

Imagen hiperespectral-NIR y machine learning para la clasificación no-invasiva de quimiotipos en *cannabis sativa* L

La actual tasa de aceptación pública hacia la viabilidad del cannabis medicinal ha provocado un aumento mundial en la producción de esta especie de planta. Sin embargo, el marco legal actual, en continua transformación, no exime al conocimiento sobre el cultivo de cannabis de su origen clandestino, donde se carece de normas o estándares de garantía y control de calidad para los correctos procesos de fabricación y manufacturación.

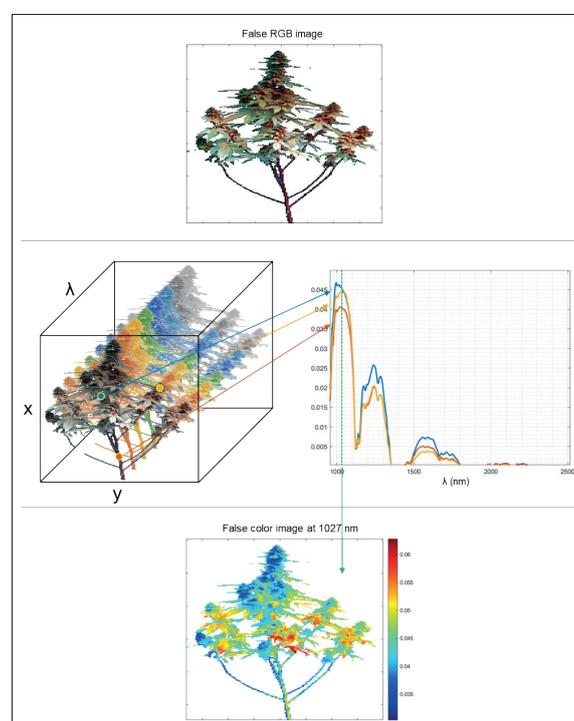
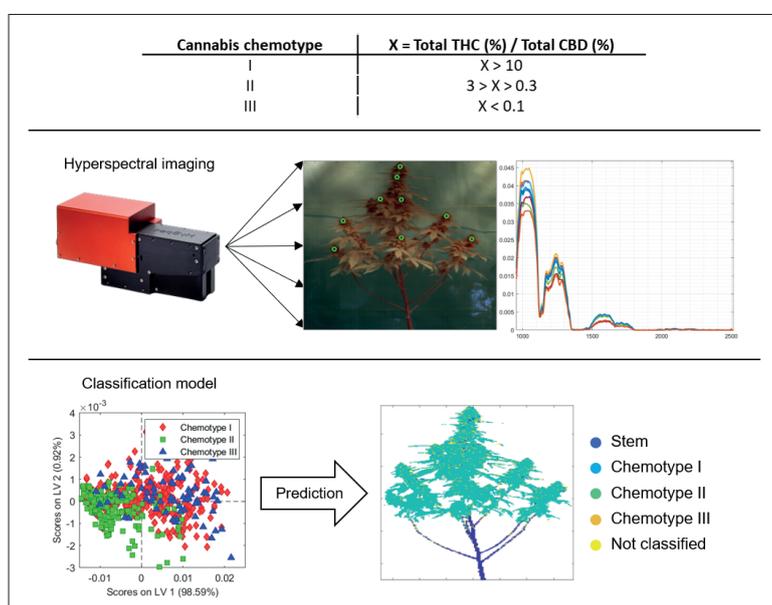


Figura 1: Visualización de una imagen hiperespectral NIR de una planta de cannabis obtenida en el rango espectral de 930-2500 nm

(San Nicolas et al., 2024) el marco legal actualmente en transformación no impide el conocimiento originalmente ilícito en torno al cultivo de cannabis, el cual carece de regulaciones de control de calidad o estándares para procesos de fabricación correctos, hecho que posteriormente podría dar lugar a productos de cultivo incontrolados e incluso dañinos. En esta línea, el objetivo de este trabajo ha sido desarrollar una metodología no invasiva para la clasificación del quimiotipo del cannabis en diferentes cultivares durante el proceso de cultivo de la planta, con el fin de mantener un control de producción indudable sobre los cultivos de cannabis. Para ello, se han utilizado técnicas de imagen hiperespectral (HSI).

MARKEL SAN NICOLAS ORUETXEBARRIA
DIRECTOR TÉCNICO DE SOVEREIGN FIELDS S.L.

La ausencia de dichas medidas causa preocupación a las agencias reguladoras pertinentes, ya que este hecho puede, subsecuentemente, dar lugar a productos de cultivo no controlados e incluso perniciosos. Por ello, con el fin de mantener un indudable control de producción sobre los cultivos con fines medicinales, se realizó un trabajo de investigación para demostrar la prueba de concepto de una metodología no invasiva y fácilmente aplicable para la rápida determinación del quimiotipo de diferentes variedades de *Cannabis sativa* L. durante el proceso de cultivo. Así pues,

esta prueba de concepto se demostró mediante la imagen hiperespectral (Figura 1), combinado con diferentes metodologías de análisis multivariante de datos, así como el análisis de componentes principales (PCA) o el análisis discriminante de mínimos cuadrados parciales (PLS-DA), que permitieron el análisis *in-situ* no invasivo de las plantas.

Procedimiento experimental

Para este trabajo se tomaron imágenes NIR-hiperespectrales de 57 individuos de cannabis, divididos en diferentes variedades: 10 individuos de Dairy Queen (DQ), 9 de Futura 75 (FT), 10 de Remedy (RE), 8 de Roma

(RO), 10 de Tel Aviv (TA) y 10 de White Widow (WW). Se enumeró cada uno ellos. En la Figura 2 pueden observarse algunos ejemplos de las imágenes visualizadas como imágenes RGB falsas.

Se extrajo un total de 502 espectros de las imágenes hiperespectrales-NIR mediante la selección manual de píxeles en las inflorescencias: 249 del quimiotipo I, 178 del quimiotipo II y 75 del quimiotipo III. Además, se extrajo un número acumulado de 219 espectros de tallos de todas las imágenes mediante la selección manual de píxeles de las regiones del tallo de la planta, como puede observarse en la Figura 3.



NACEMOS. CRECEMOS.
EVOLUCIONAMOS.
Y SEGUIMOS CRECIENDO.

Devecan es ahora Evocan:
un cambio para adaptarse al futuro.

El mercado del cannabis medicinal en Europa crece de forma exponencial.
Para afrontarlo nace Evocan, una nueva compañía más fuerte y mejor preparada
para ayudarte a hacer realidad tu proyecto de cannabis legal con un soporte 360°.

Evoluciona tú también con Evocan.



THE LEGAL CANNABIS EXPERTS

evocan.es

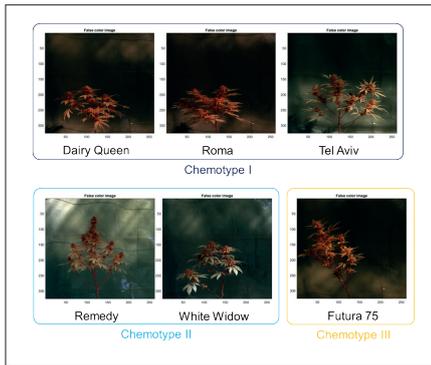


Figura 2: Ejemplos de imágenes hiperespectrales-NIR visualizadas como imágenes RGB falsas de individuos de las diferentes variedades empleadas en el estudio

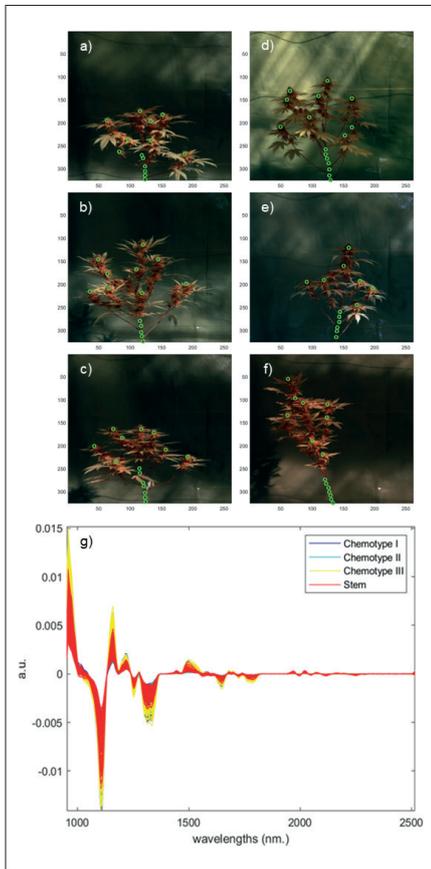


Figura 3: Selección manual de píxeles de flores y tallos en las imágenes hiperespectrales-NIR de varios individuos para obtener espectros y entrenar el modelo de clasificación correspondiente. Variedades: a) Dairy Queen b) Tel Aviv c) Roma d) Remedy e) White Widow f) Futura 75. e) Espectros NIR de los píxeles de flor y tallo utilizados en la construcción de los modelos de clasificación

Resultados y discusión

El quimiotipo de cada variedad se definió previamente mediante el análisis destructivo de cannabinoides en sus inflorescencias secas por cromatografía líquida de alto rendimiento acoplado a un detector ultravioleta-visible (HPLC-UV/Vis.). Las concentraciones correspondientes de cannabinoides principales se muestran en la *Tabla 1*, a partir de la cual las variedades Dairy Queen, Tel

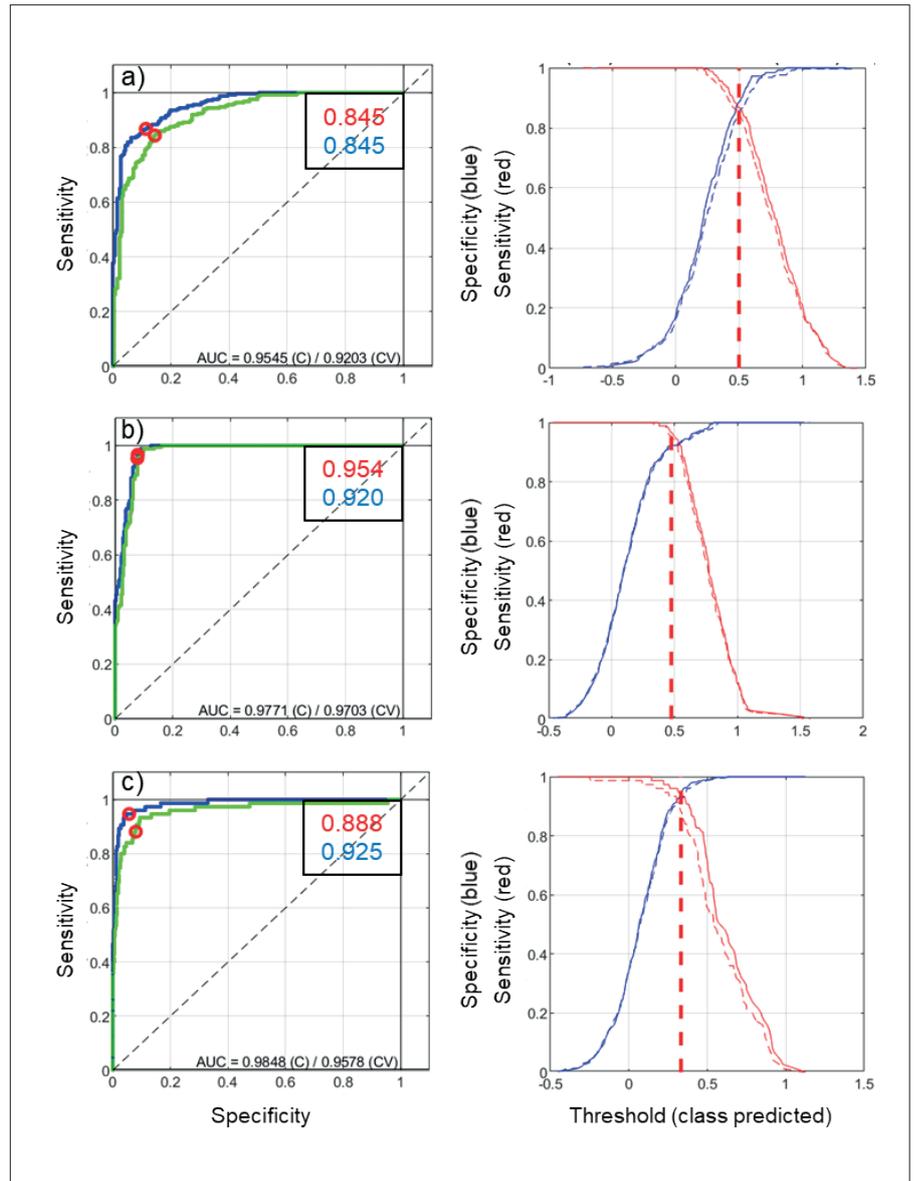


Figura 4: Curvas ROC estimadas (línea azul) y cross-validadas (línea verde) y curvas de respuesta estimadas (línea continua) y cross-validadas (línea discontinua) a) Quimiotipo I b) Quimiotipo II c) Quimiotipo III

TABLA 1: CONCENTRACIONES DE THC Y CBD TOTALES EN INFLORESCENCIAS SECAS DE DIFERENTES VARIEDADES DE CANNABIS, DETERMINADOS POR HPLC-UV/VIS

Variedad	Concentración de cannabinoides (w/w. %)			Clase
	THC Total	CBD Total	CTotal THC / CTotal CBD	
White Widow	5.6	7.5	0.75	Quimiotipo II
Dairy Queen	20.2	0.7	28.86	Quimiotipo I
Tel Aviv	16.9	0.2	84.50	Quimiotipo I
Roma	17.3	0.2	86.50	Quimiotipo I
Futura 75	0.7	15.6	0.04	Quimiotipo III
Remedy	6.8	8.8	0.77	Quimiotipo II

Aviv y Roma se definieron como pertenecientes al quimiotipo I; White Widow y Remedy se definieron como quimiotipo II; y Futura 75 se definió como quimiotipo III.

Aviv y Roma se definieron como pertenecientes al quimiotipo I; White Widow y Remedy se definieron como quimiotipo II; y Futura 75 se definió como quimiotipo III.

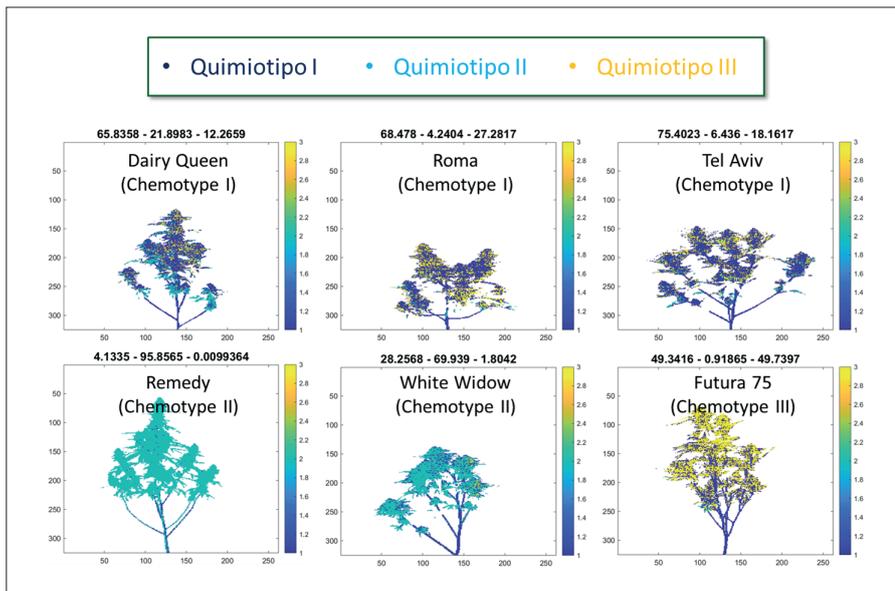


Figura 5: Ejemplos de predicción de quimiotipo en algunos individuos de las variedades empleadas en el estudio.

El modelo de clasificación de los tres quimiotipos se entrenó utilizando 8 variables latentes. La validación cruzada del modelo se realizó mediante submuestreo aleatorio, ya que las tres clases definidas poseían diferentes pesos estadísticos en el modelo, con 10 divisiones de datos permutadas en 5 iteraciones. Los valores de sensibilidad y especificidad optimizados mediante validación cruzada fueron los siguientes para cada clase: 0.845/0.845 para el Quimiotipo I, 0.954/0.920 para el Quimiotipo II y 0.888/0.925 para el Quimiotipo III. En la *Figura 4* se muestran las curvas ROC de cada clase.

Con el modelo de clasificación obtenido se predijo el quimiotipo en los demás píxeles que constituían cada imagen hiperespectral-NIR para realizar la validación externa del modelo, y así determinar mediante la metodología no-invasiva el quimiotipo de la planta completa. En la *Figura 5* pueden observarse los ejemplos de predicción de quimiotipo de algunos de los individuos. Las probabilidades de quimiotipo predichas para cada individuo mediante el modelo de clasificación se presentan en la *Tabla 2*.

El umbral mínimo para predecir el quimiotipo correspondiente en cada píxel también se fijó en un 50 % de probabilidad. Así, para la predicción externa del quimiotipo de un individuo vegetal, se estableció una regla de restricción, según la cual al menos el 50 % de los píxeles que constituyen una imagen debían clasificarse en uno de los quimiotipos. De acuerdo con esta afirmación, los quimiotipos predichos sobre las clases verdaderas de los individuos se muestran en la *Tabla 3*.

En los resultados presentados en la *Tabla 3* se puede observar que 4 individuos se clasificaron incorrectamente (7,02 % de tasa de clasificación errónea). A su vez, otros dos individuos no se clasificaron en ninguna de las clases (ratio de no-clasificación del 3,51 %), por lo que 51 de los 57 individuos se clasificaron correctamente (89,47 % de veracidad). En esta línea, cabe destacar que todos los individuos pertenecientes al quimiotipo II fueron correctamente clasificados (100 % de veracidad). Se esperaba que los mejores resultados correspondieran a esta clase, ya que mostró los mejores valores globales de sensibilidad y especificidad en el entrenamiento del modelo (0.954 y 0.920, respectivamente). Además, el volumen de individuos del quimiotipo I analizados fue el mayor entre las tres clases (n=28). Esta clase estaba compuesta por tres variedades diferentes, lo que podría llevar a una mayor probabilidad de clasificación errónea debido a la mayor variabilidad biológica entre los individuos. Esta podría ser la razón por la que, de los 28 individuos, hubo 25 individuos correctamente clasificados (89,29 % de veracidad). Por último, el modelo obtuvo los peores resultados en la predicción de plantas quimiotipo III. En este caso, sólo 6 de 9 individuos fueron clasificados correctamente (66,67 % de veracidad). No obstante, debido a la dispersión de los píxeles de cada clase en las imágenes de individuos pertenecientes al quimiotipo III, se dedujo que las clasificaciones incorrectas se debieron a la gran diferencia de peso estadístico entre este grupo y la clase de quimiotipo I.

Conclusiones

Así pues, este trabajo concluyó que, mediante el análisis de imágenes hiperespectrales-NIR de *Cannabis sativa* L., se pueden clasificar de forma representativa diferentes variedades pertenecientes a los quimiotipos I, II y III, evitando técnicas analíticas invasivas, demostrando así la prueba de concepto de una metodología fácil de aplicar y no-destructiva. Por lo tanto, la técnica basada en imagen hiperespectral-NIR para la clasificación de quimiotipos abordó adecuadamente el principal hándicap del análisis de cannabinoides mediante espectroscopía de infrarrojo cercano (NIRS), que es la humedad presente en el tejido vegetal fresco, permitiendo un análisis representativo directamente en un individuo vegetal vivo completo.

De este modo, al haberse alcanzado el objetivo principal de este trabajo, se demostró la prueba de concepto para el control exhaustivo de la calidad de los cultivos, y más concretamente del cannabis, mediante una técnica analítica no invasiva. Este hecho podría allanar el camino para el desarrollo de tecnologías no-invasivas en el control de calidad agrícola, ya que el objetivo planteado se consiguió evitando técnicas analíticas habituales como la cromatografía o la espectroscopia NIR convencional. Sin embargo, para hacerlo posible, el modelo de clasificación necesitaría datos de entrenamiento mucho más amplios. Para mejorar la generalizabilidad de los resultados, se necesitaría un número de muestras mayor, con una mayor cobertura de diversidad de variedades a través de diferentes tiempos de cultivo. Por lo tanto, una propuesta ideal sería ampliar los datos de calibración del modelo con más variedades de los tres quimiotipos, equilibrando idealmente el peso estadístico de las clases, y añadiendo espectros de diferentes tejidos de la planta, como las hojas, y complementándolo con espectros en diferentes etapas de crecimiento. De esta forma, se deduciría una tendencia regresiva en función de los distintos momentos de crecimiento, haciendo posible la predicción del quimiotipo correspondiente en individuos en fases tempranas de crecimiento ●

Referencias

San Nicolas, M., Villate, A., Alvarez-Mora, I., Olivares, M., Aizpurua-Olaizola, O., Usobiaga, A., & Amigo, J. M. (2024). NIR-hyperspectral imaging and machine learning for non-invasive chemotype classification in *Cannabis sativa* L. *Computers and Electronics in Agriculture*, 217. Scopus. <https://doi.org/10.1016/j.compag.2023.108551>

TABLA 2: RESULTADOS DE PROBABILIDAD EN LA PREDICCIÓN DE LA CLASE DE QUIMIOTIPO EN CADA PLANTA COMPLETA

Individuo	Tipo	Predicción de clases (Probabilidad, %)		
		Quimiotipo I	Quimiotipo II	Quimiotipo III
DairyQueen_1	Quimiotipo I	80.01	12.99	7.00
DairyQueen_2	Quimiotipo I	75.87	19.48	4.65
DairyQueen_3	Quimiotipo I	65.84	21.90	12.27
DairyQueen_4	Quimiotipo I	80.62	16.14	3.23
DairyQueen_5	Quimiotipo I	68.55	22.98	8.47
DairyQueen_6	Quimiotipo I	60.35	29.75	9.90
DairyQueen_7	Quimiotipo I	56.43	36.51	7.06
DairyQueen_8	Quimiotipo I	56.23	37.94	5.83
DairyQueen_9	Quimiotipo I	51.94	41.61	6.45
DairyQueen_10	Quimiotipo I	48.18	48.81	3.00
Futura75_1	Quimiotipo III	49.34	0.92	49.74
Futura75_2	Quimiotipo III	50.44	1.76	47.80
Futura75_3	Quimiotipo III	59.15	2.43	38.41
Futura75_4	Quimiotipo III	39.80	0.71	59.49
Futura75_5	Quimiotipo III	38.04	2.31	59.64
Futura75_7	Quimiotipo III	35.63	2.99	61.38
Futura75_8	Quimiotipo III	35.98	1.72	62.29
Futura75_9	Quimiotipo III	38.59	2.26	59.14
Futura75_10	Quimiotipo III	30.87	3.81	65.33
Remedy_1	Quimiotipo II	4.13	95.86	0.01
Remedy_2	Quimiotipo II	0.99	99.01	0.00
Remedy_3	Quimiotipo II	1.51	98.49	0.00
Remedy_4	Quimiotipo II	1.89	98.11	0.00
Remedy_5	Quimiotipo II	1.23	98.77	0.00
Remedy_6	Quimiotipo II	2.39	97.61	0.00
Remedy_7	Quimiotipo II	3.42	96.58	0.00
Remedy_8	Quimiotipo II	3.47	96.53	0.00
Remedy_9	Quimiotipo II	7.18	92.82	0.00
Remedy_10	Quimiotipo II	4.88	95.12	0.00
Roma_1	Quimiotipo I	68.48	4.24	27.28
Roma_2	Quimiotipo I	75.88	10.88	13.23
Roma_3	Quimiotipo I	83.59	6.31	10.10
Roma_4	Quimiotipo I	67.25	9.31	23.44
Roma_5	Quimiotipo I	70.84	7.02	22.14
Roma_6	Quimiotipo I	72.46	6.08	21.46
Roma_7	Quimiotipo I	70.49	14.01	15.51
Roma_8	Quimiotipo I	74.94	15.59	9.47
TelAviv_1	Quimiotipo I	75.40	6.44	18.16
TelAviv_2	Quimiotipo I	50.79	46.12	3.09
TelAviv_3	Quimiotipo I	8.05	91.95	0.00
TelAviv_4	Quimiotipo I	9.02	90.98	0.00
TelAviv_5	Quimiotipo I	76.39	2.15	21.47
TelAviv_6	Quimiotipo I	74.94	3.83	21.23
TelAviv_7	Quimiotipo I	78.74	2.76	18.50
TelAviv_8	Quimiotipo I	74.21	0.92	24.87
TelAviv_9	Quimiotipo I	72.42	1.21	26.37
TelAviv_10	Quimiotipo I	70.12	2.40	27.48
WhiteWidow_1	Quimiotipo II	28.26	69.94	1.80
WhiteWidow_2	Quimiotipo II	23.42	75.85	0.74
WhiteWidow_3	Quimiotipo II	22.96	76.54	0.50
WhiteWidow_4	Quimiotipo II	24.52	74.59	0.89
WhiteWidow_5	Quimiotipo II	20.32	79.24	0.44
WhiteWidow_6	Quimiotipo II	15.40	84.59	0.01
WhiteWidow_7	Quimiotipo II	12.81	87.15	0.04
WhiteWidow_8	Quimiotipo II	10.21	89.78	0.01
WhiteWidow_9	Quimiotipo II	7.66	92.34	0.00
WhiteWidow_10	Quimiotipo II	4.95	95.03	0.02

TABLA 3: CLASES VERDADERAS Y PREDICHAS DE LOS 57 INDIVIDUOS DE LAS 6 VARIETADES EMPLEADAS, PERTENECIENTES A LOS TRES QUIMIOTIPOS, MEDIANTE EL MODELO DE CLASIFICACIÓN ENTRENADO

Clase verdadera		Clase predicha (Probabilidad de quimiotipo > 50 %)				
		Quim. I	Quim. II	Quim. III	No clasificado	
Clase verdadera	Quimiotipo I (n=28)	25	2	0	1	28
	Quimiotipo II (n=20)	0	20	0	0	20
	Quimiotipo III (n=9)	2	0	6	1	9
		27	22	6		Suma = 57



TAIB®

A world of solutions



◆ Hacia un mañana más seguro y confiable

Desde la investigación hasta la producción, aseguramos productos confiables que cumplen con los más altos estándares.

Ingeniería, robótica, fabricación y automatización



CALIDAD

Superamos expectativas.



EFICIENCIA

Impulsamos resultados.



COMPROMISO

Ofrecemos soluciones duraderas.



www.taib.es