-H y Amida II, a la combinación de flexión N-H y Amida, y a la combinación de Amida I y Amida III respectivamente (Ali et al., 2001; Tüker-Kaya, 2017). La banda en 2017 nm según Budić-Leto et al (2011) corresponde al estiramiento de N-H y C-O y/o a la combinación de Amida I y Amida II. En cuanto a la banda 2041 nm en literatura la banda en 2080 nm se asocia estiramiento O-H y deformación C-H atribuido a azúcares. La banda en 2277 nm se asocia a estiramientos de C-H alifáticos (Stenberg et al., 2010).

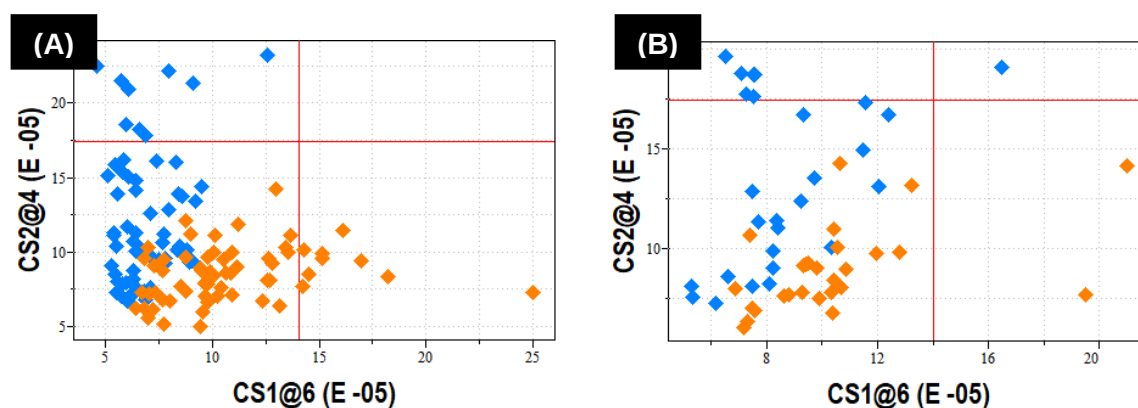


Figura 24. Gráfico de Coomans de modelo SIMCA NIR portátil, umbrales para un 95% de confianza. (A) set de entrenamiento y (B) set de validación. Cuadros azules corresponden a muestras de la Cordillera de la Costa (CS1) y cuadros naranjos a la Cordillera de los Andes (CS2).

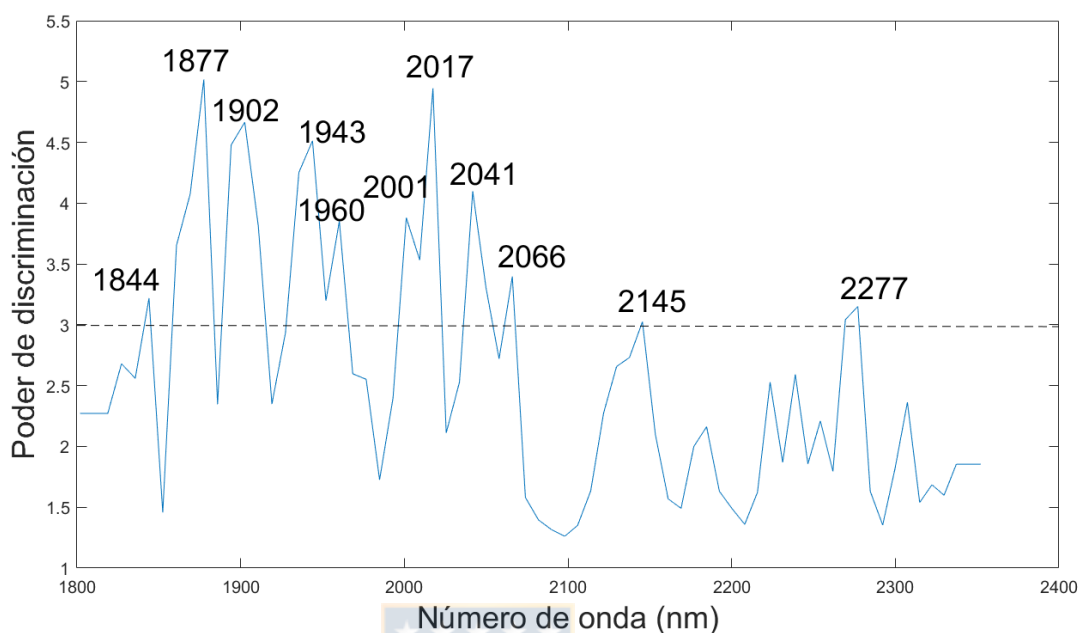


Figura 25. Poder de discriminación de variables en modelo SIMCA en rango NIR portátil. Solo se señalan longitudes con valor de discriminación mayor a 3.

En definitiva, la información espectral adquirida en el rango NIR con equipo portátil, permitió observar una tendencia de separación de ambas clases por el PC2 (12.5% de varianza). A su vez los modelos K-NN y SIMCA, lograron discriminar plántulas de andes y costa con una precisión de clasificación del 85 y 89% respectivamente. La discriminación lograda se atribuyó principalmente a proteínas y carbohidratos, donde presumiblemente las muestras costeras presentan mayor contenido de carbohidratos y las andinas mayor contenido proteico.

4. Rango FTIR (3999 - 650 cm^{-1})

Los espectros adquiridos en la región del IR medio no presentaron extremos ruidosos que debiesen ser excluidos, no obstante, este rango se caracteriza por presentar una zona muy influenciada por la humedad ambiental (zona espectral de absorción de O-H y C-O, 2300 cm^{-1}) (Jaya et al., 2015). Considerando que esta la zona espectral 2600 - 2000 cm^{-1} no presenta bandas que provean información química de las muestras, para fines de esta investigación fue excluido (Figura 26). En cuanto a las bandas de absorción de los espectros de las plántulas de *A. araucana*, se observa que son múltiples bandas las cuales se atribuyen a los grupos funcionales las principales macromoléculas (proteínas, lípidos, polisacáridos y ácidos nucleicos) (Beekes et al., 2007; Dorado et al., 2001; Invernizzi et al., 2018; Kačuráková et al., 2000; Liu et al., 2013; Pandey & Pitman, 2003). En los espectros se logran distinguir bandas en las longitudes 3410, 2965, 2933, 2855, 1246, 1103, 1067 cm^{-1} , sumado a ellas, la mayoría de las bandas observadas se encuentran en la región de huella dactilar (1800 – 500 cm^{-1}). La banda en 3410 cm^{-1} corresponde a la señal de estiramiento de O-H y N-H, atribuido a carbohidratos, proteínas, alcoholes y compuestos fenólicos (Pandey & Pitman, 2003; Tüker-Kaya, 2017). La banda en la longitud 2965 cm^{-1} se presenta la señal de estiramiento asimétrico CH₃, bandas en 2933 y 2855 cm^{-1} se atribuyen a estiramientos asimétricos y simétricos respectivamente de CH₂. Estas bandas principalmente se atribuyen a lípidos con una pequeña contribución

de proteínas, carbohidratos y ácidos nucleicos (Beekes et al., 2007; Malek et al., 2014; Severcan et al., 2003). Continuando, la banda en 1246 cm^{-1} Pandey & Pitman (2003) reportan que en la longitud 1244 cm^{-1} se observa el anillo de la unidad siringil y estiramiento C-O de ligninas y xilanos. En literatura en el rango $1110 - 1105\text{ cm}^{-1}$ se observa la banda de absorción de estiramiento simétrico C-O-C de la cutina (Tüker-Kaya, 2017). La banda observada en 1076 cm^{-1} correspondería a estiramientos de C-O de alcoholes secundarios, atribuibles a grupo éster metoxilos en celulosa, hemicelulosas y ligninas (Lopes et al., 2018). En la Figura 27 se presentan los espectros pretratados según se indicó en la Tabla 2.

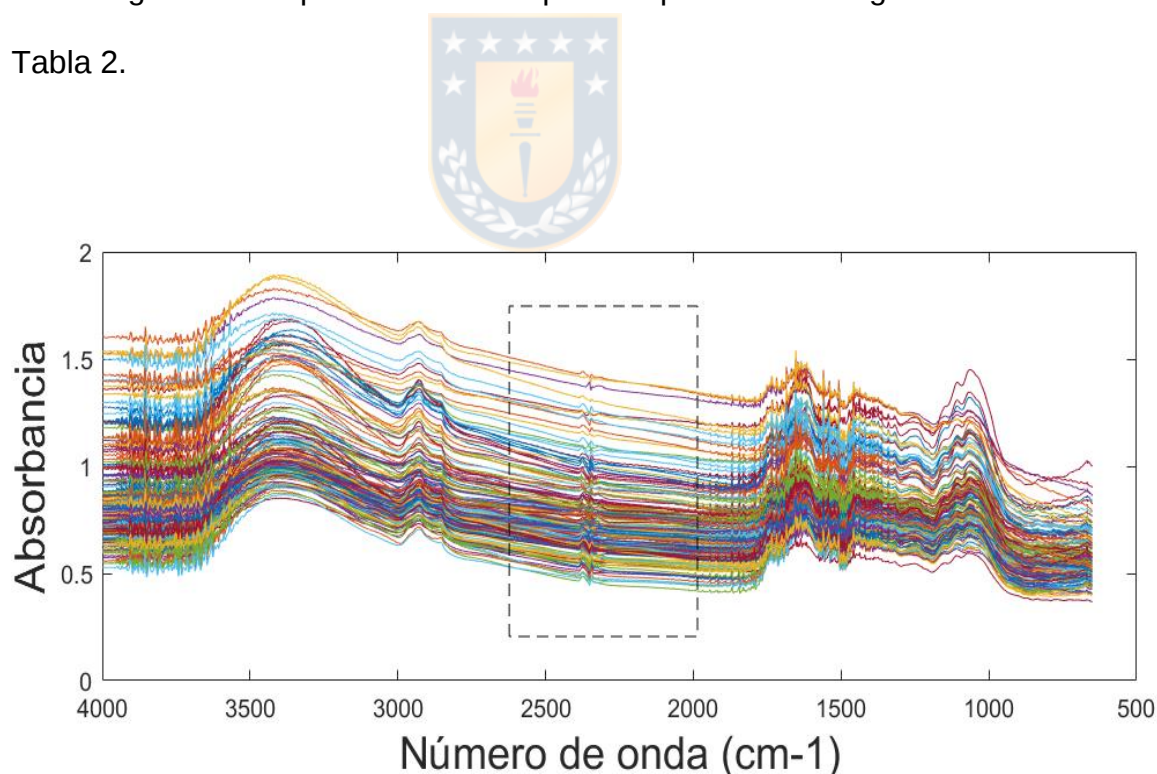


Figura 26. Espectros de absorbancia FTIR plántulas de *A. araucana*, se señala en recuadro punteado rango de $2600\text{ a }2000\text{ cm}^{-1}$ excluido.

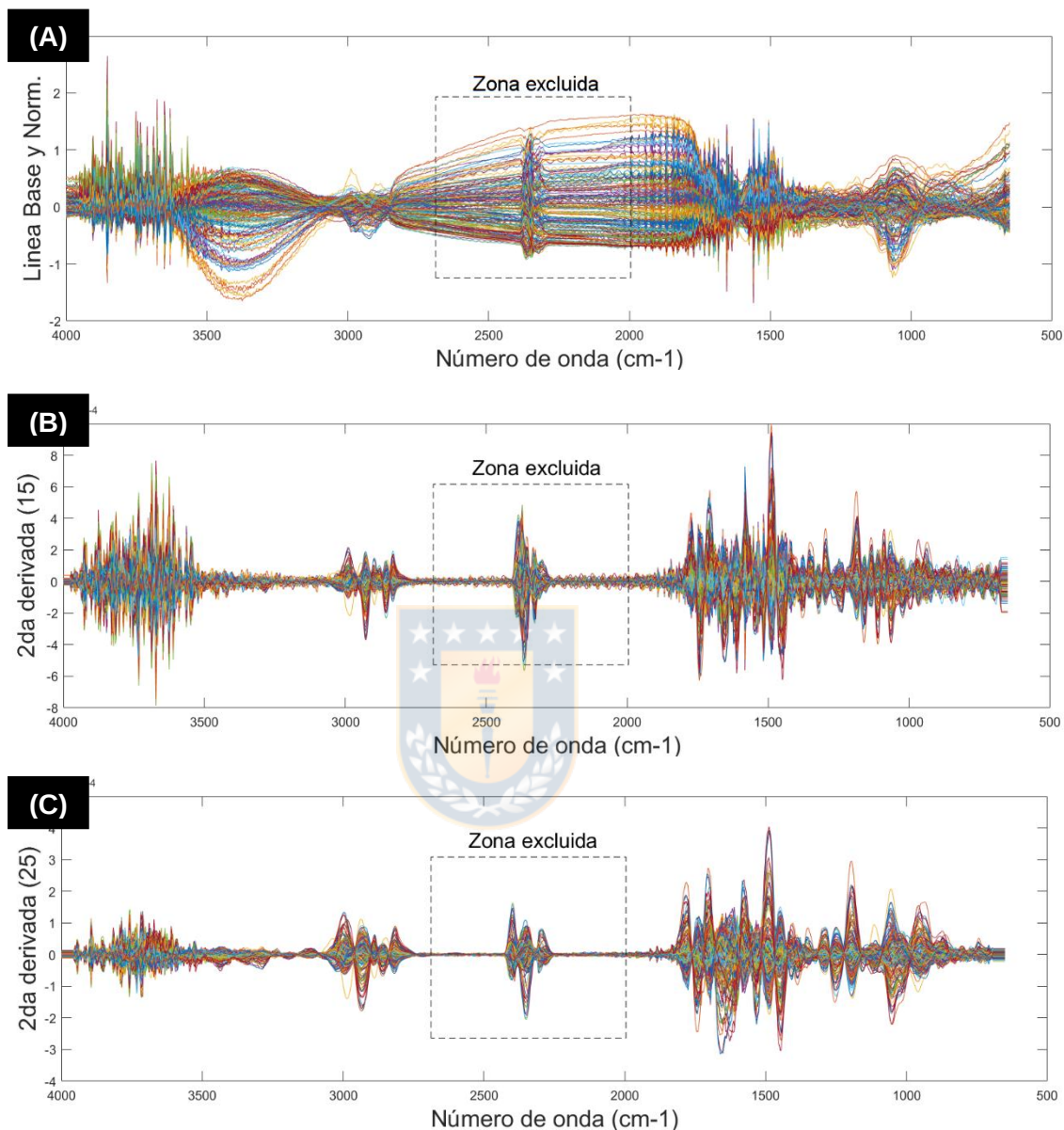


Figura 27. Espectros en región FTIR pre-tratados para (A) PCA con corrección de línea base, normalizado, (B) K-NN con 2^{da} derivada de 45 puntos y (C) SIMCA con 2^{da} derivada de 25 puntos. Todos fueron centrados en la media.

3 muestras outliers fueron identificadas mediante la distancia de Mahalanobis y el residuo de las muestras, excluyéndose de los análisis. Según el porcentaje de varianza explicada y acumulada (Tabla 6), se aprecia que los 3 primeros PCs

logran explicar un 81.6% de la varianza. En la Figura 28 se presentan los gráficos de puntuación de los 3 primeros PCs y sus respectivos gráficos de cargas. Se aprecia que el PC2 (12.7%) es el PC que mejor explica una tendencia de separación entre muestras andinas y costeras.

Tabla 6. Componentes principales calculados para matriz espectral FTIR.

Componente principal	Varianza explicada (%)	Varianza acumulada (%)
PC1	59.69	59.69
PC2	12.69	72.39
PC3	9.25	81.64
PC4	5.57	87.21
PC5	4.14	91.35

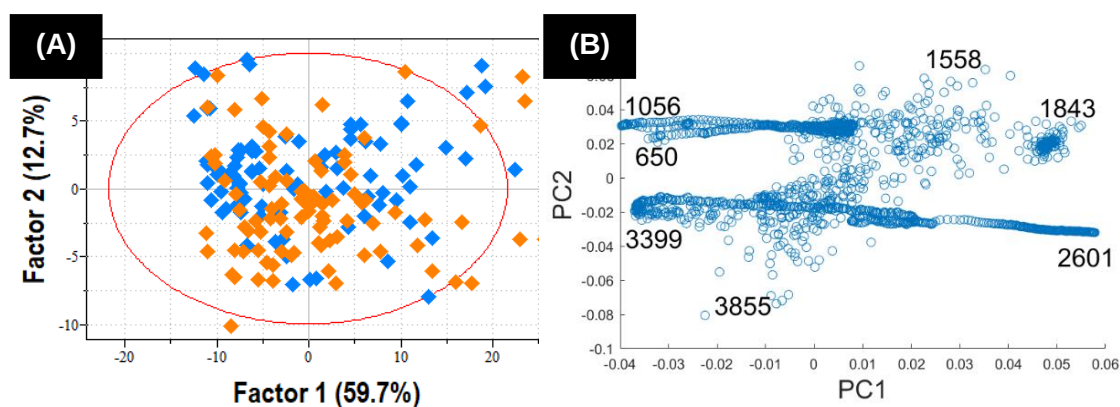


Figura 28. Gráfico de puntuaciones entre los primeros 3 PCs y sus respectivos gráficos de cargas. Cuadros azules y naranja corresponden a muestras de la costa y andes respectivamente.

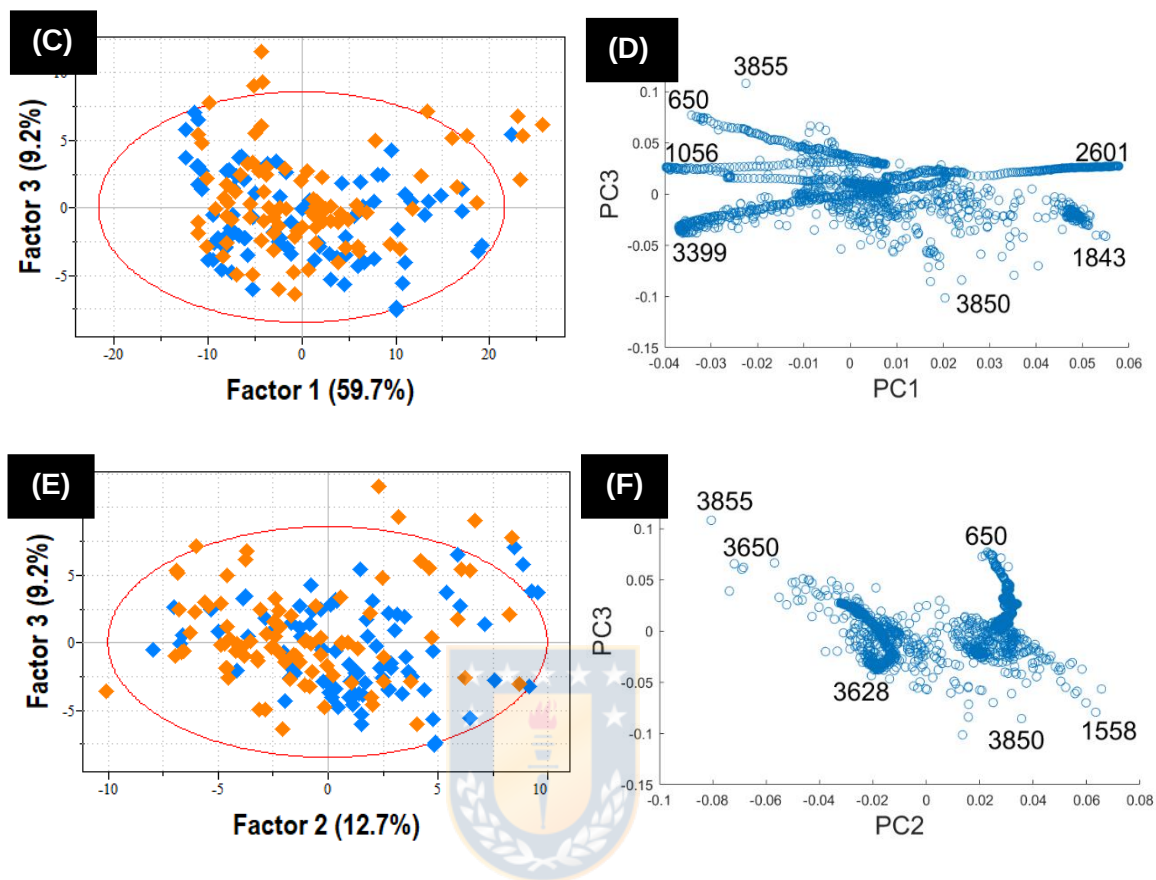


Figura 29. Gráfico de puntuaciones entre los primeros 3 PCs y sus respectivos gráficos de cargas. Cuadros azules y naranja corresponden a muestras de la costa y andes respectivamente.

Se analizó el gráfico de cargas el PC2 por separado, con la finalidad de identificar las bandas que más aportan a dicho componente (Figura 29). A grandes rasgos las longitudes entre $4000 - 3380 \text{ cm}^{-1}$ (principalmente por debajo del cero) influyen a las muestras pertenecientes a CS2 y el rango $2000 - 650 \text{ cm}^{-1}$ a CS1. Al analizar ambos rangos por separado, en el rango mayor energía se

observa que las longitudes 3897, 3850, 3813, 3742, 3731, 3709, 3686, 3668, 3645 y 3627 cm^{-1} influyen a la CS1 (Figura 29B). Por otro lado, CS2 se ve influenciada por las longitudes 3906, 3856, 3822, 3753, 3737, 3712, 3692, 3677, 3651, 3630, 3621, 3589 y 3569 cm^{-1} . Dichas longitudes se asocian a estiramientos simétricos de O-H de ácidos o metanol atribuibles a todas las fibras vegetales. No obstante, en el rango 3639 – 3029 cm^{-1} se observan estiramientos simétricos de O-H y N-H de proteínas (Liu et al., 2013; Sablinskas et al., 2003). Al observar la Figura 29C, es posible identificar que las bandas en 1735, 1701, 1655 y 1560 cm^{-1} influyen a CS2, donde la banda 1735 cm^{-1} se asocia al estiramiento de C=O de ésteres saturados de lípidos y ácidos grasos (Farghi et al., 2019; Liu et al., 2013; Malek et al., 2014) y las bandas en 1655 y 1560 cm^{-1} se asocian a proteínas (Amida I y Amida II respectivamente). En cambio, la CS1 se encuentra principalmente influenciada por bandas en el rango 2000-650 cm^{-1} , donde las bandas con mayor poder de discriminación por bibliografía se atribuyen a proteínas (absorben principalmente Amida I y II) (Beekes et al., 2007; Liu et al., 2013; Sablinskas et al., 2003).

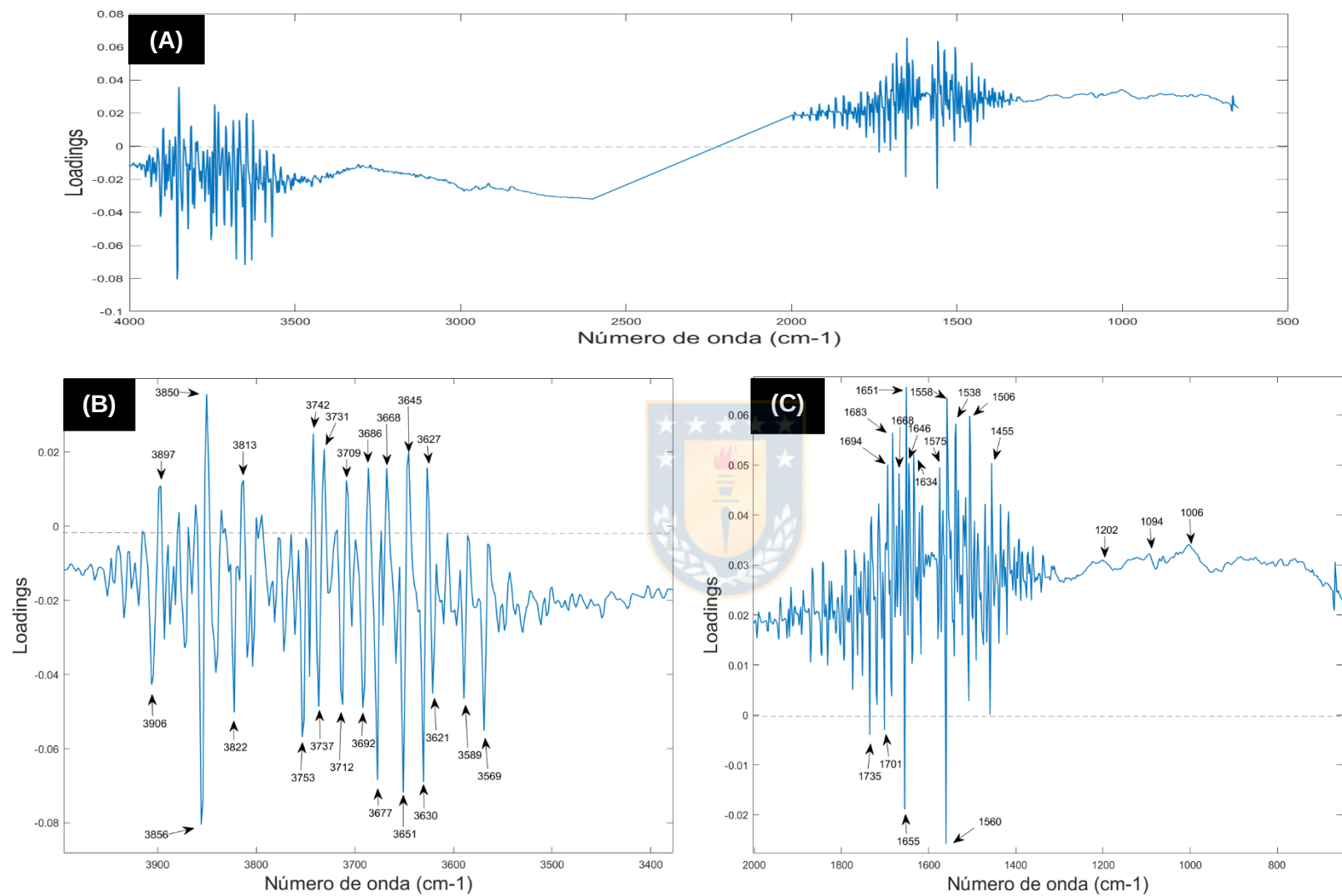


Figura 30. Gráfico de cargas de PC2 (FTIR). (A) Gráfico rango 4000 - 650 cm^{-1} . (B) Gráfico rango 4000 - 3380 cm^{-1} . (C) Gráfico rango 2000 - 650 cm^{-1} .

Otra forma de visualizar e identificar diferencias entre orígenes filogeográficos se realizó mediante el análisis de los promedios espectrales de cada clase con sus respectivas desviaciones estándar (Figura 30). Las bandas observadas en el rango $4000 - 3400 \text{ cm}^{-1}$ (Figura 30A), como se mencionó anteriormente se atribuyen a estiramientos simétricos de O-H y N-H asociados fibras vegetales y proteínas. Estas bandas presentan mayor absorbancia en las muestras andinas que en las costeras. Lo anterior podría deberse a que las plántulas andinas necesitan mayor endurecimiento dadas las mayores condiciones adversas en las que se desarrollan. En la región espectral $2000 - 600 \text{ cm}^{-1}$ (Figura 30B) se observan 8 bandas cuyos máximos de absorción se encuentran asociados a una clase, como es en el caso de las bandas en $1556, 1550, 1543, 1362, 829$ y 769 cm^{-1} , las cuales presentan mayor absorbancia en muestras andinas. Por su parte las bandas en $1556, 1550$ y 1543 cm^{-1} se atribuyen a proteínas, siendo señales de estiramientos de C=N y N-H (Beratto et al., 2017). Cercano a 1360 cm^{-1} se observan bandas de Amina terciaria aromática, estiramiento C-N por lo que también se encuentra asociado a proteínas (Nandiyanto et al., 2019). En cuando a la banda observada en 829 cm^{-1} , en literatura se reporta que la banda en 831 cm^{-1} corresponde a estiramientos asimétricos de O-P-O de ADN. Las diferencias observadas en cuanto a ácidos nucleicos sugiere diferencias en las conformaciones del ADN (Lin et al., 2011). Finalmente, en la longitud 769 cm^{-1} se observan vibraciones de movimiento conocido como “wagging” de C-H. Por otro

lado, para CS1 solo se visualiza una banda influyente en 1423 cm^{-1} , asignada a flexión asimétrica de C-H de proteínas.

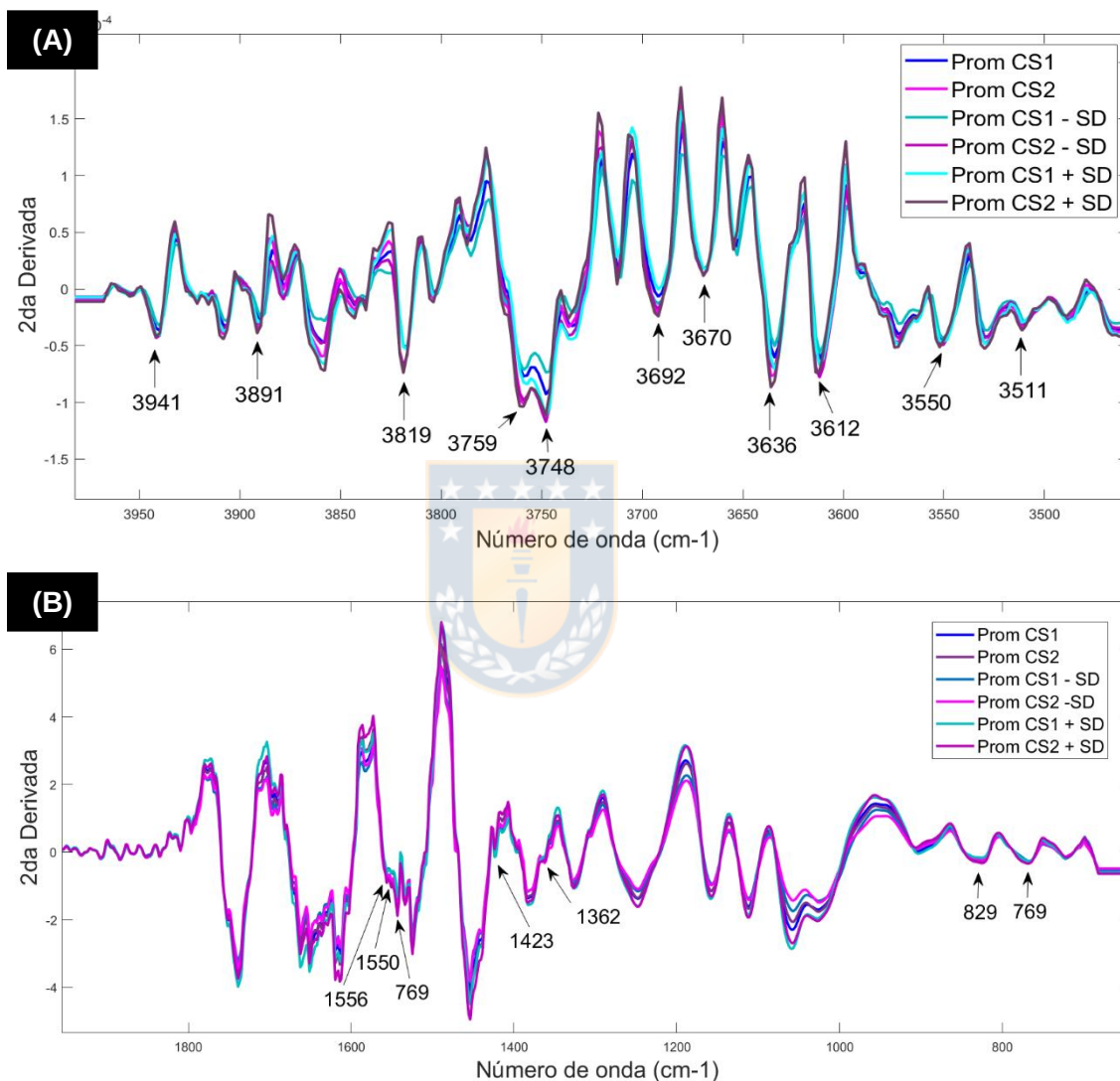


Figura 31. Segunda derivada del promedio de espectros FTIR de cada clase. (A) Rango $4000 - 3400\text{ cm}^{-1}$ y (B) Rango $2000 - 600\text{ cm}^{-1}$. Prom CS1 y Prom CS2: Promedio espectros clase 1 y clase 2, respectivamente. - SD: Promedio menos desviación estándar. + SD: Promedio más desviación estándar.

Continuando con los SPRM, como resultado de la clasificación de K-NN (K=8) con el set de entrenamiento se obtuvo un 77.24% de clasificación exitosa y en la validación un 79.25% (Para ver resultados de clasificación revisar Anexo 1). Por su parte, para SIMCA, modelado con 5 PCs para cada clase, con una distancia interclase de 0.39, se obtuvo un 93.4% de clasificación correcta con el set de entrenamiento (Para ver resultados de clasificación revisar Anexo 2). En cuanto a la validación del modelo con el set externo, se obtuvo un 90.6% de clasificación exitosa. En la Figura 31 se observan los respectivos gráficos de Coomans, en el cual se observa que la mayoría de las muestras se encuentran por debajo de ambos umbrales.

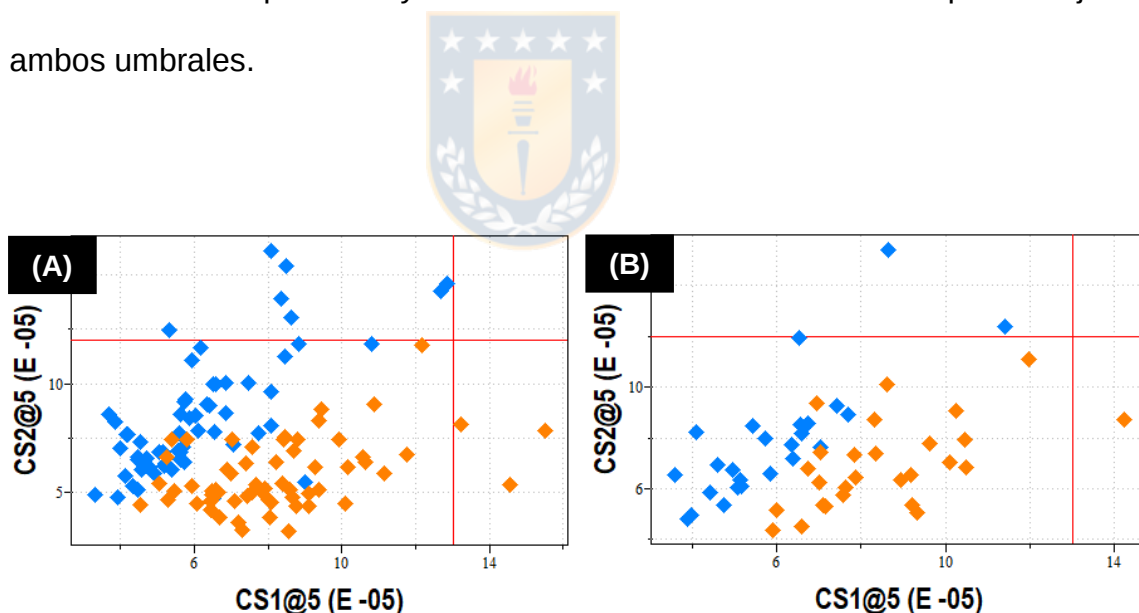


Figura 32. Gráfico de Coomans de modelo SIMCA FTIR, umbrales para un 95% de confianza. (A) set de entrenamiento y (B) set de validación. Cuadros azules corresponden a muestras de la Cordillera de la Costa (CS1) y cuadros naranjos a la Cordillera de los Andes (CS2).

En la Figura 32A se observa la gráfica de poder de discriminación de las variables en el modelo SIMCA generado. Se aprecia que hay múltiples bandas que presentan gran poder de discriminación entre orígenes filogeográficos de plántulas de *A. araucana*. En las Figuras 32B y 32C se señalan las longitudes con mayor poder de discriminación (solo se mencionan aquellas longitudes con poder de discriminación mayores a 2.5) cuya asignación química por bibliografía se presenta en la Tabla 7. Se aprecia que las bandas que más destacan son 3034, 2870, 1703, 1585, 1414, 1347 y 1293 cm^{-1} . Stenberg et al (2010) señala que aproximadamente a los 3030 cm^{-1} se observan estiramientos C-H aromáticos. La longitud 2870 cm^{-1} se ha asignado a estiramientos simétricos de CH_3 principalmente de lípidos y proteínas con pequeñas contribuciones de carbohidratos y ácidos nucleicos. Las bandas entre 1705 - 1690 cm^{-1} se han asociado a estiramientos C=O de bases de ácidos nucleicos (Malek et al., 2014). En literatura a 1581 cm^{-1} se observan las vibraciones de estiramiento de C=C de anillos aromáticos de ligninas. Por su parte, las flexiones O-H de ácidos carboxílicos absorben en 1414 cm^{-1} (Severcan et al., 2003). La banda en 1347 cm^{-1} se atribuye a deformaciones simétricas de C-H₂ y CH_3 , asociadas en literatura a celulosas, hemicelulosas y aceites (Lopes et al., 2018). Finalmente en cuanto a la banda en 1293 cm^{-1} , según lo indicado por Beekes et al (2007) entre 1310 - 1240 cm^{-1} se observa la banda de Amida III de proteínas.

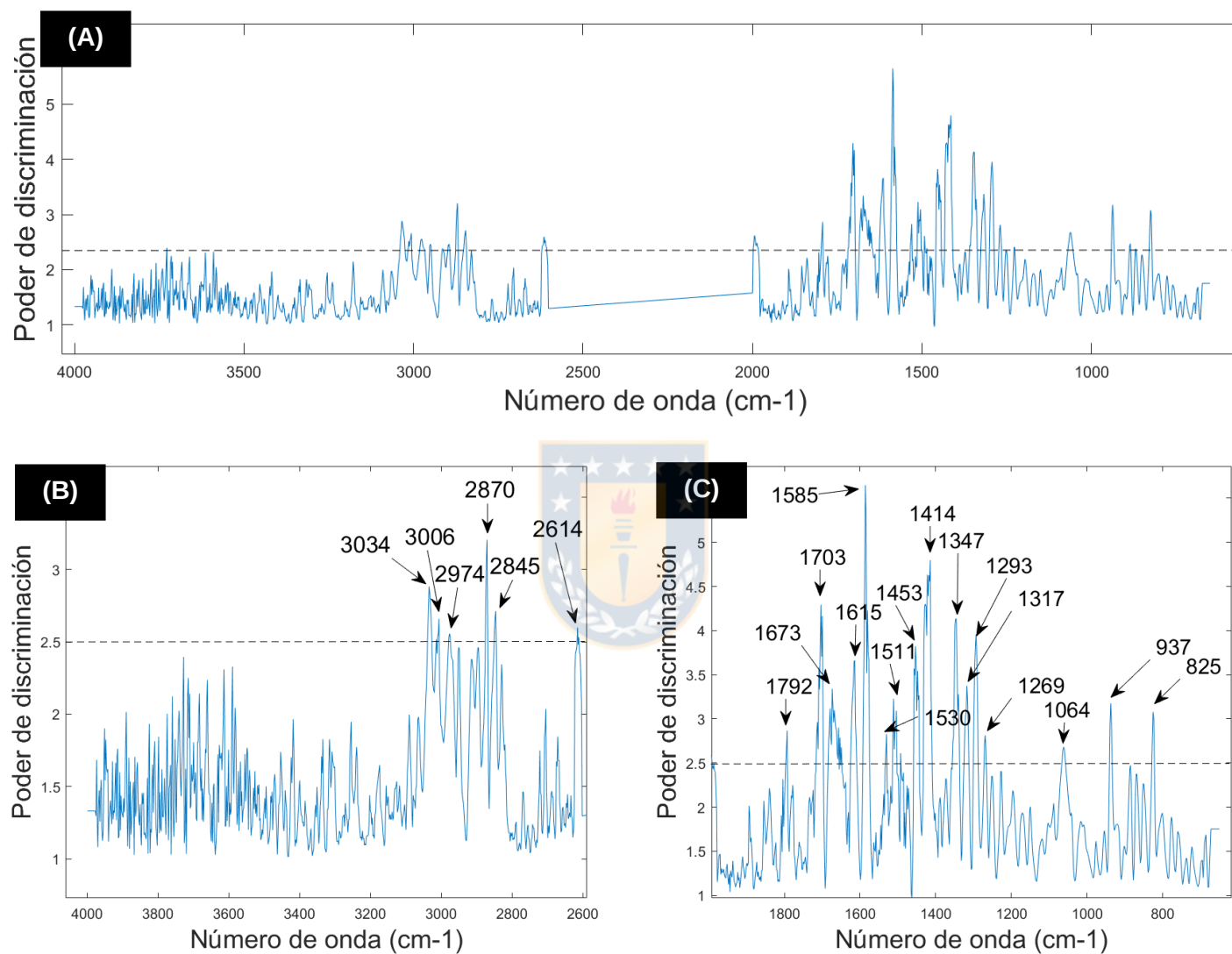


Figura 33. Poder de discriminación de variables (longitudes de onda, cm^{-1}) en modelo SIMCA en rango FTIR. Solo se señalan longitudes con mayor poder de discriminación. (A) Rango completo, (B) 4000 - 2600 cm^{-1} y (C) 2000 - 657 cm^{-1} .

Tabla 7. Asignación de bandas que presentan diferencias encontradas en la caracterización espectroscópica y análisis multivariante de las muestras. sym: simétrico, asym: asimétrico, str: estiramiento (stretching), ben: flexión (bending).

λ (cm ⁻¹) investigación	λ (cm ⁻¹) literatura	Asignación	Grupo funcional	Referencia
3034,3006	3034,3006	C-H (cis) sym str	Lípidos (enlace doble cis-oleofínicos =CH), ácidos grasos insaturados	(Irudayaraj et al., 2001; Poiana et al., 2015)
2974		CH ₂ asym str	Principalmente lípidos con pequeñas contribuciones de proteínas, carbohidratos y ácidos nucleicos	(Muthuselvi et al., 2018)
2870	2870	CH ₃ sym str	Lípidos y proteínas	(Malek et al., 2014)
2845	2860-2840	CH ₂ sym str	Principalmente lípidos con pequeñas contribuciones de proteínas, carbohidratos y ácidos nucleicos	(Lazar et al., 2012; Severcan et al., 2003; Tüker-Kaya, 2017)
1792	1792	C=O sym str		(Sclavons et al., 2005)
1703	1705-1690	C=O str	Bases de ácidos nucleicos	(Malek et al., 2014)
1673	1670-1660	C=O str + NH ben (Amida I)	Proteínas	(Malek et al., 2014)
1615	1610	N-H ben (Amida I)	Proteínas	(Stenberg et al., 2010)
1585	1581	C=C str aromáticos	Lignina	(Janardhana et al., 2013; Prégent et al., 2014)
1530	1560-1540	C=N y N-H str (Amida III)	Proteínas	(Tüker-Kaya, 2017)
1511	1514	Vibración esqueleto aromático (C=C)	Lignina, guaiacilo> siringilo	(Lopes et al., 2018)
1453,1414, 1347	1450,1418,1347	O-CH str y C-C-H str o O-H str de grupo C-OH	Carbohidratos	(Kedzierska-Matysek, Monika et al., 2018)
1317	1317	Amida III	Proteínas	(Gupta et al., 2015)
1293	1310-1240	Amida III	Proteínas	(Beekes et al., 2007)
1269	1268	Anillo guaiacilo, C-O str y C-O de metoxilos aromáticos	Lignina	(Pandey & Pitman, 2003)
1064,937	1200-900	Vibraciones de C-O-C y C-O en anillos	Principalmente carbohidratos	(Beekes et al., 2007)
825	825	PO ₂ ⁻ Frecuencia vibratoria de azúcar-fosfato	Ácidos nucleicos (Ribosa-fosfodiéster)	(Pereira et al., 2003; Tajmir-Riahi & Theophanides, 1985)

En retrospectiva, con la información adquirida en IR medio mediante Transformada de Fourier, se logró observar una tendencia de separación entre las muestras de ambos orígenes filogeográficos con el PC2 (explica el 12.92% de la varianza) atribuible a diferencias en el contenido de lípidos y proteínas. Por otro lado, los modelos supervisados, K-NN y SIMCA, lograron una precisión de clasificación utilizando un set externo de un 79 y 91% respectivamente. Por su parte, las variables con mayor poder de discriminación entre clases se atribuyen a lípidos, carbohidratos y proteínas.



5. Rango espectral con mayor poder de discriminación

Como se logró apreciar, el análisis y tratamiento de la información colectada en los distintos rangos espectrales permitió captar diferencias en bandas específicas en función del origen filogeográfico de las muestras. El análisis PCA en los cuatro rangos espectrales analizados no se observó una clara separación de grupos, pero si, una tendencia de separación entre plántulas de origen costero y andino, atribuible a pequeñas diferencias en su composición química principalmente en el contenido de lípidos y proteínas. El porcentaje de varianza explicado por los componentes que explican dicha tendencia es mayor en el caso de VIS-NIR (PC1=51.6%) y FT-NIR (PC2=25.9%), que el porcentaje de varianza explicado

obtenido mediante las herramientas moleculares SSRs en Martín et al (2014) del 16%.

En cuanto a los SPRM desarrollados, estos presentaron diferencias en cuanto a la precisión de clasificación y, por ende, diferencias en los parámetros de tasa de error, sensibilidad y especificidad. En la Tabla 8 se presenta un resumen de los parámetros obtenidos por cada modelo desarrollado. Se observa que los modelos SIMCA presentan mayores porcentajes de precisión de clasificación. Lo anterior puede deberse a que a diferencia de K-NN, SIMCA implica producir modelos estadísticos siendo un tipo de modelo probabilístico, el cual calcula un modelo PCA para cada clase y luego calcula sus límites comúnmente con una probabilidad del 95%. Por su parte K-NN es un modelo mucho más simple que no requiere de cálculos estadísticos elaborados (Brereton, 2003; Dago et al., 2008). A su vez, se logra identificar que los rangos que presentan los mejores resultados de clasificación SIMCA son VIS-NIR y FTIR, con %CA de 90 y 91% respectivamente. Por su parte, la espectroscopia VIS-NIR permitió desarrollar modelos K-NN y SIMCA capaces de predecir el origen filogeográfico de plántulas de *A. araucaria* con una precisión de clasificación del 90.5%. El %CA obtenido resultó ser menor al reportado en la identificación de fuentes semilleras de *P. sylvestris* a través de modelación SIMCA y PLS-DA, donde se lograron obtener precisiones de clasificación del 100% (Tigabu et al., 2005). Por otro lado, Atkinson

et al (1997) mediante análisis discriminante (DA) de información espectral VIS-NIR lograron distinguir exitosamente entre dos especies de *Betula* (*B. pendula* y *B. pubescens*) y sus híbridos logrando una clasificación correcta entre el 95 - 100%, porcentaje él dependía de la función discriminante utilizada. Las diferencias de absorbancias observadas en el rango VIS-NIR de plántulas de *A. araucana* indican que las muestras de ambos orígenes difieren en cuanto a su composición química específicamente en el contenido de pigmentos fotosintéticos (Richardson et al., 2004). Lo anterior concuerda con lo reportado por Gong et al (1997), quienes a través de espectroscopia en el VIS lograron discriminar entre distintas especies de coníferas debido a variaciones en su contenido de pigmentos. Las bandas más discriminantes en este rango se encuentran entre los 641 y 713 nm, es decir, en la región del rojo. Por su parte Danusevičius et al (2014) determinó que resultan ser efectivamente las bandas en la región del rojo las más informativas para la discriminación de árboles durante su crecimiento activo, indicando que factores como el origen de la fuente semillera podrían influenciar notablemente el contenido de estos pigmentos. En cuanto a la espectroscopia FTIR ha sido ampliamente utilizada para el análisis e investigación de relaciones genéticas de distintas especies (Kim et al., 2004). Por ejemplo ha permitido la determinación de relaciones filogenéticas entre poblaciones aisladas de *U. elongata* (Jianguo et al., 2012). Así mismo, ha permitido distinguir entre distintas especies de café (*Coffea canephora* var. robusta y *C. arabica*) (Kemsley et al., 1995). Al ser VIS-NIR e IR medio los rangos

más discriminantes, se corrobora lo descrito por Wang & Yu (2015) quienes indican que tras la aplicación de espectroscopia en dichos rangos es posible lograr distinguir entre especies y orígenes geográficos.

En cuanto a los parámetros de Sn y Sp, se aprecia que en la mayoría de los modelos la clase costera presenta valores de Sn más bajos en relación con la clase andina, y por ende este último presentan menores valores de especificidad, principalmente al observar los valores en los modelos SIMCA VIS-NIR y FTIR, Sn presenta valores de 0.8 para clase costera. Lo anterior se traduce en que varias muestras costeras no son reconocidas en su clase y si en la clase andina, afectando la especificidad de esta última clase. La baja sensibilidad que presenta la clase costera se puede atribuir a la gran diferencia genética que se ha evidenciado entre las poblaciones del Parque Nacional Nahuelbuta (PNN) y Villa las Araucarias (VL) (Bekessy et al., 2002; Martín et al., 2014). Bekessy et al (2002) mediante análisis de marcadores genéticos RAPDs (Random Amplified Polymorphic DNA) determinaron que la población de *A. araucana* de VL era significativamente diferente genéticamente a la población del PNN. Además, indican que justamente esta población tan diferente en términos genéticos también lo es en términos ecológicos, desarrollándose en suelos graníticos/metamórficos y a menores altitudes que las demás poblaciones analizadas (Nahuelbuta y poblaciones andinas). Estos resultados abalan la hipótesis de divergencia racial reportada por Delmastro & Donoso (1980) y

sugiriendo que dicha diferencia se podría deber a adaptaciones locales. No obstante, en bibliografía se señala que ambas poblaciones de la costa (VL y PNN) son diferentes significativamente, con un 9,24% de varianza explicada. Ahora bien, la baja Sp que presentó la clase andina se presume tiene relación con el amplio rango ecológico que cubre dicha clase, la cual está conformada por 10 localidades o poblaciones de *A. araucana* en este estudio. De esta forma, con muestras originarias de un amplio rango ecológico es de esperar que se cuente con un rango más amplio de variaciones genéticas y, por ende, de un rango más amplio de respuestas químicas. Por otro lado, estudios han revelado que la población PNN es más similar genéticamente a poblaciones Andinas que a VL. A través de análisis de aloenzimas de *A. araucana* se logró observar que las poblaciones de baja elevación de los andes (PNV, LON y PNC) eran significativamente más cercanas a las poblaciones costeras, sugiriendo que dicho fenómeno sería resultado de dispersiones de larga distancia, por araucanos o roedores (Ruiz et al., 2007). De esta forma, los autores hacen la suposición de que las poblaciones de baja elevación en los andes habrían sido colonizadas por refugios costeros después del UMG. Es así como toma sentido que las muestras pertenecientes a la Cordillera de la Costa tiendan a ser clasificadas en la clase Andina, y por consecuencia, causando que dicha clase presente niveles de Sp bajos.

Tabla 8. Evaluación de modelos supervisados desarrollados en con las 4 matrices espectrales analizadas. Sn: sensibilidad, Sp: Especificidad, CA: precisión de clasificación y ER: tasa de error.

Matriz	Clase	K-NN				SIMCA			
		Sn	Sp	CA	ER	Sn	Sp	CA	ER
VIS-NIR	Costa	1.00	0.85	0.90	0.10	0.81	0.96	0.90	0.10
	Andes	0.85	1.00	0.90	0.10	0.96	0.81	0.90	0.10
	<i>Promedio CA</i>			0.90				0.90	
	<i>Promedio ER</i>				0.10				0.10
FT-NIR	Costa	1.00	0.74	0.86	0.14	0.83	0.93	0.88	0.12
	Andes	0.74	1.00	0.86	0.14	0.93	0.83	0.88	0.12
	<i>Promedio CA</i>			0.86				0.88	
	<i>Promedio ER</i>				0.14				0.12
NIR	Costa	0.85	0.74	0.85	0.15	0.96	0.85	0.89	0.11
	Andes	0.85	0.85	0.85	0.15	0.85	0.96	0.89	0.11
	<i>Promedio CA</i>			0.85				0.89	
	<i>Promedio ER</i>				0.15				0.11
FTIR	Costa	0.69	0.89	0.79	0.21	0.81	1.00	0.91	0.09
	Andes	0.89	0.69	0.79	0.21	1.00	0.81	0.91	0.09
	<i>Promedio CA</i>			0.79				0.91	
	<i>Promedio ER</i>				0.21				0.09

Ahora bien, ya identificados los modelos con mejores resultados (SIMCA VIS-NIR y FTIR) se procedió a evaluar la robustez, cuyos resultados se observan en la Tabla 9. El parámetro de robustez entrega información sobre la susceptibilidad del modelo quimiométrico a ciertas variaciones, en este caso las muestras del set de entrenamiento y validación. Una buena robustez entrega confiabilidad en sus resultados, permitiendo discernir si se está frente a un modelo inestable y poco fiable (Amigo, 2007; Moncayo, 2017). Para esta investigación una buena robustez da credibilidad a las predicciones de clases. Para VIS-NIR se observa que los modelos K-NN validados presentan una variación en la precisión de

clasificación del 7.17%, cuyo máximo porcentaje es de un 90.5%. En cuanto a SIMCA para este mismo rango el %CA se encontró en un rango de 85.71-90.5%, es decir una variación del 4.79%. Para FTIR, K-NN presento %CA en la validación en un rango de 71.69-81.13%, es decir una variación de 9.44% y por último SIMCA presentó %CA en un rango de 88.68-90.60, variando solo un 1.92%. De esta forma los resultados demuestran que es factible desarrollar modelos robustos de reconocimiento de patrones para la clasificación según origen filogeográfico de plántulas de *A. araucana* a partir de sus propiedades químicas. Se determinó que los modelos SIMCA presentaron la menor variación para ambos rangos espectrales, siendo el modelo en el rango FTIR el de mayor robustez con solo un 1.92% de variación en %CA. En este sentido los 4 modelos desarrollados presentan una buena robustez, siendo los modelos SIMCA los más robustos, por ende, los modelos más confiables, cuyos resultados no se ven altamente influenciados por la selección de las muestras para los sets de entrenamiento y validación.

Tabla 9. Evaluación de robustez de modelos K-NN y SIMCA en rango VIS-NIR con equipo portátil y FTIR, según %CA.

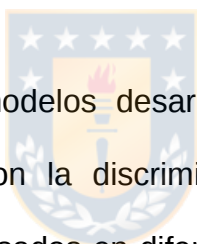
Matriz espectral	%CA K-NN		%CA SIMCA	
	Set entrenamiento	Set validación	Set entrenamiento	Set validación
VIS-NIR 1	88.89	90.50	98.99	90.50
VIS-NIR 2	82.83	85.71	98.99	85.71
VIS-NIR 3	89.89	83.33	98.99	85.71
FTIR 1	77.24	79.25	93.40	90.60
FTIR 2	78.86	71.69	92.68	88.68
FTIR 3	75.60	81.13	91.87	90.57

6. Discriminación de plántulas de *A. araucana* según origen filogeográfico

La variación química en la composición de las hojas de *A. araucana* entre ambos orígenes filogeográficos podría ser resultado de las variaciones genéticas que se han ido acumulando durante la historia evolutiva de la especie. Sork et al (1994) indican que variaciones en caracteres químicos entre subpoblaciones de *Quercus rubra* L. tienen más probabilidades de deberse a diferenciación genética que simplemente al efecto materno del sitio. Las variaciones genéticas recaen en la variación de rasgos para la adaptación local, lo cual sería resultado de la evolución de las poblaciones a través de la deriva genética y de la selección natural. Lo anterior, adquiere mayor relevancia al tratarse de especies que presentan una dispersión limitada y restringida de sus semillas, y que a su vez lidian con fuertes presiones selectivas en sus hábitats (Premoli et al., 2007). Se ha visto que el órgano de las plantas más sensible a los factores ambientales resultan ser las hojas, siendo el órgano más idóneo para estudiar la filogeografía de las especies vegetales (Zhang et al., 2008). Ciertos autores señalan que variaciones en las condiciones de suelo, humedad, nutrientes o interacciones competitivas podrían ser responsables de dichas variaciones intraespecíficas (Sork et al., 1994; Zhang et al., 2008). Por ejemplo, para la especie *Nothofagus pumilio* (Poepp. & Endl.) Krasser se tiene conocimiento de que el tipo de suelo en el que habita influye directamente en el contenido lignina de sus hojas, aumentando su contenido en suelos pobres (Premoli & Mathiasen, 2011). Por

otro lado, Jianguo et al (2012) sugieren que factores climáticos son los principales responsables de la composición química de las hojas, como la temperatura que es un factor determinante y esencial para la síntesis de celulosa y almidón. A su vez, otro factor del que se tiene evidencia de su influencia en la diferenciación de composición química de hojas es la latitud. En la especie *Pistacia chinensis* Bunge se observó que a mayor latitud menor era su contenido de ácidos grasos específicamente de ácido oleico, esteárico y linoleico (Hou et al., 2010). Las técnicas como marcadores genéticos isoenzimas han permitido documentar diferencias genéticas interpoblacionales debido a variaciones en latitudes y elevaciones (Sork et al., 1994). El estudio de aloenzimas reportado para las poblaciones chilenas de *A. araucana* permitió identificar una alta variabilidad genética, explicado principalmente por la variabilidad intrapoblacional y en menor proporción por la diversidad genética entre poblaciones (Ruiz et al., 2007). Por su parte, Donoso (1987) planteó que las variaciones intraespecíficas entre poblaciones de distintas latitudes tendrían un fuerte componente genético además del ambiental. Lo anterior, por el efecto de las discontinuidades geográficas las cuales funcionan como barreras para el flujo de genes, o a su vez por el aislamiento en distintos refugios glaciares (Premoli et al., 2012). Ambas situaciones son las que presenta *A. araucana* en Chile. Por un lado, el aislamiento geográfico entre sus dos núcleos poblacionales (Cordillera de los Andes y Cordillera de la Costa) generado por la depresión intermedia (Villagrán & Armesto, 2005) y la existencia de refugios glaciares en ambos cordones

montañosos en el UMG. Dichas condiciones y eventos son los responsables de la estructura genética actual conocida para *A. araucana*, así como para otras coníferas del centro-sur de Chile como *Fitzroya cupressoides* (Molina) I.M. Johnst (Allnutt et al., 1999). Los datos filogeográficos reportados para las poblaciones de estas coníferas indican altos valores de diversidad genética intra e interpoblacional. Específicamente para la especie declarada monumento natural, su mayor variación genética se da dentro de las poblaciones, lo que concuerda con la gran nube de dispersión observada en los distintos PCA generados.



Ahora bien, los mejores modelos desarrollados (correspondientes al rango VIS-NIR y FTIR) permitieron la discriminación de origen filogeográfico de plántulas de *A. araucana* basados en diferencias en los máximos de absorción en distintas longitudes de onda asociadas a distintos grupos funcionales o macromoléculas. Por su parte, SIMCA VIS-NIR permitió identificar que ambas clases se diferencian principalmente en el contenido de pigmentos fotosintéticos (Chla, Chlb y carotenoides (xantófilas)). Presumiblemente, las plántulas de la Cordillera de los Andes presentaron mayor contenido de compuestos carotenoides (535 nm) los cuales otorgan coloración amarilla-anaranjada a las hojas. Dentro de los carotenoides se encuentran aquellos que no contienen un oxígeno en sus anillos terminales y otros que sí, conocidos estos últimos como xantófilas (Carranco et al., 2011). Las xantófilas absorben en el rango de 400 -

530 nm y actúan como protectores frente a excesos lumínicos, disipando la energía en forma de calor mediante el ciclo de las xantófilas (Varela, 2005). De esta forma se podría atribuir que el mayor contenido de estos pigmentos carotenoides se encuentra en las poblaciones andinas dado a que estas se desarrollan a mayores alturas en comparación a VL y PNN. Esto quiere decir que en las poblaciones andinas se encuentran sometidos a mayores estreses lumínicos siendo necesario contar con un mayor contenido de estos pigmentos fotoprotectores. Por si fuera poco, en Rivas (2020) se menciona que mayores concentraciones de carotenoides contribuyen a una mayor capacidad antioxidante, de esta forma la planta está más preparada para tolerar un estrés abiótico. En este sentido cabe mencionar que las plantas de *A. araucaria* de la Cordillera de los Andes se encuentran bajo mayores estreses abióticos, como ya se mencionó se encuentra bajo estrés lumínico donde las altas intensidades lumínicas aumentan la probabilidad de la formación de especies reactivas de oxígeno (ROS), pero también se encuentran bajo estrés hídrico. Lo anterior dado a que esta especie pasa gran parte del año sin agua disponible al estar ésta en forma de nieve, requiriendo así mayor protección frente a las ROS. Sumado a lo anterior, según Aherne et al (2009) el contenido de pigmentos carotenoides está controlado por atributos tanto fisiológicos, genéticos y bioquímicos e incluso podría estar influenciado por la ubicación geográfica. Por otro lado, en cuanto a clorofila a y b, las plántulas de origen costero presentaría un mayor contenido de Chlb en comparación con las de origen andino. Para la Chla, pigmento que

presenta más bandas de discriminación asociadas, hay bandas que contribuyen a muestras andinas (<690nm) y otras a las muestras costeras (>700nm). La Chla resulta ser el pigmento más importante y principal de la fotosíntesis, a diferencia de la Chlb que se considera un pigmento accesorio (se considera parte del complejo antena) (Jaramillo-Salazar et al., 2018). Se conoce que frente a distintas condiciones ambientales las plantas responden ya sea morfológica o fisiológicamente modificando entre ellos su contenido de pigmentos. De esta forma, por lo general cuando las plantas están sometidas bajo estrés (hídrico, salino, por patógenos, entre otros) tienden a disminuir el contenido de clorofila en sus hojas, privilegiando el resistir frente a crecer. Otro factor que se ha descrito en literatura que afecta o contribuye a la variación de contenido de clorofila foliar es el contenido de nitrógeno (N). El N forma parte de las moléculas de clorofila existiendo así una estrecha correlación positiva entre sus concentraciones (Sánchez et al., 2018). Teniendo en consideración que los suelos de la Cordillera de los Andes y de la Costa son muy distintos desde el punto de vista nutricional es de esperar que presentaran variaciones en cuanto al contenido de clorofila foliar. Los suelos de la Cordillera de la costa se caracterizan por presentar una escasa profundidad y por ser muy pobres en nutrientes (entre ellos N) (Bekessy et al., 2004). Es por lo anterior que se presume que las plántulas de origen andino presentaron mayor número de bandas discriminadoras relacionadas a estos pigmentos, presentando mayor contenido de estos. Resumiendo, en literatura se ha reportado que el contenido de clorofilas varía según el desbalance nutricional,

las condiciones ambientales a la cual se ven expuestas las plantas, tales como temperatura, sequía, radiación, salinidad, entre otros (A. González et al., 2014).

Continuando con el modelo SIMCA FTIR, este permitió discriminar entre ambos orígenes filogeográficos gracias a diferencias químicas asociadas principalmente al contenido de lípidos, proteínas, carbohidratos. Específicamente, las plántulas de origen costero presentaron mayor absorbancia de bandas atribuidas a carbohidratos y proteínas, presentando presumiblemente mayor contenido de estas macromoléculas. Los principales carbohidratos o polisacáridos que componen las plantas son la celulosa, hemicelulosas, ligninas y almidón (Poisa et al., 2011). Billow et al (1994) mencionan que frente a la limitación de nutrientes las plantas tienden a presentar mayores concentraciones de carbohidratos en comparación a aquellas que crecen bajo una alta disponibilidad de nutrientes. Por otro lado, un mayor contenido de celulosa podría relacionarse con el factor temperatura, debido a que las temperaturas en la Cordillera de la Costa no son tan extremas como en la Cordillera de los Andes por lo que tienden a ser más cálidas, favoreciendo así la síntesis de estos polisacáridos (Jianguo et al., 2012). En cuanto a la lignina, en literatura se ha descrito que frente a limitantes nutricionales las plantas suelen presentar contenidos más altos de lignina y compuestos de defensa como terpenos y compuestos fenólicos (Billow et al., 1994). De hecho, en la especie *N. pumilio* se evidenció que aquellos individuos

desarrollados en suelos pobres presentaban mayor contenido de lignina en sus hojas (Premoli & Mathiasen, 2011). Según Campbell & Sederoff (1996) el contenido de lignina es una variable que tiende a presentar variaciones del 25 y 30% entre poblaciones de una misma especie, esto se ha reportado para *P. monticola* y *P. palustris*. Lo anterior concuerda con que la banda asociada a lignina resultara útil como variable discriminadora de origen filogeográfico. Continuando con las proteínas, en literatura se ha reportado un gran número de proteínas que participan en los mecanismos de defensas de las plantas frente a factores bióticos (Blanco-Labra & Aguirre Mancilla, 2002). Estas proteínas de defensa pueden actuar como factores constitutivos de defensa o ser parte de la cascada de respuesta. En cuanto a las poblaciones de ambos orígenes filogeográficos de *A. araucaria* se tiene conocimiento de que se diferencian, entre otros factores, por la composición de especies en sus bosques (Delmastro & Donoso, 1980). Así como se diferencian entre especies arbóreas como distintas especies de *Nothofagus*, también la biota de fúngica es diferente, encontrándose ambas cordilleras expuestas a distintos patógenos. Cabe mencionar que justamente las pequeñas poblaciones de la Cordillera de la Costa son las que presentan el peor estado de conservación de todo el país (Wolodarsky-Franke & Díaz, 2011), pudiendo ser poblaciones más susceptibles y que requieren de mayor contenido de compuestos de defensa. A pesar de la gran relevancia de los compuestos de defensa asociados a moléculas protectoras, Moran et al (2017) señalan que estos rasgos han sido ignorados en estudios de

procedencias. Por su parte, las plántulas de origen andino presentaron mayor absorbancia en las longitudes asociada lípidos y ácidos grasos con una pequeña contribución de proteínas. Estos mismos compuestos resultaron útiles para la diferenciación de procedencias y fuentes semilleras de *P. abies* y *P. sylvestris* (Farhadi et al., 2017; Tigabu et al., 2005). Continuando, el contenido de lípidos y ácidos grasos de las plantas es afectado por factores ambientales, entre ellos el suelo y el clima, pero en cierta medida también se ven influenciados genéticamente. Sumado a ello, Tillman-Sutela et al (1995) reportan que el contenido total de lípidos, esto en semillas de *P. sylvestris*, tiende a aumentar con la latitud. Por el contrario, Hou et al (2010) evidenciaron una relación inversa entre la latitud y el contenido de ácidos grasos presentes en las hojas de *P. chinensis*. Otras investigaciones reportan que los ácidos linoleico, palmítico y oleico presentan un alto poder de discriminación geográfica (Uncu & Ozen, 2016). En cuanto a la pequeña contribución de proteínas en las bandas asociadas (no son las mismas bandas influyentes en la clase costera), podrían atribuirse a la presencia de proteínas de choque térmico (HSPs). La Cordillera de los Andes se caracteriza por fenómenos de sequías y temperaturas extremas, por lo que los individuos que ahí se desarrollan deben aclimatarse para que estas condiciones no sean letales. Dentro de los procesos de aclimatación se encuentra la síntesis de proteínas en respuesta a estas temperaturas extremas las cuales confieren termotolerancia. Al enfrentarse a estas condiciones algunas plantas tienden a disminuir la síntesis de proteínas normales, privilegiando la síntesis de HSPs

como mecanismo de defensa celular (Yepes & Buckeridge, 2011). Lo anterior es una especulación, una posible explicación de la influencia de proteínas en plántulas de origen andino, pero no se registran investigaciones sobre HSPs en *A. araucana*.

Finalmente, los espectros no son más que la expresión fenotípica de señales relacionadas con la composición química y estructural de las hojas, resultantes de la variación de factores ambientales y principalmente de la variación genética derivada de la historia evolutiva o filogeografía de la especie en cuestión. Es así como las plántulas de *A. araucana* provenientes de distintos orígenes filogeográficos o UES han desarrollado diferentes mecanismos funcionales para la adquisición y aprovechamiento de sus recursos, diferencias que se reflejan en las propiedades espectrales adquiridas. En otras palabras, se sugiere que las plántulas han adaptado su funcionamiento fisiológico a las condiciones de su sitio de origen. Por lo que es preciso enfatizar en la importancia de conocer el origen filogeográfico de los materiales de propagación a utilizar con fines de restauración.

V CONCLUSIÓN

El uso de la espectroscopia IR y métodos quimiométricos mostraron excelentes resultados para la diferenciación de plántulas de *Araucaria araucana* según su origen filogeográfico. Los análisis exploratorios presentaron tendencias de separación entre muestras de la Cordillera de la Costa y Cordillera de los Andes, presentando mayores porcentajes de varianza explicada (51.6% en VIS-NIR) que los reportados mediante herramientas moleculares (16% por SSRs). Los modelos supervisados SIMCA en los distintos rangos presentaron los mayores porcentajes de predicción (88 – 90%) en comparación con K-NN. En este sentido, los resultados muestran que es factible diferenciar entre orígenes filogeográficos en todos los rangos espectrales analizados y que, a su vez, es posible utilizar herramientas espectrales portátiles para evaluaciones en terreno. Sin embargo, los mejores resultados se obtuvieron en los rangos VIS-NIR y FTIR, logrando porcentajes de predicción del 90% en modelos predictivos validados con sets externos. Ahora bien, para las muestras mal clasificadas (10% restante) se podría realizar un análisis más profundo, analizando la georeferenciación de las muestras y sus condiciones ambientales, volver a analizar la misma muestra (partiendo desde el preprocesamiento) por posibles errores experimentales o

incluso aplicar herramientas moleculares, dado que aquellas muestras podrían ser híbridos. Con respecto a los compuestos químicos responsables de la discriminación entre orígenes *A. araucana*, se identificó que las muestras costeras presentan mayor contenido de pigmentos carotenoides (xantófilas), Chlb, carbohidratos y proteínas y menor contenido de Chla, lípidos y ácidos grasos respecto a las muestras de origen andino. Estas variaciones químicas entre UES son el resultado de las variaciones ambientales y genéticas derivadas de la historia evolutiva de *A. araucana*. Mediante esta investigación se evidencia que las diferencias reportadas a nivel genético se expresan a nivel químico.



En conclusión, los espectros en la región VIS-NIR y FTIR obtenidos desde plántulas de *Araucaria araucana*, tratados a través de métodos quimiométricos, permitieron captar la señal filogeográfica que separa las procedencias de la Cordillera de la Costa con respecto a la Cordillera de los Andes. Se sugiere la aplicación de estas herramientas como alternativas a los métodos convencionales (moleculares, morfológicos, entre otros) para la diferenciación de origen filogeográfico de *A. araucana*. Esto, tras los excelentes resultados obtenidos en esta investigación y considerando las ventajas de estas herramientas, tales como los tiempos de análisis, el carácter mínimamente destructivo, la no utilización de químicos, la facilidad de uso y bajo costo.

Finalmente, se debe considerar que los modelos resultantes de esta investigación podrían ser mejorados tras el aumento del número de muestras mejorando así los valores de robustez y sensibilidad de los modelos, los cuales podrían ser utilizados en futuros programas de restauración de *A. araucana* para rastrear o autenticar el origen del germoplasma a utilizar, reduciendo los riesgos asociados a trabajar con material foráneo.



VI REFERENCIAS

- Aherne, A., Jiwan, M., Daly, T., & O'Brien, N. (2009). Geographical Location has Greater Impact on Carotenoid Content and Bioaccessibility from Tomatoes than Variety. *Plant foods for human nutrition (Dordrecht, Netherlands)*, 64, 250–256. <https://doi.org/10.1007/s11130-009-0136-x>
- Ali, M., Emsley, A. M., Herman, H., & Heywood, R. J. (2001). Spectroscopic studies of the ageing of cellulosic paper. *Polymer*, 42(7), 2893–2900. [https://doi.org/10.1016/S0032-3861\(00\)00691-1](https://doi.org/10.1016/S0032-3861(00)00691-1)
- Alía, R., Alba, N., Agúndez, D., & Iglesias, S. (2005). *Manual para la comercialización y producción de semillas y plantas forestales. Materiales de base y de reproducción* (Organismo Autónomo Parques Nacionales. Ministerio de Medio Ambiente). Serie Forestal.
- Allnutt, T. R., Newton, A. C., Lara, A., Premoli, A., Armesto, J. J., Vergara, R., & Gardner, M. (1999). Genetic variation in *Fitzroya cupressoides* (alerce), a threatened South American conifer. *Molecular Ecology*, 8, 975–987. <https://doi.org/10.1046/j.1365-294x.1999.00650.x>
- Amigo, J. (2007). *Desarrollo y Aplicación de Métodos Quimiométricos Multidimensionales al Estudio de Sistemas Enzimáticos* [Tesis Doctoral]. Universidad Autónoma de Barcelona.
- Armesto, J., Donoso, C., & Villagrán, C. (1994). Desde la era glacial a la industria: La historia del bosque templado chileno. *Ambiente y Desarrollo*.
- Azcon-Bieto, J., I, F., Aranda, X., & Gomez-Casanovas, N. (2008). Fotosíntesis, factores ambientales y cambio climático. En *Fundamentos de Fisiología Vegetal*.
- Bannister, J., González, M., Little, C., Gutiérrez, A., Donoso, P., Mujica, R., Müller-Using, S., Lara, A., Bustamante Sánchez, M., Bannister, A., Caracciolo, A., Echeverría, J., Suárez, J., & Zambrano, C. (2013). Experiencias de restauración en los bosques nativos del sur de Chile: Una mirada desde la Isla Grande de Chiloé. *Bosque Nativo*, 52, 35–43.

- Beekes, M., Lasch, P., & Naumann, D. (2007). Analytical applications of Fourier transform-infrared (FT-IR) spectroscopy in microbiology and prion research. *Recent Progress in Prion Research*, 123(4), 305–319. <https://doi.org/10.1016/j.vetmic.2007.04.010>
- Bekessy, S. A., Allnutt, T., Premoli, A., Lara, A., Ennos, R., Burgman, M., Cortes, M., & Newton, A. (2002). Genetic variation in the vulnerable and endemic Monkey Puzzle tree, detected using RAPDs. *Heredity*, 88, 243–249. <https://doi.org/10.1038/sj.hdy.6800033>
- Bekessy, S. A., Lara, A., González, M., Cortes, M., Gallo, L., Premoli, A. C., & Newton, A. C. (2004). Variación en *Araucaria araucana* (Molina) K. Koch (Araucaria o Pehuén). En *Variación Intraespecífica en las especies arbóreas de los bosques templados de Chile y Argentina* (Claudio Donoso, Andrea Premoli, Leonardo Gallo y Roberto Ipinza, p. 434). Editorial Universitaria.
- Beratto, A., Agurto, C., Freer, J., Peña-Farfal, C., Troncoso, N., Agurto, A., & Castillo, R. del P. (2017). Chemical Characterization and Determination of the Anti-Oxidant Capacity of Two Brown Algae with Respect to Sampling Season and Morphological Structures Using Infrared Spectroscopy and Multivariate Analyses. *Applied Spectroscopy*, 71(10), 2263–2277. <https://doi.org/10.1177/0003702817715654>
- Bhamondez, A. (2016). *Bosque nativo, comunidades y cambio climático: Avances en el Diseño y Preparación de la Estrategia Nacional de Cambio Climático y Recursos Vegetacionales de Chile (ENCCRV), y propuestas para su implementación.*
- Billow, C., A., M., & Bond, B. (1994). Seasonal biochemical changes in coniferous forest canopies and their response to fertilization. *Tree physiology*, 14, 563–574. <https://doi.org/10.1093/treephys/14.6.563>
- Blanco, M., & Villarroja, I. (2002). NIR spectroscopy: A rapid-response analytical tool. *TrAC Trends in Analytical Chemistry*, 21(4), 240–250. [https://doi.org/10.1016/S0165-9936\(02\)00404-1](https://doi.org/10.1016/S0165-9936(02)00404-1)
- Blanco-Labra, A., & Aguirre Mancilla, C. (2002). Proteínas Involucradas en los Mecanismos de Defensa de Plantas. *Acta Universitaria*, 12(3), 3–28. Redalyc.
- Brereton, R. (2003). Pattern Recognition. En *Chemometrics* (pp. 183–269). John Wiley & Sons, Ltd. <https://doi.org/10.1002/0470863242.ch4>

- Broadhurst, L. M., Lowe, A., Coates, D. J., Cunningham, S. A., McDonald, M., Vesk, P. A., & Yates, C. (2008). Seed supply for broadscale restoration: Maximizing evolutionary potential. *Evolutionary Applications*, 1(4), 587–597. PubMed. <https://doi.org/10.1111/j.1752-4571.2008.00045.x>
- Budić-Leto, I., Gajdoš Kljusurić, J., Goran, Z., Humar, I., Banović, M., Kurtanjek, Z., & Lovrić, T. (2011). Usefulness of near infrared spectroscopy and chemometrics in screening of the quality of dessert wine Prošek. *Croatian Journal of Food Science and Technology*, 3, 9–15.
- Campbell, M. M., & Sederoff, R. R. (1996). Variation in Lignin Content and Composition (Mechanisms of Control and Implications for the Genetic Improvement of Plants). *Plant Physiology*, 110(1), 3–13. PubMed. <https://doi.org/10.1104/pp.110.1.3>
- Carranco, M., Calvo, M., & Pérez-Gil, F. (2011). Carotenoids and their antioxidant function: A review. *Archivos latinoamericanos de nutrición*, 61, 233–241.
- Caruso, G., Broglia, V., & Pocoví, M. (2015). Diversidad genética. Importancia y aplicaciones en el mejoramiento vegetal. *Lhawet*, 4(1).
- Chalmers, J. M., Edwards, H. G. M., & Hargreaves, M. D. (2012). Vibrational Spectroscopy Techniques: Basics and Instrumentation. En *Infrared and Raman Spectroscopy in Forensic Science* (pp. 9–44). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119962328.ch2>
- Dago, A., Morales, Osorio, A., Fernández, R., & Dennes, E. (2008). Desarrollo de un modelo SIMCA para la clasificación de kerosinas mediante el empleo de la espectroscopía infrarroja. *Química Nova*, 31. <https://doi.org/10.1590/S0100-40422008000600049>
- Danezis, G., Tsagkaris, A., Brusic, V., & Georgiou, C. (2016). Food Authentication: State of the Art and Prospects. *Current Opinion in Food Science*, 10. <https://doi.org/10.1016/j.cofs.2016.07.003>
- Danusevičius, D., Masaitis, G., & Mozgeris, G. (2014). Visible and near infrared hyperspectral imaging reveals significant differences in needle reflectance among Scots pine provenances. En *Silvae genetica. Warsaw: De Gruyter, 2014, Vol. 63, iss. 4. Žemės ūkio akademija*. <https://doi.org/10.1515/sg-2014-0022>
- Datt, B. (1998). Remote Sensing of Chlorophyll a, Chlorophyll b, Chlorophyll a+b, and Total Carotenoid Content in Eucalyptus Leaves. *Remote Sensing of Environment*, 66(2), 111–121. [https://doi.org/10.1016/S0034-4257\(98\)00046-7](https://doi.org/10.1016/S0034-4257(98)00046-7)

- Delmastro, R., & Donoso, C. (1980). Review of distribution, variation and utilization of gene resources of *Araucaria araucana* (Mol.) Koch in Chile. *Simposio. IUFRO en mejoramiento genético e productividad de espécies florestais de rápido crescimento*, Brasil.
- Donoso, C. (1987). Variación Natural en Especies de *Nothofagus* en Chile. *Bosque*, 8(2), 85–97. <https://doi.org/10.4206/bosque.1987.v8n2-03>
- Donoso, C., Premoli, A., Gallo, L., & Ipinza, R. (2004). *Variación Intraespecífica En Las Especies Arbóreas De Los Bosques Templados De Chile Y Argentina*.
- Dorado, J., Almendros, G., Field, J., & Sierra-Alvarez, R. (2001). Infrared spectroscopy analysis of hemp (*Cannabis sativa*) after selective delignification by *Bjerkandera* sp. At different nitrogen levels. *Enzyme and microbial technology*, 28, 550–559. [https://doi.org/10.1016/S0141-0229\(00\)00363-X](https://doi.org/10.1016/S0141-0229(00)00363-X)
- Drake, F., Martín, A., Machuca, M., Martínez, J., Drake-Martin, F., & Martin, L. (2009). Networking sampling of *Araucaria araucana* (Mol.) K. Koch in Chile and the bordering zone of Argentina: Implications for the genetic resources and the sustainable management. *iForest : Biogeosciences and Forestry*, 2. <https://doi.org/10.3832/ifor0524-002>
- Dufour, É. (2009). Chapter 1—Principles of Infrared Spectroscopy. En D.-W. Sun (Ed.), *Infrared Spectroscopy for Food Quality Analysis and Control* (pp. 1–27). Academic Press. <https://doi.org/10.1016/B978-0-12-374136-3.00001-8>
- Farghl, A., El-Sheekh, M., & Mousa, H. (2019). *Extraction and characterization of antimicrobial active substance from cyanobacteria Nostoc carneum and Anabaena circinalis*.
- Farhadi, M., Tigabu, M., Pietrzykowski, M., Danusevičius, D., & Odén, P. C. (2017). Application of near infrared spectroscopy for authentication of *Picea abies* seed provenance. *New Forests*, 48(5), 629–642. <https://doi.org/10.1007/s11056-017-9589-1>
- Fernández, I., Morales San Martin, N., Olivares, L., Salvatierra, J., Gómez, M., & Montenegro, G. (2010). *Restauración Ecológica para Ecosistemas Nativos Afectados por Incendios Forestales*. <https://doi.org/10.13140/RG.2.1.3191.0887>
- Foley, W., Mcilwee, A., Lawler, I., Aragones, L., Woolnough, A., & Berding, N. (1998). Ecological applications of near infrared reflectance spectroscopy—

A tool for rapid, cost-effective prediction of the composition of plant and animal tissues and aspects of animal performance. *Oecologia*, 116, 293–305. <https://doi.org/10.1007/s004420050591>

- G. Campos, J., Gutiérrez Caro, B., Brand, M., Ipinza, R., & Quilodrán, F. (2016). Herramientas tecnológicas complementarias para la fiscalización de tala ilegal y delitos forestales contra los bosques naturales. *CIFOR*, 22.
- Garrido, A., García-Olmo, J., & De-Pedro Sanz, E. (2000). Espectroscopía de Infrarrojo Cercano (NIRS). Una metodología para implementar en sistemas de aseguramiento de la calidad y trazabilidad de productos derivados del cerdo ibérico. *Solo Cerdo Ibérico*, 4, 93–102.
- Godbout, J., Bomal, C., Farr, K., Williamson, M., & Isabel, N. (2018). Genomic tools for traceability: Opportunities, challenges and perspectives for the Canadian forestry sector. *The Forestry Chronicle*, 94(01), 75–87. <https://doi.org/10.5558/tfc2018-010>
- Gong, P., Pu, R., & Yu, B. (1997). Conifer species recognition: An exploratory analysis of in situ hyperspectral data. *Remote Sensing of Environment*, 62(2), 189–200. [https://doi.org/10.1016/S0034-4257\(97\)00094-1](https://doi.org/10.1016/S0034-4257(97)00094-1)
- González, A., Briceño, H., Chirinos, J., Buonocore, R., & Villarreal, A. (2014). Variación de la concentración de clorofila a, b, clorofila total y tasa de fotosíntesis en *Avicennia germinans* en el manglar de Punta de Palmas, municipio Miranda, estado Zulia, Venezuela. *Revista Investigaciones Científicas UNERMB*, 5, 67/82.
- González, M., & Montaña, L. (2015). La espectroscopia y su tecnología: Un repaso histórico y su importancia para el siglo XXI. *Latin-American Journal of Physics Education*, 9, 13.
- Gupta, B. S., Jelle, B. P., & Gao, T. (2015). Application of ATR-FTIR Spectroscopy to Compare the Cell Materials of Wood Decay Fungi with Wood Mould Fungi. *International Journal of Spectroscopy*, 2015, 521938. <https://doi.org/10.1155/2015/521938>
- Hallett, R., Hornbeck, J., & Martin, M. (1997). Predicting elements in white pine and red oak foliage with near infrared reflectance spectroscopy. *Journal of Near Infrared Spectroscopy - J NEAR INFRARED SPECTROSC*, 5. <https://doi.org/10.1255/jnirs.101>
- Hancock, N., & Hughes, L. (2014). Turning up the heat on the provenance debate: Testing the 'local is best' paradigm under heatwave conditions. *Austral Ecology*, 39(5), 600–611. <https://doi.org/10.1111/aec.12122>

- Hoffer, R. (1978). Biological and physical considerations in applying computer-aided analysis techniques to remote sensor data. En *Remote Sensing: The Quantitative Approach* (P. H Swain, S. M. Davis, pp. 55–96). McGraw-Hill.
- Hou, X. C., Mou, H. X., & Jian, Y. Z. (2010). Geographical regularity of oil and fatty acid content of bioenergy plant *Pistacia chinensis* Bunge. *Ecology and Environmental Sciences*, 19, 2773–2777.
- IBM Knowledge Center. (2021, marzo 15). *Overview (KNN command)*. https://www.ibm.com/support/knowledgecenter/SSLVMB_sub/statistics_reference_project_ddita/spss/base/syn_knn_overview.html
- INFOR. (2009). *Elaboración proyectos forestación / reforestación bajo mecanismo de desarrollo limpio en Chile. A/R MDL Protocolo de Kyoto* (Manual N° 176).
- Invernizzi, C., Rovetta, T., Licchelli, M., & Malagodi, M. (2018). Mid and Near-Infrared Reflection Spectral Database of Natural Organic Materials in the Cultural Heritage Field. *International Journal of Analytical Chemistry*, 2018, 7823248. <https://doi.org/10.1155/2018/7823248>
- Irudayaraj, J., Sivakesava, S., Kamath, S., & Yang, H. (2001). Monitoring Chemical Changes in Some Foods Using Fourier Transform Photoacoustic Spectroscopy. *Journal of Food Science*, 66, 1416–1421. <https://doi.org/10.1111/j.1365-2621.2001.tb15224.x>
- Janardhana, K., Ravindrachary, V., P C, Dr. R. K., Umesh, G., K B, Dr. M., & Ismayil. (2013). Third order optical nonlinearity of a novel chalcone derivative through Z scan technique. *Indian Journal of Pure and Applied Physics*, 51, 844–850.
- Jaramillo-Salazar, M., Marin-Giraldo, Y., & Ocampo-Serna, D. (2018). Efectos en el nivel fotosintético en tres especies de plantas acuáticas sometidas a un tratamiento con agua residual de origen minero. *Boletín Científico. Centro de Museos. Museo de Historia Natural*, 22(1), 43–57.
- Jaya, S., Mani, S., Kandala, C., & Holser, R. (2015). Application of NIR Reflectance Spectroscopy on Rapid Determination of Moisture Content of Wood Pellets. *American Journal of Analytical Chemistry*, 06, 923–932. <https://doi.org/10.4236/ajac.2015.612088>
- Jianguo, G., Yuhuan, W., Gendi, X., Wenqiao, L., Guohao, Y., Jun, M., & Liu, P. (2012). Phylogeography of *Ulmus elongata* based on Fourier transform-infrared spectroscopy (FTIR), thermal gravimetric and differential thermal

analyses. *Biochemical Systematics and Ecology*.
<https://doi.org/10.1016/j.bse.2011.10.018>

Kačuráková, M., Capek, P., Sasinková, V., Wellner, N., & Ebringerová, A. (2000). FT-IR study of plant cell wall model compounds: Pectic polysaccharides and hemicelluloses. *Carbohydrate Polymers*, 43(2), 195–203. [https://doi.org/10.1016/S0144-8617\(00\)00151-X](https://doi.org/10.1016/S0144-8617(00)00151-X)

Kedzierska-Matysek, Monika, Matwijczuk, Arkadiusz, Florek, Mariusz, Barłowska, Joanna, Wolanciuk, Anna, Matwijczuk, Alicja, Chruł'sciel, Edyta, Walkowiak, Radosław, Karcz, Dariusz, & Gladyszewska, Bożena. (2018). Application of FTIR spectroscopy for analysis of the quality of honey. *BIO Web Conf.*, 10, 02008. <https://doi.org/10.1051/bioconf/20181002008>

Keller, M., Kollmann, J., & Edwards, P. J. (2000). Genetic introgression from distant provenances reduces fitness in local weed populations. *Journal of Applied Ecology*, 37(4), 647–659. <https://doi.org/10.1046/j.1365-2664.2000.00517.x>

Kemsley, E. K., Ruault, S., & Wilson, R. H. (1995). Discrimination between *Coffea arabica* and *Coffea canephora* variant robusta beans using infrared spectroscopy. *Food Chemistry*, 54(3), 321–326. [https://doi.org/10.1016/0308-8146\(95\)00030-M](https://doi.org/10.1016/0308-8146(95)00030-M)

Kim, S., Ban, S., Cho, S., Chung, H., Choi, P. S., Yoo, O., & Liu, J. (2004). Taxonomic discrimination of flowering plants by multivariate analysis of Fourier transform infrared spectroscopy data. *Plant cell reports*, 23, 246–250. <https://doi.org/10.1007/s00299-004-0811-1>

Lazar, G., Ureche, D., Ifrim, I., Stamate, M., Ureche, C., Nedeff, V., Nistor, I.-D., Finaru, A.-L., & Iuliana, L. (2012). Effects of the environmental stress on two fish populations revealed by statistical and spectral analysis. *Environmental Engineering and Management Journal*, 11, 109–124. <https://doi.org/10.30638/eemj.2012.016>

Lerma-García, M., Cortés, V., Talens, P., & Barat, J. (2018). Variety Discrimination of Fruits, Edible Plants, and Other Foodstuffs and Beverages by Infrared Spectroscopy. En *Comprehensive Analytical Chemistry* (Vol. 80, pp. 127–163). <https://doi.org/10.1016/bs.coac.2018.03.004>

Li, Y., Sun, Y., Jiang, J., & Liu, J. (2019). Spectroscopic determination of leaf chlorophyll content and color for genetic selection on *Sassafras tzumu*. *Plant Methods*, 15. <https://doi.org/10.1186/s13007-019-0458-0>

- Lin, D., Lin, J., Wu, Y., Feng, S., Li, Y., Yu, Y., Xi, G., Zeng, H., & Chen, R. (2011). Investigation on the interactions of lymphoma cells with paclitaxel by Raman spectroscopy. *Journal of Spectroscopy*, 25. <https://doi.org/10.3233/SPE-2011-0497>
- Lin, M., A Rasco, B., G Cavinato, Anna, & Al-Holy, M. (2009). Chapter 6—Infrared (IR) Spectroscopy—Near-Infrared Spectroscopy and Mid-Infrared Spectroscopy. En D.-W. Sun (Ed.), *Infrared Spectroscopy for Food Quality Analysis and Control* (pp. 119–143). Academic Press. <https://doi.org/10.1016/B978-0-12-374136-3.00006-7>
- Liu, J., Mukherjee, J., Hawkes, J., & Wilkinson, S. (2013). Optimization of lipid production for algal biodiesel in nitrogen stressed cells of *Dunaliella salina* using FTIR analysis. *Journal of Chemical Technology & Biotechnology*, 88, 1807–1814. <https://doi.org/10.1002/jctb.4027>
- Loewe M., V., Navarro Cerrillo, R., García-Olmo, J., Riccioli, C., & Sánchez-Cuesta, R. (2016). Discriminant analysis of Mediterranean pine nuts (*Pinus pinea* L.) from Chilean plantations by Near Infrared Spectroscopy. *Food Control*. <https://doi.org/10.1016/j.foodcont.2016.09.012>
- Lohumi, S., Lee, S., Lee, H., & Cho, B.-K. (2015). A review of vibrational spectroscopic techniques for the detection of food authenticity and adulteration. *Trends in Food Science & Technology*, 46(1), 85–98. <https://doi.org/10.1016/j.tifs.2015.08.003>
- Loo, J. (2011). *Manual de genética de la conservación. Principios aplicados de genética para la conservación de la diversidad biológica* (1ª ed.).
- Lopes, J. de O., Garcia, R. A., & Souza, N. D. de. (2018). Infrared spectroscopy of the surface of thermally-modified teak juvenile wood. *Maderas. Ciencia y Tecnología*, 20, 737–746.
- Lucio-Gutiérrez, J. R. (2012). *Aplicación de Métodos Quimiométricos para la Caracterización y Control de Calidad de Plantas Medicinales* [Tesis Doctoral, Universidad Autónoma de Barcelona]. <http://hdl.handle.net/10803/96257>
- Luykx, D., & Ruth, S. (2008). An overview of analytical methods for determining the geographical origin of food products. *Food Chemistry*, 107, 897–911. <https://doi.org/10.1016/j.foodchem.2007.09.038>
- Malek, K., Wood, B., & Bambery, K. (2014). FTIR Imaging of Tissues: Techniques and Methods of Analysis. En *Optical Spectroscopy and Computational*

Methods in Biology and Medicine (pp. 419–473).
https://doi.org/10.1007/978-94-007-7832-0_15

- Manley, M. (2014). Near-Infrared Spectroscopy and Hyperspectral Imaging: Non-Destructive Analysis of Biological Materials. *Chemical Society reviews*, 43. <https://doi.org/10.1039/c4cs00062e>
- Manzano-Agugliaro, F., Garrido-Cárdenas, J. A., & Jiménez-Camínaro, E. (2018). *Técnicas de investigación avanzadas en agronomía y ciencias afines*. Universidad de Almería.
- Markgraf, V., Mcglone, M., & Hope, G. (1995). Neogene Paleoenvironmental and Paleoclimatic Change In Southern Temperate Ecosystems—A Southern Perspective. *Trends in ecology & evolution*, 10, 143–147. [https://doi.org/10.1016/S0169-5347\(00\)89023-0](https://doi.org/10.1016/S0169-5347(00)89023-0)
- Martín, M. A., Mattioni, C., Lusini, I., Molina, J. R., Cherubini, M., Drake, F., Herrera, M. A., Villani, F., & Martín, L. M. (2014). New insights into the genetic structure of *Araucaria araucana* forests based on molecular and historic evidences. *Tree Genetics & Genomes*, 10(4), 839–851. <https://doi.org/10.1007/s11295-014-0725-1>
- Massart, D., Vandeginste, B. G. M., Buydens, L. M. C., De Jong, S., Lewi, P. J., & Smeyers-Verbeke, J. (1998). *Handbook of Chemometrics and Qualimetrics—Part A*. Elsevier Science Ltd.
- Matesanz, S., Ramos-Muñoz, M., Blanco, M., García-Fernández, A., Sánchez, A., & Escudero, A. (2019). Migration, genetic variation and phenotypic plasticity of gypsophiles and their role on climate change responses. *Ecosistemas*, 28, 48–59. <https://doi.org/10.7818/ECOS.1525>
- McGarigal, K., Cushman, S., & Stafford, S. (2000). *Multivariate Statistics for Wildlife Ecology Research*. <https://doi.org/10.1007/978-1-4612-1288-1>
- McKay, J. K., Christian, C. E., Harrison, S., & Rice, K. J. (2005). “How Local Is Local?”—A Review of Practical and Conceptual Issues in the Genetics of Restoration. *Restoration Ecology*, 13(3), 432–440. <https://doi.org/10.1111/j.1526-100X.2005.00058.x>
- Merzlyak, M., Solovchenko, A., & Gitelson, A. (2003). Reflectance spectral features and non-destructive estimation of chlorophyll, carotenoid and anhocyanin content in apple fruit. *Postharvest Biology and Technology*, 27, 197–211. [https://doi.org/10.1016/S0925-5214\(02\)00066-2](https://doi.org/10.1016/S0925-5214(02)00066-2)

- MMA. (2020). *Listado de Especies Clasificadas desde el 1º al 15º Proceso de Clasificación RCE (actualizado a julio de 2020)*. Clasificación de Especies. Ministerio del Medio Ambiente. <https://clasificacionespecies.mma.gob.cl/>
- Moncayo, S. (2017). *Desarrollo y aplicación de métodos quimiométricos para el estudio de muestras mediante Espectroscopia de Ablación Láser (LIBS)* [Tesis Doctoral, Universidad Complutense de Madrid]. <https://eprints.ucm.es/id/eprint/42642/1/T38776.pdf>
- Monsef, Z. (2012). Reflectance IR Spectroscopy. En *Infrared Spectroscopy—Materials Science, Engineering and Technology* (Theophanides Theophile, pp. 233–244).
- Moore, D. S., Jepsen, P. U., & Volka, K. (2014). Principles of Vibrational Spectroscopic Methods and their Application to Bioanalysis. En *Handbook of Spectroscopy* (pp. 1037–1078). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9783527654703.ch27>
- Moran, E., Lauder, J., Musser, C., Stathos, A., & Shu, M. (2017). The genetics of drought tolerance in conifers. *New Phytologist*, 216. <https://doi.org/10.1111/nph.14774>
- Muthuselvi, C., Pandiarajan, S., Ravikumar, B., Athimoolam, S., Srinivasan, N., Rajaputi, V., & Krishnakumar, V. (2018). FT-IR and FT-Raman spectroscopic analyzes of indeno quinoxaline dervative crystal. *Asian Journal of Applied Sciences*, 11, 83–91.
- Nandiyanto, A., Oktiani, R., & Ragadhita, R. (2019). How to Read and Interpret FTIR Spectroscopy of Organic Material. *Indonesian Journal of Science and Technology*, 4, 97–118. <https://doi.org/10.17509/ijost.v4i1.15806>
- Pandey, K., & Pitman, A. (2003). FTIR studies of the changes in wood chemistry following decay by brown-rot and white-rot fungi. *International Biodeterioration & Biodegradation*, 52, 151–160. [https://doi.org/10.1016/S0964-8305\(03\)00052-0](https://doi.org/10.1016/S0964-8305(03)00052-0)
- Pasquini, C. (2003). Near Infrared Spectroscopy: Fundamentals, practical aspects and analytical applications. *Journal of the Brazilian Chemical Society*, 14, 198–219.
- Pereira, L., Sousa, A., Coelho, H., Amado, A. M., & Ribeiro-Claro, P. J. A. (2003). Use of FTIR, FT-Raman and ¹³C-NMR spectroscopy for identification of some seaweed phycocolloids. *Marine Biotechnology: Basics and Applications*, 20(4), 223–228. [https://doi.org/10.1016/S1389-0344\(03\)00058-3](https://doi.org/10.1016/S1389-0344(03)00058-3)

- Poiana, M.-A., Alexa, E., Munteanu, M.-F., Gligor, R., Moigradean, D., & Mateescu, C. (2015). Use of ATR-FTIR spectroscopy to detect the changes in extra virgin olive oil by adulteration with soybean oil and high temperature heat treatment. *Open Chemistry*, 13, 689–698. <https://doi.org/10.1515/chem-2015-0110>
- Poisa, L., Adamovics, A., Platane, R., & Teirumnieka, Ē. (2011). *Evaluation of the Factors that Affect The Lignin Content in The Reed Canarygrass (Phalaris arundinacea L.) in Latvia* (p. 231). <https://doi.org/10.3384/ecp11057224>
- Prégent, J., Vandsburger, L., Blanchard, V., Blanchet, P., Riedl, B., Sarkissian, A., & Stafford, L. (2014). Modification of hardwood samples in the flowing afterglow of N₂–O₂ dielectric barrier discharges open to ambient air. *Cellulose*, 22. <https://doi.org/10.1007/s10570-014-0496-8>
- Premoli, A. C., Acosta, M. C., Mathiasen, P., & Donoso Z, C. (2012). Variación genética en *Nothofagus* (subgénero *Nothofagus*). *Bosque (Valdivia)*, 33, 115–125.
- Premoli, A., & Mathiasen, P. (2011). Respuestas ecofisiológicas adaptativas y plásticas en ambientes secos de montaña: *Nothofagus pumilio*, el árbol que acaparó los Andes australes. *Ecología austral*, 21, 251–269.
- Premoli, A., Raffaele, E., & Mathiasen, P. (2007). Morphological and phenological differences in *Nothofagus pumilio* from contrasting elevations: Evidence from a common garden. *Austral Ecology*, 32, 515–523. <https://doi.org/10.1111/j.1442-9993.2007.01720.x>
- Pustjens, A. M., Muilwijk, M., Weesepeel, Y., & Ruth, S. M. van. (2016). Advances in authenticity testing of geographical origin of food products. En *Advances in Food Authenticity Testing* (pp. 339–367). Elsevier Inc. Academic Press. <https://edepot.wur.nl/518074>
- Quiroz, I., Gutiérrez, B., & García, E. (2012). *Bases para un reglamento de semillas y plantas de especies forestales utilizadas en Chile. Proyecto 028/2010 “Fundamentos para una Normativa sobre Origen de Semillas y Calidad de Plantas Forestales Nativas” Financiado por el Fondo de Investigación del Bosque Nativo*. <https://bibliotecadigital.infor.cl/bitstream/handle/20.500.12220/20103/30636.pdf;jsessionid=E92F5721489CAC9D99E2F6AEBB68E3E2?sequence=1>
- Reusch, W. (2013). *Visible and Ultraviolet Spectroscopy*. Chemistry. <https://www2.chemistry.msu.edu/faculty/reusch/virttxtjml/spectrpy/uv-vis/spectrum.htm>

- Richardson, A. D., Reeves III, J. B., & Gregoire, T. G. (2004). Multivariate analyses of visible/near infrared (VIS/NIR) absorbance spectra reveal underlying spectral differences among dried, ground conifer needle samples from different growth environments. *New Phytologist*, 161(1), 291–301. <https://doi.org/10.1046/j.1469-8137.2003.00913.x>
- Rivas, G. (2020). *Efecto de la inoculación de hongos micorrícicos arbusculares y endófitos en la tolerancia al estrés por sequía en plántulas de Araucaria araucana* [Universidad de Concepción]. http://repositorio.udec.cl/jspui/bitstream/11594/534/1/Tesis_efecto_de_la_inoculacion.pdf
- Rodríguez, R., Matthei, O., & Quezada, M. (1983). *Flora arbórea de Chile*. Editorial de la Universidad de Concepción.
- Rodríguez, W., García, P., & Fajardo, A. (2016). Aplicaciones de técnicas espectroscópicas para el análisis de suelos. *Revista Facultad de Ciencias Básicas*, 12. <https://doi.org/10.18359/rfcb.2030>
- Rohman, A., Windarsih, A., Hossain, M. A., Johan, M. R., Ali, M. E., & Fadzilah, N. A. (2019). Application of near- and mid-infrared spectroscopy combined with chemometrics for discrimination and authentication of herbal products: A review. *Journal of Applied Pharmaceutical Science*, 9, 137–147. <https://doi.org/10.7324/JAPS.2019.90319>
- Romero-Saritama, J. M. (2018). Conservación de semillas: Una alternativa inmediata para almacenar germoplasma forestal y recuperar los bosques secos amenazados del Ecuador. *Neotropical Biology and Conservation*, 13(1), 74–85.
- Ruiz, E., González, F., Torres, C., Fuentes arce, G., Mardones, M., Stuessy, T., Samuel, R., Becerra, J., & Silva, M. (2007). Genetic Diversity and Differentiation within and among Chilean Populations of *Araucaria araucana* (Araucariaceae) Based on Allozyme Variability. *Taxon*, 56. <https://doi.org/10.2307/25065913>
- Sablinskas, V., Steiner, G., & Hof, M. (2003). Applications. En *Handbook of Spectroscopy* (pp. 89–168). John Wiley & Sons, Ltd. <https://doi.org/10.1002/3527602305.ch6>
- Sánchez, E., Ruiz, J. M., Romero, L., Preciado-Rangel, P., Flores-Córdova, M., & Márquez-Quiroz, C. (2018). ¿Son los pigmentos fotosintéticos buenos indicadores de la relación del nitrógeno, fósforo y potasio en frijol ejotero? *Ecosistemas y recursos agropecuarios*, 5, 387–398.

- Sandak, A., Sandak, J., & Negri, M. (2011). Relationship between near-infrared (NIR) spectra and the geographical provenance of timber. *Wood Science and Technology*, 45, 35–48. <https://doi.org/10.1007/s00226-010-0313-y>
- Schwanninger, M., Rodrigues, J., & Fackler, K. (2011). A Review of Band Assignments in near Infrared Spectra of Wood and Wood Components. *Journal of near Infrared Spectroscopy*, 19, 287–308. <https://doi.org/10.1255/jnirs.955>
- Sclavons, M., Laurent, M., Devaux, J., & Carlier, V. (2005). Maleic anhydride-grafted polypropylene: FTIR study of a model polymer grafted by ene-reaction. *Polymer*, 46, 8062–8067. <https://doi.org/10.1016/j.polymer.2005.06.115>
- Severcan, F., Kaptan, N., & Turan, B. (2003). Fourier Transform Infrared Spectroscopic Studies of Diabetic Rat Heart Crude Membranes. *Journal of Spectroscopy*, 17. <https://doi.org/10.1155/2003/472834>
- Sherman, H. (1997). Infrared Spectroscopy. En *Handbook of Instrumental Techniques for Analytical Chemistry* (pp. 247–283). Prentice Hall: Upper Saddle River.
- Skoog, D., Holler, J., & Crouch, S. (2008). *Principios de Análisis Instrumental* (Sexta). CengageLearning.
- Society for Ecological Restoration (SER) International, Grupo de trabajo sobre ciencia y políticas. (2004). Principios de SER International sobre la restauración ecológica. www.ser.org y Tucson: Society for Ecological Restoration
- International.Sork, V., Stowe, K., & Hochwender, C. (1994). Evidence for Local Adaptation in Closely Adjacent Subpopulations of Northern Red Oak (*Quercus rubra* L.) Expressed as Resistance to Leaf Herbivores. *The American naturalist*, 142, 928–936. <https://doi.org/10.1086/285581>
- Stenberg, B., Viscarra Rossel, R., Mouazen, A., & Wetterlind, J. (2010). Chapter Five—Visible and Near Infrared Spectroscopy in Soil Science. *Advances in Agronomy*, 107, 163–215. [https://doi.org/10.1016/S0065-2113\(10\)07005-7](https://doi.org/10.1016/S0065-2113(10)07005-7)
- Stuart, H. (2004). *Infrared Spectroscopy: Fundamentals and Applications*. John Wiley & Sons, Ltd.
- Su, W.-H., He, H.-J., & Sun, D.-W. (2017). Non-destructive and Rapid Evaluation of Staple Foods Quality by Using Spectroscopic Techniques: A Review.

- Critical reviews in food science and nutrition*, 57, 1039–1051.
<https://doi.org/10.1080/10408398.2015.1082966>
- Tajmir-Riahi, H. A., & Theophanides, T. (1985). An FT-IR Study of DNA and RNA Conformational Transitions at Low Temperatures. *Journal of Biomolecular Structure and Dynamics*, 3(3), 537–542.
<https://doi.org/10.1080/07391102.1985.10508441>
- Tigabu, M., Oden, P., & Lindgren, D. (2005). Identification of seed sources and parents of *Pinus sylvestris* L. using visible–near infrared reflectance spectra and multivariate analysis. *Trees*, 19, 468–476.
<https://doi.org/10.1007/s00468-005-0408-5>
- Tillman-Sutela, E., Johansson, A., Laakso, P., & Mattila, T. (1995). Triacylglycerols in the seeds of northern Scots pine, *Pinus sylvestris* L., and Norway spruce, *Picea abies* (L.) Karst. *Trees*, 10, 40–45.
- Townsend, A., & McIntosh, M. (1993). Variation among full-sib progenies of red maple in growth, autumn leaf color, and leafhopper injury. *Journal of Environmental Horticulture*, 11(2), 72–75.
- Tsuchikawa, S., & Kobori, H. (2015). A review of recent application of near infrared spectroscopy to wood science and technology. *Journal of Wood Science*, 61. <https://doi.org/10.1007/s10086-015-1467-x>
- Tüker-Kaya, S. (2017). A Review of Mid-Infrared and Near-Infrared Imaging: Principles, Concepts and Applications in plant Tissues Analysis. *Molecules*, 22(1), 168. <https://doi.org/10.3390/molecules22010168>
- UICN. (2020). *Monkey Puzzle. Araucaria araucana*. The IUCN Red List of Threatened Species.
<https://www.iucnredlist.org/es/species/31355/2805113>
- Uncu, O., & Ozen, B. (2016). Geographical differentiation of a monovarietal olive oil using various chemical parameters and mid-infrared spectroscopy. *Anal. Methods*, 8. <https://doi.org/10.1039/C6AY01290F>
- Varela, S. (2005). *Respuestas morfológico-fisiológicas ante distintos tipos e intensidades de estrés en dos especies patagónicas del género Nothofagus con aptitud comercial* [Tesis de Magíster, Universidad Nacional del Comahue]. <https://core.ac.uk/download/pdf/33998273.pdf>
- Vargas, O., & Sc, M. (2011). Restauración Ecológica: Biodiversidad y Conservación. Ecological Restoration: Biodiversity and Conservation. *Acta Biológica Colombiana*, 16, 221–246.

- Villagrán, C. (2018). Biogeografía de los bosques subtropical-templados del sur de sudamérica. Hipótesis Históricas. *Magallania (Punta Arenas)*, 46, 27–48.
- Villagrán, C., & Armesto, J. (2005). Fitogeografía histórica de la Cordillera de la Costa de Chile. En *Historia, biodiversidad y ecología de los bosques costeros de Chile.*, (Cecilia Smith-Ramírez, Juan J. Armesto y Claudio Valdovinos, p. 708 páginas). Editorial Universitaria. <http://repositorio.uchile.cl/handle/2250/120075>
- Vitale, R., Bevilacqua, M., Bucci, R., Magrì, A., Magrì, A., & Marini, F. (2013). A rapid and non-invasive method for authenticating the origin of pistachio samples by NIR spectroscopy and chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 121, 90–99. <https://doi.org/10.1016/j.chemolab.2012.11.019>
- Walsh, K. B., Blasco, J., Zude-Sasse, M., & Sun, X. (2020). Visible-NIR ‘point’ spectroscopy in postharvest fruit and vegetable assessment: The science behind three decades of commercial use. *Postharvest Biology and Technology*, 168, 111246. <https://doi.org/10.1016/j.postharvbio.2020.111246>
- Wang, P., & Yu, Z. (2015). Species authentication and geographical origin discrimination of herbal medicines by near infrared spectroscopy: A review. *Journal of Pharmaceutical Analysis*, 5(5), 277–284. <https://doi.org/10.1016/j.jpha.2015.04.001>
- Williams, P., & Norris, K. (2001). *Near-Infrared Technology: In the Agricultural and Food Industries* (2nd ed.).
- Wolodarsky-Franke, A., & Díaz, S. (2011). *Cordillera de Nahuelbuta. Reserva Mundial de Biodiversidad*. http://awsassets.panda.org/downloads/nahuelbuta2011_1.pdf
- Yepes, A., & Buckeridge, M. S. (2011). Respuestas de las plantas ante factores ambientales del cambio climático global: Revisión. *Colombia Forestal*, 14, 213–232.
- Zamorano-Elgueta, C., Cayuela, L., González-Espinosa, M., Lara, A., & Parra-Vázquez, M. R. (2012). Impacts of cattle on the South American temperate forests: Challenges for the conservation of the endangered monkey puzzle tree (*Araucaria araucana*) in Chile. *Biological Conservation*, 152, 110–118. <https://doi.org/10.1016/j.biocon.2012.03.037>

- Zarco-Tejada, P., Miller, J., Harron, J., Hu, B., Noland, T., Goel, N., Mohammed, G., & Sampson, P. (2004). Needle Chlorophyll content estimation through model inversion using hyperspectral data from Boreal Conifer Forest Canopies. *Remote Sensing of Environment*, 89, 189–199. <https://doi.org/10.1016/j.rse.2002.06.002>
- Zhang, Z., Liu, P., Kang, H., Liao, C., Chen, Z., & Xu, G. (2008). [A study of the diversity of different geographical populations of *Emmenopterys henryi* using FTIR based on principal component analysis and cluster analysis]. *Guang pu xue yu guang pu fen xi = Guang pu*, 28, 2081–2086.



ANEXO 1: TABLA RESUMEN DE LOS RESULTADOS DEL CÁLCULO DE MODELOS K-NN

Rango espectral	N° muestras	K	Entrenamiento		N° muestras	Validación	
			Clasificado correctamente	Mal clasificado		Clasificado correctamente	Mal clasificado
VIS-NIR	99	2	88	11	42	38	4
FT-NIR	116	2	96	20	51	44	7
NIR	126	5	114	12	54	46	8
FTIR	123	8	101	22	53	42	11

ANEXO 2: TABLA RESUMEN DE LOS RESULTADOS DEL CÁLCULO DE MODELOS SIMCA

Rango espectral	N° muestras	N° PCs		Entrenamiento			N° muestras	Validación		
		CS1	CS2	Clasificado correctamente	Mal clasificado	Sin coincidencia		Clasificado correctamente	Mal clasificado	Sin coincidencia
VIS-NIR	99	4	4	96	1	0	42	38	4	0
FT-NIR	116	3	3	109	7	0	51	45	6	0
NIR	126	6	4	112	14	0	54	48	5	1
FTIR	123	5	5	115	8	0	53	48	5	0