

**Caracterización de la diversidad fitoquímica del cannabis colombiano mediante  
análisis estadísticos y pruebas de clasificación por quimiotipos**

**Carlos D. Penagos S.**

UNIVERSIDAD ICESI

Facultad de Ingeniería, Diseño y Ciencias Aplicadas

PROGRAMA DE QUÍMICA FARMACÉUTICA

Santiago de Cali

2024

**Caracterización de la Diversidad fitoquímica del cannabis colombiano mediante  
análisis estadísticos y pruebas de clasificación por quimiotipos**

**Carlos D. Penagos S.**

**TRABAJO DE GRADO PARA OPTAR AL TITULO DE QUIMICO  
FARMACEUTICO**

**Tutor: Diego E. Valencia MSc**

**Santiago de Cali**

**2024**

**AVAL PARA LA ENTREGA DE PDG II**

**Firma del Asesor:**

A handwritten signature in black ink, enclosed within a large, horizontal oval shape. The signature is cursive and appears to read 'Diego Enríquez'.

---

**Diego Javier Enríquez Valencia**

**Firma del estudiante 1**

A handwritten signature in black ink that reads 'Carlos D. Penagos S.'.

---

**Carlos David Penagos Serrano**

## Resumen

*Cannabis sativa* se clasifica en cinco quimiotipos basados en sus cannabinoides principales:  $\Delta^9$ -THC, CBD y CBG. Estudios globales identifican tres de los cinco quimiotipos principales: THC-dominante (Tipo I), CBD-dominante (Tipo III) y balanceado (Tipo II). En EE.UU. e India prevalece el quimiotipo I, mientras que en Brasil domina el quimiotipo III. En Colombia, un estudio de 2015 reveló que el 98% de 335 muestras pertenecían al Tipo I. Un estudio de la Universidad Icesi en 2022 descubrió variedades del quimiotipo CBG-dominante (Tipo IV) algo nuevo en contraste con estudios anteriores. Sin embargo, a la fecha la diversidad fitoquímica del cannabis colombiano aún no se ha establecido completamente. El objetivo de este trabajo es determinar la diversidad fitoquímica del cannabis colombiano mediante la clasificación de quimiotipos. Para ello, se implementó una metodología en cinco pasos: obtención de certificados de análisis de cannabis, estructuración de datos en Excel, procesamiento en R studio, clasificación de quimiotipos y análisis estadísticos PCA y Spearman. De 107 certificados de análisis evaluados, se encontraron contenidos mayoritarios de CBDA y THCA. La data se agrupó en tres quimiotipos: CBD dominante (62.6%), THC dominante (23.4%) y balanceado (14%). El análisis de correlación de Spearman mostró una fuerte correlación positiva entre  $\Delta^9$ -THC y CBG, consistente con lo reportado en la literatura. Por último, el análisis por componentes principales permitió agrupar las muestras según sus quimiotipos: THC dominante, balanceado y CBD dominante.

**Palabras claves:** Cannabinoides, quimiotipos, diversidad fitoquímica, PCA, correlación Spearman

## INTRODUCCIÓN

*Cannabis sativa*, ha sido objeto de numerosos estudios quimiotaxonómicos. Actualmente, esta planta se clasifica según las concentraciones de sus cannabinoides principales: tetrahidrocannabinol ( $\Delta^9$ -THC), cannabidiol (CBD) y cannabigerol (CBG)<sup>[1]</sup>. Inicialmente dividida en dos subespecies, *Cannabis sativa* (*C. sativa*) y *Cannabis Indica*, ahora reconocida como única especie (*C. sativa*)<sup>[2]</sup>. Esta planta, originaria de Asia, ha conquistado diversos ambientes gracias a su extraordinaria capacidad de adaptación, dando origen a cepas únicas con perfiles cannabinoides distintivos los cuales podrían exhibir potenciales propiedades medicinales como antiemético, anticonvulsivos, antiinflamatorio y antioxidante<sup>[1]</sup>.

El aumento del interés en comprender las propiedades del cannabis, junto con la notable variabilidad fitoquímica de la planta, ha resultado en la elaboración de una clasificación basada en las concentraciones de sus cannabinoides principales. Estos compuestos, que funcionan como marcadores quimiotaxonómicos, son metabolitos secundarios exclusivos del cannabis y facilitan la categorización de la planta en cinco quimiotipos primordiales. Estos quimiotipos se subdividen en: I) plantas tipo droga con predominio de THC (THC>0.3% y CBD<0.5%); II) plantas medicinales con equilibrio entre THC y CBD; III) plantas tipo fibra con concentración elevada de CBD (THC <1%); IV) plantas tipo fibra con CBG como cannabinoide principal; V) cáñamo de fibra industrial con escasos cannabinoides (**Tabla 1**) <sup>[3,4]</sup>. Por ende, se ha propuesto una clasificación que abarque todos sus componentes, garantizando una distinción quimiotaxonómica más precisa.

En los últimos años se han realizado estudios quimiotaxonómicos en diferentes partes del mundo con el fin de verificar la diversidad de esta planta. Investigaciones exhaustivas, como el estudio iraní de 2022 que abarcó 161 muestras de 20 regiones geográficas, han revelado la presencia de distintos quimiotipos, destacando la prevalencia de cepas THC-predominantes (tipo droga) <sup>[5]</sup>. A su vez, en India (2023), se realizó un análisis con 66 muestras; se evidenció que el fenotipo de cannabis tipo droga predominaba <sup>[6]</sup>. En Estados Unidos (con un impresionante conjunto de 89,923 muestras) <sup>[7,8]</sup> y Brasil (con 35 muestras) comprobaron la diversidad de cannabinoides predominante en cada una de sus regiones, este último país encontró especies CBD-dominantes <sup>[9]</sup>.

En cuanto a la variabilidad en el territorio nacional colombiano, en el año 2015 se ejecutó un análisis por parte del ministerio de Justicia y Derecho. Ellos recolectaron 335 muestras en ciudades principales como Cali, Medellín y Bogotá; donde se obtuvieron resultados semejantes a los estudios en el mundo entero, con el 98% de las muestras agrupadas en Tipo I (droga) y solo 7 muestras poseen un perfil tipo medicinal o fibra <sup>[4]</sup>. Sin embargo, en un estudio realizado por la universidad Icesi en 2022, se seleccionaron, 156 muestras de 17 empresas cultivadoras de cannabis diferentes de Colombia. Esta vez, se obtuvieron resultados prometedores no solo en sus variaciones morfológicas como, forma, tamaño, color, estructura, aroma si no también en su perfil

químico; en donde se encontró hasta el quimiotipo CBG-Dominante, algo significativo y comparable con respecto a los otros estudios realizados en otras partes del mundo <sup>[10]</sup>.

A pesar de que existen más de 20 laboratorios acreditados en el país para la cuantificación de cannabinoides con fines medicinales investigativos e industriales <sup>[11]</sup>, a la fecha no se han realizado estudios estadísticos que evalúen toda la data existente en el país. Esta situación subraya la necesidad de identificar la diversidad de cannabinoides presentes en las variedades de cannabis colombiano y su potencial para enriquecer la investigación sobre el cannabis a nivel global. Por lo tanto, el objetivo de este estudio es analizar e identificar los resultados de los análisis de muestras obtenidas por laboratorios analíticos de Colombia.

### **Objetivo general:**

Establecer la diversidad fitoquímica del cannabis colombiano mediante pruebas de clasificación por quimiotipos aplicados a los resultados de análisis de potencia de cannabinoides obtenidos por un laboratorio analítico.

### **Objetivos específicos**

- 1) Identificar quimiotipos específicos (I, II, III, IV u otros) a partir de los datos reportados en certificados analíticos de cuantificación de cannabinoides.
- 2) Determinar correlaciones entre los cannabinoides analizados en la data suministrada, empleando la prueba de correlación de Spearman y análisis de componentes principales.

## **METODOLOGÍA**

La metodología empleada se fundamenta principalmente en cinco (5) etapas que se describen a continuación:

**1. Adquisición de datos:** Se recibieron los datos de un laboratorio analítico de Colombia, que incluyen certificados de análisis sobre perfiles de cannabinoides. Esta información de laboratorio consiste en mediciones de cannabinoides obtenidos mediante cromatografía líquida de alta resolución (HPLC) y espectrometría de masas (GC-MS). Esta información se procesó en una matriz de Excel donde posteriormente sirvió como data frame (df) para el software R studio.

**2. Procesamiento de datos:** Este apartado incluye un tratamiento de la matriz de datos creada a partir de la data anterior, aquí se emplearon scripts (**Ver anexo 1**) y librerías del software estadístico R studio con el fin de eliminar faltantes, reestructurar la matriz, limpiar archivos sin procesar.

**3. Categorización por quimiotipos:** En este apartado se empleará la concentración total obtenida de cada cannabinoide a partir de los certificados de análisis, para clasificar así cada muestra según la **Tabla 1**.

**Tabla 1.** Clasificación del cannabis por quimiotipos.

| Perfil químico (quimiotipo) | Cannabinoide predominante                |
|-----------------------------|--|
| I                           | THC -Dominante (proporción 5:1, THC:CBD) |
| II                          | THC:CBD (proporción 1:1) Balanceado      |
| III                         | CBD-dominante (proporción 1:5, THC:CBD)  |
| IV                          | CBG-dominante                            |
| V                           | Sin cannabinoides predominantes          |

Para eso, primero, se realizó el cálculo de los porcentajes totales de CBD, THC y CBG que es la cantidad máxima que podemos obtener de estos cannabinoides si descarboxilamos adecuadamente la muestra. A continuación, se tuvo en cuenta un factor de conversión teniendo presente la descarboxilación (pérdida de CO<sub>2</sub>) de la molécula ácida. Por ejemplo, cuando el THCA pierde una molécula de CO<sub>2</sub>, se transforma en THC, que es una molécula más pequeña (el THCA pesa 358 g/mol mientras que el THC pesa 314 g/mol). Considerando esta pérdida, por cada molécula de THCA, se obtienen aproximadamente  $358/314 = 0.877$  moléculas de THC <sup>[12]</sup>. Ahora, con ese factor se procede a calcular el total para el THC como se evidencia a continuación:

$$\%THC_{total} = \%THC + (\%THCA \times 0.877)$$

**Ecuación 1.** Fórmula para calcular el % Total THC en los datos.

A partir de la **Ecuación 1**, se calculó también para el %CBD<sub>total</sub>, %CBG<sub>total</sub> y %CBC<sub>total</sub>. Con los totales calculados se estableció que: especies cuyo THC supera el CBD, en relación (5:1, THC-CBD) son clasificadas como THC-dominantes, correspondientes al quimiotipo I. Aquellas con una relación (1:1, THC-CBD), se entienden que son Balanceado y se categorizan como tipo II. El tipo III se asigna a especies donde relación es (1:5, THC-CBD), estas se refieren a las CBD-dominantes. Aquellas en las que

prevalece el cannabigerol (CBG) se identifican como tipo IV, y el tipo V se refiere a aquellas sin una cantidad detectable de cannabinoides. <sup>[13]</sup>

**4. Correlación de Spearman:** La correlación de Spearman cuantifica la fuerza y dirección de la relación lineal entre dos conjuntos de datos, proporcionando un coeficiente que varía de -1 a 1. Un valor cercano a 1 indica una relación positiva fuerte, mientras que un valor cercano a -1 señala una relación negativa. Un valor de 0 sugiere ausencia de relación lineal. También, se estableció la fuerza de la correlación dado el valor del coeficiente de correlación de Spearman (R) obtenido, con el fin de interpretar el comportamiento de las variables. Por lo tanto, si se obtiene un valor entre 0 y 0,3 se entiende que es una correlación débil. Si esta entre 0,3 y 0,7 es una correlación moderada. Y si esta entre 0,7 y 1.0 es una correlación fuerte <sup>[14]</sup>. En el contexto, este método se empleó para examinar la posible correlación entre variables como la concentración de compuestos específicos en las muestras. En el estudio realizado, se empleó para determinar la correlación principalmente entre los contenidos totales de CBD vs THC; THC vs CBG; CBD vs CBG y CBC vs CBG. Se empleó el intervalo de confianza del 95 %,  $p = 0.05$ .

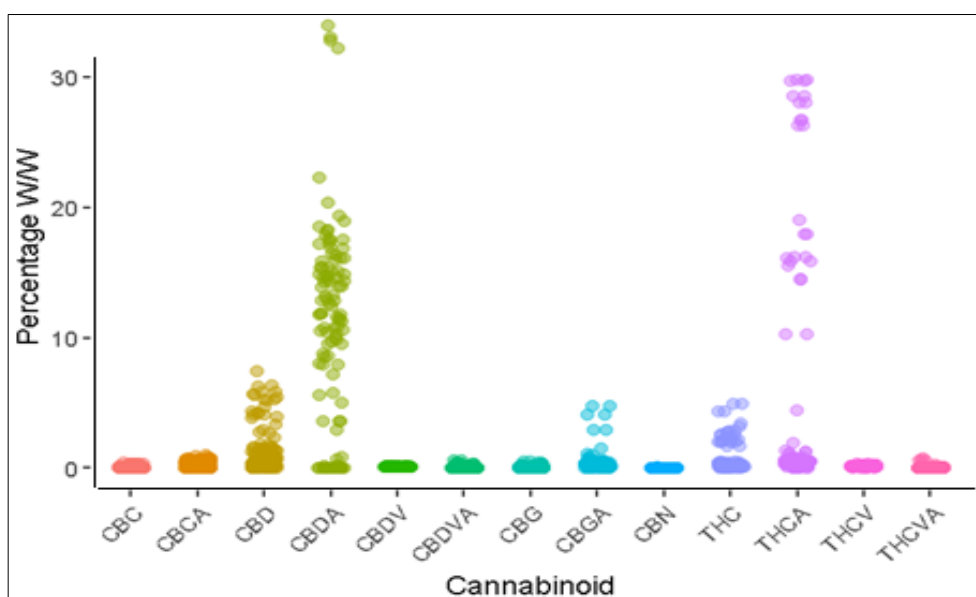
**5. Análisis de componentes principales (PCA) <sup>[15]</sup>:** En este punto se ejecutó una técnica de reducción de dimensionalidad que se utiliza comúnmente en análisis de datos para explorar y visualizar la estructura subyacente de un conjunto de datos. Este revelará la correlación entre diferentes cannabinoides. Los cannabinoides que están altamente correlacionados se ubicarán cerca unos de otros en el espacio de las componentes principales. En este caso los cannabinoides que se emplearan en análisis principales son trece (13):  $\Delta$ -9-tetrahidrocannabinol ( $\Delta^9$ -THC), cannabidiol (CBD) y cannabigerol (CBG), ácido tetrahydrocannabinólico (THCA), tetrahydrocannabivarina (THCV), ácido tetrahydrocannabivarina (THCVA), cannabicromeno (CBC), ácido cannabicromeno (CBCA), ácido cannabidiólico (CBDA), cannabidivarina (CBDV), ácido cannabidivarínico (CBDVA), ácido cannabigerólico (CBGA), cannabinol (CBN). Por último, se realizó una clusterización para identificar si hay subgrupos de muestras que se agrupen y expliquen el enfoque por quimiotipos que se planteó en el trabajo.

## RESULTADOS Y DISCUSIÓN

A continuación, se presentan los resultados obtenidos a partir de los certificados de análisis evaluados.

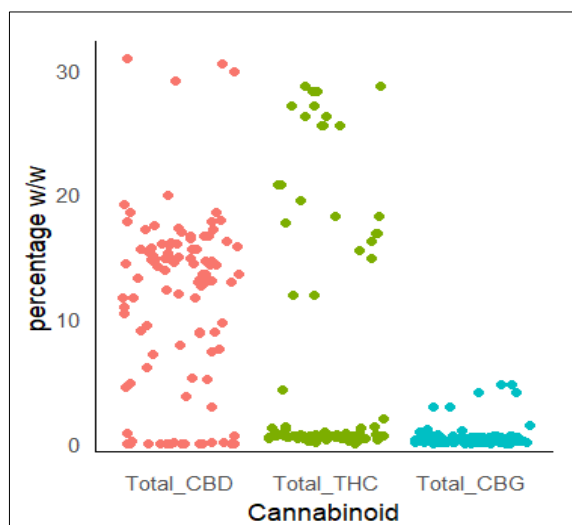
**1. Adquisición de datos:** se recopilaron 107 certificados de análisis con información de 13 cannabinoides en total. Dando como resultado una matriz de 107 observaciones con 13 variables cada una. Ver **Anexo 2**.

**2. Procesamiento de datos:** en este apartado se ejecutaron los códigos descritos en la metodología dando como resultado: una matriz sin datos con “NA” y una fila extra titulada “Quimiotipo” tal y como se evidencia en el **Anexo 3**. Una vez se tenga la matriz sin datos “NA”, se procede a ejecutar la línea de código para generar el gráfico de puntos para todos los cannabinoides recuperados. En total son 13 cannabinoides y su porcentaje w/w se evidencia en la **Figura 1**.



**Figura 1.** Distribución de muestras por porcentaje de cannabinoide (% w/w) para n=107 muestras

Como se puede apreciar, la mayor densidad de muestras se encuentran repartidas entre CBD, CBDA, THC y THCA. Probablemente, nos encontremos frente a muestras de especies CBD y THC dominantes.



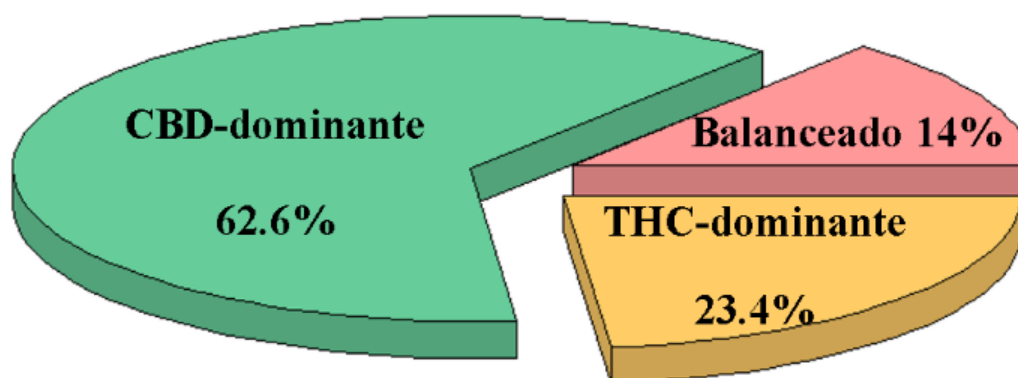
**Figura 2.** Distribución de muestras por porcentaje de CBD, THC y CBG total (%w/w) para n=107 muestras.

Con respecto al **Figura 2**, se evidenció que, al calcular los porcentajes totales por cada cannabinoide de interés (THC, CBD y CBG) (**Ver Ecuación 1**), se denota que la mayoría de las muestras tienen cantidades  $CBD_{total}$  y  $THC_{total}$  altas y  $CBG_{total}$  bajas. Lo anterior resulta coherente dado que, desde la ruta biosintética de los cannabinoides se conoce que, la molécula de donde se obtienen los metabolitos secundarios del cannabis es el CBGA; pues, por acción de las sintasas puede producir THCA, CBDA y CBCA, que posteriormente se descarboxilan produciendo su forma neutra THC, CBD y CBC. Como se obtuvo, Total\_CBD y THC altos, se intuye que este intermediario (CBGA) fue empleado para producir CBDA y THCA mayoritariamente; no obstante, en la data recolectada se ve que existen algunas plantas, específicamente 7, que son ricas en CBG tal y como se evidencia en la Figura.

**3. Categorización por quimiotipo:** una vez se ejecutaron los códigos para este apartado, los resultados obtenidos se plasmaron en la **Tabla 2**. Luego, se generó un gráfico de torta en el cual se estableció el porcentaje de los quimiotipos encontrados en las muestras evaluadas **Figura 3**.

**Tabla 2.** Resultados de la categorización por Quimiotipo.

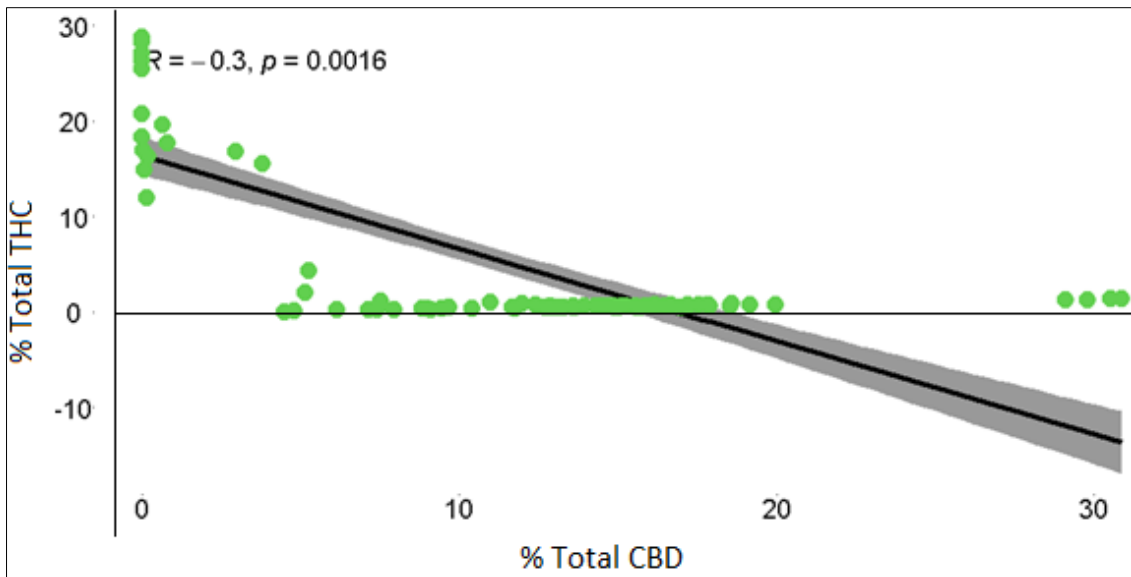
| Quimiotipo    | Número de muestras | Proporción |
|---------------|--------------------|------------|
| THC-dominante | 25                 | 23.4%      |
| CBD-dominante | 67                 | 62.6%      |
| Balanceado    | 15                 | 14%        |
| Total         | 107                | 100%       |



**Figura 3.** Proporción de los quimiotipos identificados.

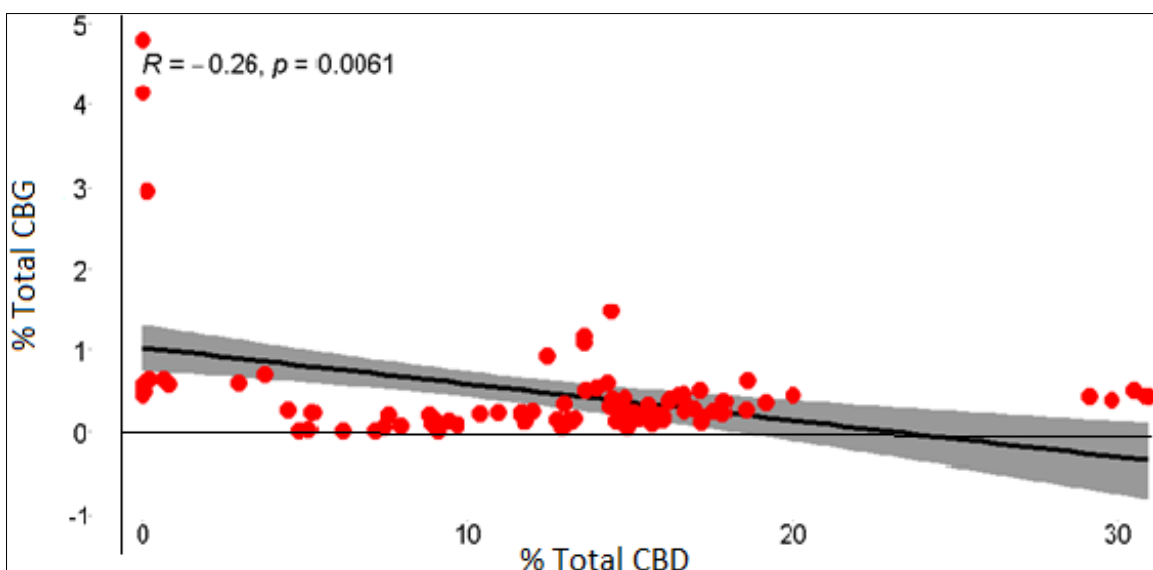
Al analizar la figura 3, se evidenció que, de los 107 datos evaluados, únicamente se encontraron 3 quimiotipos, en donde la mayoría de las muestras ( $n = 67$ ) eran CB dominante  $> 50\%$  (tipo III) y el resto se repartían entre tipo I ( $n = 25$ ) y tipo II ( $n = 14$ ). En la investigación no se encontró el tipo IV, pues como se analizó con anterioridad, la mayoría de las plantas producían mayoritariamente CBDA y THCA. Quizás se deba a la genética de las plantas, pues el interés comercial del cannabis está enfocado a plantas tipo droga (Tipo I) y plantas medicinales (Tipo III), haciendo que los cultivadores se enfoquen en estos quimiotipos. Otra posible razón, está relacionada a sus condiciones de cultivos y tiempos de cosecha, pues es posible que las condiciones favorecieran la actividad enzimática para la producción de plantas ricas en THC y CBD; así mismo, el tiempo de cosecha puede influir, ya que la concentración de CBG en la planta de cannabis puede ser más alta en etapas tempranas de crecimiento antes de que el CBGA se convierta en otros cannabinoides.

**4. Correlación de Spearman:** se presentan los resultados para cada correlación de Spearman para las parejas de cannabinoides evaluadas. En primer lugar, se evaluó la correlación entre el % Total de CBD vs el % Total THC (Ver **Figura 4**).



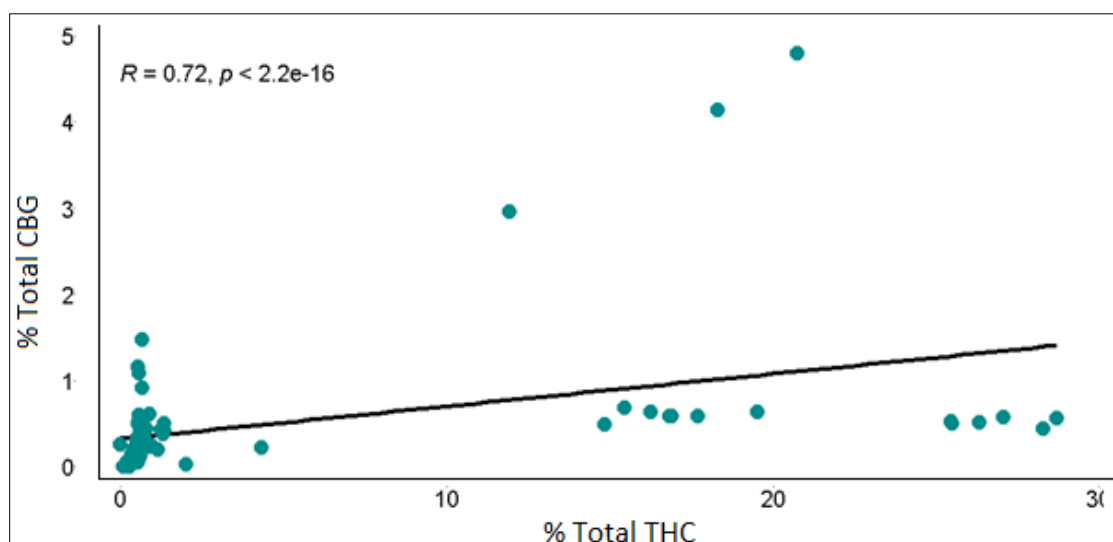
**Figura 4.** Correlación entre THC vs CBD total.

Como se aprecia en el **Figura 4**, se obtuvo un coeficiente de relación de Spearman ( $R = -0.3$ ; correlación negativa-débil). Con respecto al valor  $p = 0.0016$  este está por debajo del valor de significancia 0.05, sugiere que existe una correlación significativa entre las variables que se están analizando. Esto implica que es posible prever de manera confiable que un cambio en una variable se asocie con un cambio en la otra variable. Dado que THCA y CBDA compiten por el mismo precursor (CBGA), es probable que una planta con altos niveles de THCA tenga niveles bajos de CBDA, y viceversa. Por lo tanto, se esperaría que la cantidad total de THC y CBD (que son las formas no ácidas de THCA y CBDA) tengan una correlación negativa. <sup>[13]</sup>



**Figura 5.** Correlación entre CBG vs CBD total.

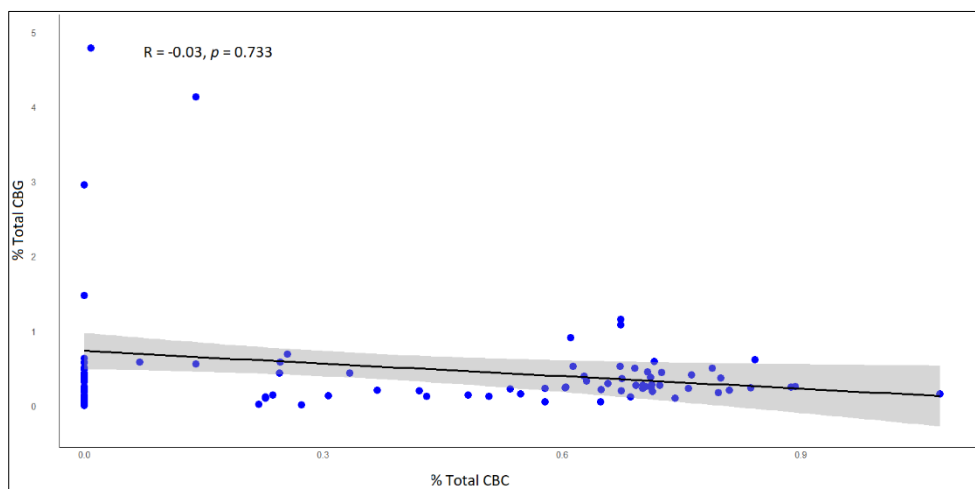
para la relación entre % Total de CBD vs % Total CBG. Se encontró que el coeficiente de correlación de Spearman ( $R = -0.26$ ) indicó una correlación negativa débil. El análisis estadístico mostró un valor  $p$  de 0.0061, que es menor que el umbral 0.05 determinando significancia estadística. Este valor  $p$  indica que la probabilidad de que la correlación observada se deba al azar es muy baja. Por lo tanto, podemos concluir que existe una correlación estadísticamente significativa entre la cantidad total de CBD y CBG. Lo anterior, no tiene sentido desde un punto de vista biológico, pues, el CBG al ser precursor del CBD debería haber una relación positiva. Probablemente, sea por la cantidad de datos, ya que un tamaño de muestra mayor tiende a proporcionar una estimación más precisa y confiable de la correlación entre CBG y CBD. Es recomendable aumentar la cantidad de datos si es posible para obtener una mejor comprensión de la relación entre estos cannabinoides y para minimizar el impacto de la variabilidad aleatoria y el sesgo muestral.



**Figura 6.** Correlación entre CBG vs THC total.

Basándonos en los resultados obtenidos en **Figura 6**, podemos asumir que existe una correlación positiva fuerte ( $R=0.72$ ) y estadísticamente significativa ( $p<0.05$ ) entre %Total THC y el % Total CBG en las muestras analizadas. Esto sugiere que, en general, las variedades de cannabis con mayores niveles de CBG también tienden a tener mayores niveles de THC. Esto probablemente se debe a que, si una planta tiene un alto contenido de CBGA, este precursor puede ser utilizado para la producción de THCA (que se convierte en THC). Por lo tanto, en muestras donde el porcentaje total de CBG

es alto, es probable que también haya una alta cantidad de CBGA disponible que podría contribuir a un mayor porcentaje total de THC [16].

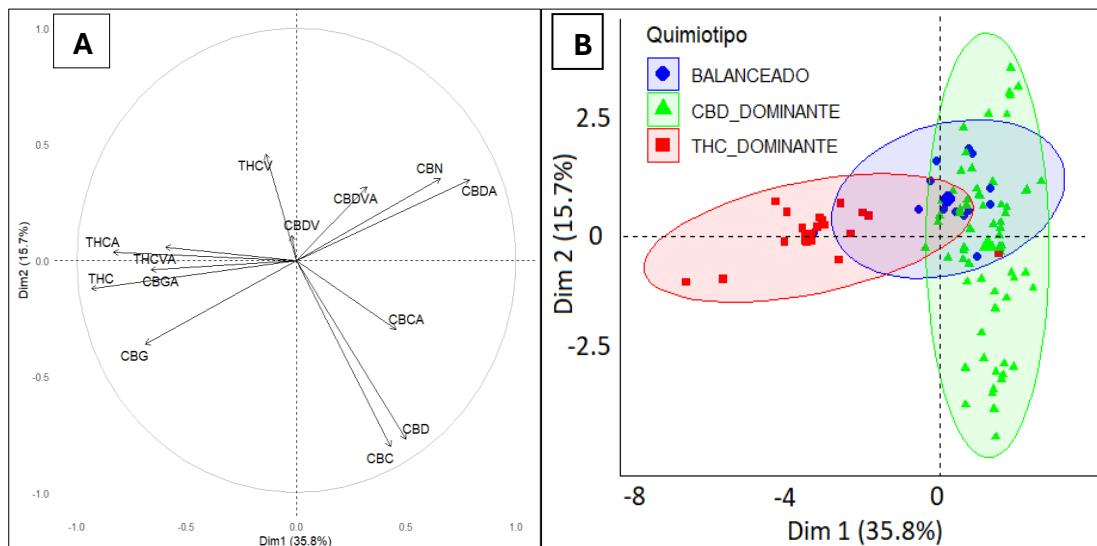


**Figura 7.** Correlación entre CBC vs CBG total.

Finalmente, la correlación observada entre CBC y CBG, indicó una relación negativa-débil ( $r = -0.03$ ) y no significativa ( $p = 0.733$ ) entre estas dos variables; lo que sugiere que la relación puede ser atribuida al azar. Varios factores pueden explicar la falta de correlación observada. Por ejemplo, los valores de 0.0 en la matriz, que representan concentraciones de cannabinoides demasiado bajas para ser detectadas, pueden introducir ruido y afectar negativamente el cálculo de la correlación. Además, la variabilidad en la actividad enzimática, influenciada por factores genéticos y ambientales, junto con diferencias en las condiciones de cultivo, tiempos de cosecha y métodos de extracción, pueden resultar en concentraciones muy variables de CBC y CBG, ocultando cualquier correlación potencial. Finalmente, la muestra de datos puede no ser lo suficientemente grande o representativa.

#### **4. Análisis de componentes principales (PCA)**

Respecto al análisis de componentes principales de los cannabinoides evaluados. Se obtuvo que el componente principal que más explicaba la varianza era el PC1 (35.8%) seguido del PC2 (15,7%). Luego se graficaron las variables que más aportan a cada una de las dimensiones como se muestra a continuación:



**Figura 8.** (A) Aporte de cada variable en el PCA (PC1 vs PC2). (B) Clusters para el PCA agrupados en quimiotipos.

La Figura 8 presenta un análisis de componentes principales (PCA) de los perfiles de cannabinoides en diferentes quimiotipos de cannabis. En la Figura 8A, se observa que los cannabinoides de la familia CBD (CBD, CBDA, CBDV, CBDVA) contribuyen significativamente a la variabilidad explicada por la primera componente principal (Dim1 = 35,8%), con vectores largos que indican una fuerte influencia en la agrupación de las muestras CBD dominantes. En contraste, los vectores de THCA y THCVA apuntan en direcciones opuestas a los de los cannabinoides tipo CBD, señalando una relación inversa y contribuyendo también de manera significativa a Dim1 y Dim2. La magnitud de los vectores en la Figura 8A sugiere que CBDA y THCA son determinantes claves en la explicación de la variabilidad para las muestras balanceadas. La Figura 8B muestra la distribución de los quimiotipos en el espacio de las dos primeras componentes principales: las variedades CBD dominantes (quimiotipo III) se agrupan principalmente en el primer y cuarto cuadrante, mientras que las variedades THC dominantes (quimiotipo I) forman un grupo compacto en el segundo y tercer cuadrante. Las muestras con contenidos balanceados de THC y CBD (quimiotipo II) se distribuyen en los cuadrantes I y II, superponiéndose parcialmente con los quimiotipos I y III, indicando una variabilidad intermedia. El PCA demuestra una clara separación entre los quimiotipos dominados por CBD y THC, con las muestras balanceadas ocupando una posición intermedia, y resalta la efectividad del análisis para distinguir las diferentes variedades en base al aporte de cada uno de sus cannabinoides.

## CONCLUSIONES

- Se estudiaron los resultados de análisis de 107 muestras de cannabis cultivadas en el Valle del Cauca y Cauca. De las muestras solo el CBDA y THCA tenían mayores cantidades (%w/w). Se encontraron 3 quimiotipos principales: CBD\_Dominante (n= 67; 62,5%), THC\_Dominante (n= 25; 23,4%) y Balanceado (n= 15; 14%).
- En cuanto a los análisis de correlación de Spearman, se encontraron correlaciones negativas débiles entre THC vs CBD total ( $R = -0.3$ ), y entre los porcentajes de CBD vs CBG total ( $R = -0.26$ ) y CBC vs CBG total ( $R = -0.03$ ). Sin embargo, se obtuvo una correlación positiva y fuerte entre CBG vs THC total ( $R = 0.72$ ,  $p < 0.05$ ) lo cual es coherente con la ruta de biosíntesis.
- Con el análisis de componentes principales, se logró: establecer cuáles eran las variables que más influían en la agrupación de los datos, así mismo se determinó que la clusterización por quimiotipos; THC\_dominante (Quimiotipo I), Balanceado (Quimiotipo II) y THC\_dominante (Quimiotipo III) resulta útil para agrupar poblaciones de muestras de cannabis

## Bibliografía

- [1] Atakan Z. Cannabis, a complex plant: different compounds and different effects on individuals. *Ther Adv Psychopharmacol.* 2012 Dec;2(6):241-54. doi: 10.1177/2045125312457586. PMID: 23983983; PMCID: PMC3736954.
- [2] ElSohly MA, Gul W (2014) Constituents of Cannabis sativa. In: Pertwee RG (ed) *Handbook of cannabis*. Oxford University Press, UK.
- [3] LEY 2204 DEL 10 DE MAYO DE 2022 regulacion cannabis col. (2022). Recuperado de <https://www.minagricultura.gov.co/Normatividad/Leyes/LEY%202204%20DEL%2010%20DE%20MAYO%20DE%202022.pdf>
- [4] Ministerio de justicia y derecho (2015). Informe técnico de la caracterización química, taxonómica y de mercado de la marihuana producida y consumida en el país. Recuperado de <https://www.minjusticia.gov.co/programas->

[co/ODC/Documents/SAT/Documentos/2015%20Caracterizacion quimica taxonomica mercado marihuana.pdf](https://co/ODC/Documents/SAT/Documentos/2015%20Caracterizacion_quimica_taxonomica_mercado_marihuana.pdf)

- [5] Mostafaei Dehnavi, M., Ebadi, A., Peirovi, A., Taylor, G., & Salami, S. A. (2022). THC and CBD Fingerprinting of an Elite Cannabis Collection from Iran: Quantifying Diversity to Underpin Future Cannabis Breeding. *Plants (Basel, Switzerland)*, 11(1), 129. <https://doi.org/10.3390/plants11010129>
- [6] Jafar Mehdi, Danish Husain, Abhilasha Srivastava, Namita Gupta, Karuna Shanker, Mohammad Talha, Raj Kishori Lal, Birendra Kumar, Anil Kumar Gupta, The genetic diversity and chemical profiling of natural populations of Indian Cannabis, *Industrial Crops and Products*, Volume 203. Recuperado de <https://doi.org/10.1016/j.indcrop.2023.117187>.
- [7] Smith, C. J., Vergara, D., Keegan, B., & Jikomes, N. (2022). The phytochemical diversity of commercial Cannabis in the United States. *PLoS ONE*, 17(5 May). <https://doi.org/10.1371/journal.pone.0267498>
- [8] Vergara, D., Gaudino, R., Blank, T., & Keegan, B. (2020). Modeling cannabinoids from a large-scale sample of Cannabis sativa chemotypes. *PLoS ONE*, 15(9 September).
- [9] Siracusa, L., Ruberto, G., & Cristino, L. (2023). Recent Research on Cannabis sativa L.: Phytochemistry, New Matrices, Cultivation Techniques, and Recent Updates on Its Brain-Related Effects (2018-2023). *Molecules (Basel, Switzerland)*, 28(8), 3387. <https://doi.org/10.3390/molecules28083387>
- [10] Cannabinoid Conference 2022. (2022). *Medical Cannabis and Cannabinoids*, 5(1), Pag 166. P-8. . <https://doi.org/10.1159/000527113>.
- [11] Ministerio de Salud y Protección Social. (2023). USUARIOS INSCRITOS ANTE EL FNE PARA REALIZAR ANÁLISIS DE CALIDAD (CUANTIFICACIÓN) E INVESTIGACIÓN DE SUSTANCIAS FISCALIZADAS. Recuperado de [https://fne.minsalud.gov.co/sustancias-fiscalizadas/SiteAssets/Paginas/Preguntas-frecuentes/Usuarios-inscritos-para-analisis-de-calidad-\(cuantificacion\)-investigacion .pdf](https://fne.minsalud.gov.co/sustancias-fiscalizadas/SiteAssets/Paginas/Preguntas-frecuentes/Usuarios-inscritos-para-analisis-de-calidad-(cuantificacion)-investigacion .pdf)
- [12] Fundación CANNA. (s.f.). Cuantificación de la concentración de THC. Fundación CANNA. <https://www.fundacion-canna.es/cuantificacion-de-la-concentracion-de-thc>

- [11] <https://doi.org/10.1371/journal.pone.0236878> Cerrato, A., Citti, C., Cannazza, G., Capriotti, A. L., Cavaliere, C., Grassi, G., Marini, F., Montone, C. M., Paris, R., Piovesana, S., & Laganà, A. (2021). Phytocannabinomics: Untargeted metabolomics as a tool for cannabis chemovar differentiation. *Talanta*, 230. <https://doi.org/10.1016/j.talanta.2021.122313>
- [12] Carvalho, V. M., de Almeida, F. G., de Macêdo Vieira, A. C., Rocha, E. D., Cabral, L. M., & Strongin, R. M. (2022). Chemical profiling of Cannabis varieties cultivated for medical purposes in southeastern Brazil. *Forensic Science International*, 335. <https://doi.org/10.1016/j.forsciint.2022.111309>
- [13] Siracusa, L., Ruberto, G., & Cristino, L. (2023). Recent Research on Cannabis sativa L.: Phytochemistry, New Matrices, Cultivation Techniques, and Recent Updates on Its Brain-Related Effects (2018-2023). *Molecules (Basel, Switzerland)*, 28(8), 3387. <https://doi.org/10.3390/molecules28083387>
- [14] Dr. Ali Raheem Alnassar. (2020). Spearman's correlation. Pag. 1-4. Recuperado de [https://uomustansiriyah.edu.iq/media/lectures/6/6\\_2020\\_05\\_12!02\\_19\\_03\\_PM.pdf](https://uomustansiriyah.edu.iq/media/lectures/6/6_2020_05_12!02_19_03_PM.pdf)
- [15] Mishra, S., Sarkar, U., Taraphder, S., Datta, S., Swain, D., & Saikhom, R. et al. (2017). Multivariate Statistical Data Analysis- Principal Component Analysis (PCA). *International Journal of Livestock Research*, 7(5), 60-78. <http://dx.doi.org/10.5455/ijlr.20170415115235>
- [16] Monti, M. C., Frei, P., Weber, S., Scheurer, E., & Mercer-Chalmers-Bender, K. (2022). Beyond  $\Delta^9$ -tetrahydrocannabinol and cannabidiol: chemical differentiation of cannabis varieties applying targeted and untargeted analysis. *Analytical and Bioanalytical Chemistry*, 414(13), 3847–3862. <https://doi.org/10.1007/s00216-022-04026-2>

## **Anexos**

### **Anexo 1.**

Listado de códigos empleados.

<https://docs.google.com/document/d/1khxgYuHnTw1iIiESvNYRuFJaD-avB9QbxEvETa-7t1Q/edit>

### **Anexo 2.**

Matriz de datos sin procesar.

<https://docs.google.com/spreadsheets/d/1qcu9wyxJP9WpYkB9B8cDh-mFTkUYQubQhS0zstEzDe8/edit?pli=1&gid=1108833027#gid=1108833027>

### **Anexo 3.**

Matriz de datos procesada.

[https://docs.google.com/spreadsheets/d/1hCPSWJdj\\_NUieIfHFYkhIYnAtlumxov6jSK\\_yOfsDr c/edit?gid=2041924464#gid=2041924464](https://docs.google.com/spreadsheets/d/1hCPSWJdj_NUieIfHFYkhIYnAtlumxov6jSK_yOfsDr c/edit?gid=2041924464#gid=2041924464)