# CLASIFICACIÓN MORFOCOLORIMÉTRICA DE SEMILLAS DE ONCE VARIEDADES DE ALUBIAS CULTIVADAS EN ÁLAVA (PAÍS VASCO, ESPAÑA)

**Gorka Castaño Uriarte & Agustí Agut Escrig**

**Banco de Germoplasma Vegetal del Jardín Botánico de Olarizu (Vitoria-Gasteiz)**
**Centro de Estudios Ambientales**
**Dpto. de Medio Ambiente y Espacio Público**
**Ayuntamiento de Vitoria-Gasteiz**
**Casa de la Dehesa de Olarizu, s/n; 01006; Vitoria-Gasteiz**
**gorkacasta@hotmail.com; aagut@vitoria-gasteiz.org**

## 1. Introducción

Aunque los rasgos morfométricos y colorimétricos de semillas se utilizan en los estudios de taxonomía de plantas (*Thellung, 1926; Valdés, 1970; Viano, 1978; Sutton, 1988*), a menudo no se suelen considerar suficientes para establecer clasificaciones entre familias o géneros. No obstante, debido a la disponibilidad de semillas en los Bancos de Germoplasma, la caracterización morfocolorimétrica de semillas es un método rápido, repetible, seguro y no destructivo que puede servir para identificar taxones mediante análisis de imagen. Su utilización en estudios taxonómicos está en auge por su eficacia para discriminar accesiones a nivel interpoblacional *(Mattana et al., 2008)*, o incluso a nivel intrapoblacional (*Dadandi et al., 2009*), gracias a la cada vez más extensa base de datos morfocolorimétrica de semillas (*Bacchetta et al., 2008; Grillo et al., 2010*), lo que permite la identificación de especies en otros grupos taxonómicos (*Bacchetta et al., 2011*). Estudios recientes están abriendo cada vez más perspectivas en la taxonomía de plantas, ofreciendo la oportunidad a los Bancos de Germoplasma de caracterizar de forma estandarizada sus accesiones.

Este estudio trata de establecer un criterio estadístico de clasificación para once variedades de alubias (*Phaseolus vulgaris* L.) cultivadas en la provincia de Álava, basado en las medidas morfocolorimétricas de sus semillas obtenidas mediante técnicas de análisis de imagen. Primero se testó la metodología utilizando ocho variedades de alubia conocidas: *Arrocera Alavesa, Frijol Negro de Basaburua, Roja de Ibeas, Marrón de Musitu, Pinta Alavesa, Redonda de Genevilla, Negra Tolosana* y *Roja Tolosana*. Posteriormente, se introdujeron tres variedades desconocidas (A, B y C) para estudiar, por una parte, si eran clasificadas en alguno de los grupos descritos (variedades conocidas) o como un nuevo grupo, y por otra parte, las relaciones taxonómicas existentes entre las distintas variedades.

## 2. Material y métodos

### 2.1. Calibración de color del escáner

Para obtener unas medidas de color estandarizadas, se realizó una calibración de color de un escáner plano *HP PSC 1610 Series,* según la norma ISO 12641, que exige el uso de una tarjeta IT-8.7/2 (*Kodak Q60*) y el archivo de referencia correspondiente. La tarjeta se escaneó a tamaño 1:1 en formato TIFF, con una resolución de 300 dpi, con todas las opciones de ajuste de color desactivadas. Una vez escaneada la tarjeta se creó un perfil de color para el escáner, utilizando el algoritmo *labclut* del software *Color Camera Calibrator* (*National Library Of Australia*).

### 2.2. Proceso de toma de imágenes

Se obtuvieron muestras de 100 semillas seleccionadas al azar de cada una de las 11 variedades de alubias incluidas en el estudio. Las semillas se colocaron sobre el escáner lo más separadas posible en un área de escaneo máxima de *29 x 21 cm*. Se obtuvieron imágenes con una resolución de 300 dpi y formato TIFF. Para cada variedad se obtuvieron dos imágenes, una sobre fondo oscuro y otra sobre fondo claro, logrando contrastar mejor las regiones claras y oscuras de las semillas (*Bacchetta et.al., 2010*). Para obtener las imágenes sobre fondo oscuro, las semillas fueron cubiertas con una caja recubierta con papel negro opaco. Para obtener las imágenes sobre fondo claro, las semillas fueron cubiertas con una caja recubierta con un papel blanco opaco, minimizando el espacio entre las semillas y el fondo de la caja para evitar sombras. En ambos casos, se evitaron al máximo las interferencias de luz ambiental.

### 2.3. Análisis de imagen

Las imágenes obtenidas se procesaron con el software libre *ImageJ*. Para obtener las medidas morfométricas se optimizó el contraste de las imágenes y posteriormente se realizó una segmentación umbral. El establecimiento de un umbral, permite que una porción se seleccione y otra no, en base a un valor de 0 a 255 para cada canal de gris. Como la mayoría de las semillas presentaban coloraciones claras y oscuras, establecer un umbral en cada imagen que separase el fondo de las semillas fue difícil, por lo que se combinaron las imágenes segmentadas sobre fondo claro y oscuro, empleando el operador *OR* para obtener una imagen única que respetase la dimensión real de las semillas.
Para obtener las medidas colorimétricas, se obtuvieron 3 imágenes en escala de grises para cada canal a partir de la imagen original RGB. Las medidas del modelo HLS se obtuvieron a partir de un *plugin Color Transformer* incluído en la librería de *ImageJ*.

### 2.4. Análisis discriminante

Las medidas morfocolorimétricas fueron sometidas a un análisis discriminante con el paquete estadístico *XLSTAT*. El análisis genera funciones discriminantes a partir de las variables, que se corresponden con las medidas. Cada función discriminante tiene un valor en porcentaje, que es la capacidad que tiene dicha función respecto del resto de funciones discriminantes de aislar un grupo en función de las variables. A su vez, cada variable contribuye en distinto grado a cada función discriminante, en función de dos estadísticos: F de Snedecor y Lambda de Wilks. A mayor valor de F y menor valor de Lambda, mayor es la contribución de cada variable a la función discriminante.

Posteriormente, la matriz de confusión da el porcentaje de acierto entre los grupos discriminados por el análisis y las variedades de alubias. Por último, la validación cruzada nos da el porcentaje de fiabilidad del análisis.

Se realizaron dos análisis discriminantes. El primero estudió si las 8 variedades de alubias descritas y comúnmente reconocidas estaban correctamente clasificadas, estableciendo aproximaciones entre ellas. El segundo estudió 11 variedades, 8 conocidas y 3 desconocidas, considerándolas como grupos distintos para el análisis. De este modo se observó con que porcentaje y fiabilidad consideraba el análisis discriminante a estas variedades, estableciendo aproximaciones taxonómicas entre todos los grupos.

**Tabla 1.** Lista de características morfométricas y colorimétricas obtenidas.

| Característica | Descripción y unidad de medida |
|---|---|
| Área | Área del objeto en mm$^2$ |
| Media | Valor de gris promedio del objeto |
| Std.Dev. | Desviación estándar del valor de gris promedio del objeto |
| Centroide | El punto central del objeto. Promedio de las coordenadas X Y de todos los píxeles del objeto (coordenadas X e Y) |
| Centro de masa | Promedio ponderado de las coordenadas X e Y de todos los píxeles del objeto (coordenadas XM e YM) |
| Perímetro | Perímetro del objeto en mm |
| Rectángulo delimitante | Rectángulo más pequeño que delimita el objeto. Utiliza 4 medidas: Anchura, altura y coordenadas BX y BY pertenecientes a la esquina superior izquierda del rectángulo. |
| Elipse delimitante | Elipse más pequeña que delimita el objeto. Utiliza 3 medidas: Eje mayor, eje menor y ángulo: ángulo entre eje primario y una línea paralela al eje x de la imagen |
| Descriptores de forma | - Circularidad: 4 π x (Área) / (Perímetro)$^2$<br>- Relación de aspecto de la elipse delimitante: Eje mayor / Eje menor<br>- Redondez: Es la inversa de la relación de aspecto. Se calcula así: 4 x Area / π x (ángulo mayor)$^2$<br>- Solidez: Área / zona convexa |
| Feret | La distancia más larga entre dos puntos del objeto (mm) |
| Feret mínimo | La distancia más corta entre dos puntos del objeto (mm) |
| Angulo de Feret | Va de cero a 180º. Ángulo que forma la dirección principal del objeto. El feret y el feret mínimo corresponderían a los catetos de un triángulo rectángulo. |
| FeretX/FeretY | La longitud de la proyección del objeto en la dirección XY |
| Densidad integrada | Suma de los valores de los píxeles del objeto. Equivale al producto del área y el valor gris promedio. |
| Densidad integrada bruta | Suma de todos los valores de los píxeles. |
| Mediana | La mediana del valor de los píxeles del objeto. |
| Asimetría | Asimetría de los valores de gris del objeto |
| Curtosis | Dispersión de los valores de gris del objeto |
| Media R | Valor promedio del canal rojo (modelo RGB)* |
| Std.Dev. R | Desviación estándar del promedio del canal rojo * |
| Media G | Valor promedio del canal verde (modelo RGB) * |
| Std.Dev. G | Desviación estándar del promedio del canal verde * |
| Media B | Valor promedio del canal azul (modelo RGB) * |
| Std.Dev. B | Desviación estándar del promedio del canal azul * |
| Media H | Valor promedio del canal de matiz (modelo HLS) * |
| Std.Dev. H | Desviación estándar del promedio del canal de matiz * |
| Media L | Valor promedio del canal luminosidad (modelo HLS) * |
| Std.Dev. L | Desviación estándar del promedio del canal luminosidad * |
| Media S | Valor promedio del canal saturación (modelo HLS) * |
| Std.Dev. S | Desviación estándar del promedio del canal saturación * |

* En escala de grises

**Bibliografía**

Bacchetta, G., Grillo, O., Mattana, E. & Venora, G. (2008). Morpho-colorimetric characterization by image analysis to identify diaspores of wild plant species. Flora. 203, 669–682.

Bacchetta, G., Grillo, O., Lovicu, G., Orrù, M., Piazza, G., Ravalli, C., Venora, G. (2010) Pips Image Analysis to support Cultivar Identification of Vitis vinifera L. CIGR Workshop on Image Analysis in Agriculture, 26-27.

Bacchetta, G.,et. al. (2011). Seed image analysis provides evidence of taxonomical differentiation within the Lavatera triloba aggregate (Malvaceae).Flora 206 (2011): 468-472.

Dadandi, M.Y., G. Kökdil, A. Ilčim and B. Özbilgin. 2009. Seed macro and micro morphology of the selected Nigella (Ranunculaceae) taxa from Turkey and their systematic significance. Biologia, 64: 261-270.

Grillo, O., Mattana, E., Venora, G. and Bacchetta, G. (2010). Statistical seed classifiers of 10 plant families representative of the Mediterranean vascular flora. Seed Sci. & Technol., 38, 455-476.

Mattana, E, Grillo, O., Venora, G and Bacchetta, G. (2008). Germplasm image analysis of Astragalus maritimus and A. verrucosus of Sardinia (subgen. Trimeniaeus, Fabaceae) Anales del Jardín Botánico de Madrid Vol. 65(1): 149-155 enero-junio 2008 ISSN: 0211-1322.

Sutton, D.A. (1988). A revision of the tribe Antirrhineae. Cambridge University Press. London & Oxford.

Thellung, A. (1926). Daucus (En alemán): Illustrierte Flora von Mittel europa. p. 1501-1526. In: G.Hegi (ed.) Verlag Paul Parey, Berlín.

Valdés, B. (1970). Revisión de las especies de Linaria con semillas aladas. Anales Universidad Hispalense, Series Ciencias, 7. Publicaciones de la Universidad de Sevilla, Sevilla.

Viano, J. (1978). Les linéaires à graines aptères du bassin méditerranéen occidental. 2.Linaria sect. Elegantes, Bipunctatae, Diffusae, Speciosae, Repentes. Candollea 33: 209-267.

Pinta alavesa

Variedad A

Variedad B

Marrón de Musitu

## 3. Resultados

Las pruebas de Box confirmaron la hipótesis de que las matrices de covarianza eran diferentes a un nivel de significancia del 0,05%, por lo que las funciones discriminantes obtenidas fueron cuadráticas. Cada medida morfocolorimétrica fue seleccionada en torno a su valor de F y Lambda de Wilks en un *ranking* de importancia para las funciones discriminantes. En la Tabla 2 y Tabla 3, se muestra el *ranking* de las 20 primeras características morfocolorimétricas para las 8 y las 11 variedades de alubias, respectivamente. Puede observarse que las características más discriminantes para ambos casos son las colorimétricas, especialmente la de los canales luminancia (L), tono (H) y saturación (S); así como la media, mediana y asimetría del valor de los píxeles.

**Tabla 2:** Ranking de 20 características morfocolorimétricas obtenidas en el análisis discriminante de las 8 variedades de alubias.

| Variable | Lambda | F |
|---|---|---|
| Std.Dev. L | 0,007 | 16634,068 |
| Asimetría | 0,043 | 2490,683 |
| Media H | 0,044 | 2473,814 |
| Mediana | 0,051 | 2093,694 |
| Media | 0,059 | 1791,860 |
| Media L | 0,064 | 1643,464 |
| Media S | 0,088 | 1178,000 |
| Media G | 0,095 | 1080,985 |
| Densidad Integrada | 0,111 | 906,801 |
| Media B | 0,140 | 695,848 |
| Std.Dev. B | 0,167 | 563,786 |
| Curtosis | 0,168 | 562,171 |
| Media R | 0,170 | 553,242 |
| Perímetro | 0,193 | 472,749 |
| Std.Dev. G | 0,198 | 458,828 |
| Circularidad | 0,209 | 429,485 |
| Std.Dev R | 0,263 | 316,493 |
| Feret mínimo | 0,269 | 307,004 |
| Eje menor | 0,292 | 274,146 |
| Area | 0,293 | 273,027 |

**Tabla 3:** Ranking de 20 características morfocolorimétricas obtenidas en el análisis discriminante de las 11 variedades de alubias.

| Variable | Lambda | F |
|---|---|---|
| Std.Dev. L | 0,006 | 17122,275 |
| Media H | 0,026 | 4115,086 |
| Mediana | 0,034 | 3140,350 |
| Media | 0,043 | 2434,358 |
| Asimetría | 0,048 | 2170,874 |
| Media L | 0,060 | 1703,112 |
| Media S | 0,072 | 1399,994 |
| Densidad Integrada | 0,111 | 869,881 |
| Media G | 0,116 | 827,262 |
| Media R | 0,181 | 493,529 |
| Perímetro | 0,192 | 458,052 |
| Std.Dev. B | 0,200 | 435,376 |
| Std.Dev. G | 0,215 | 397,956 |
| Circularidad | 0,222 | 381,935 |
| Std.Dev. H | 0,230 | 364,325 |
| Std.Dev R | 0,268 | 298,104 |
| Curtosis | 0,270 | 294,959 |
| Feret | 0,294 | 261,580 |
| Area | 0,304 | 248,846 |
| Feret mínimo | 0,313 | 238,602 |

Se obtuvieron 10 funciones discriminantes para obtener el 100% de la varianza. A continuación (Figura 1), se muestran gráficamente las dos funciones discriminantes principales, que representan el 89,83% de la varianza total entre ambas, y la agrupación de las distintas variedades de alubias en función de la contribución de sus medidas morfocolorimétricas a las dos funciones discriminantes.

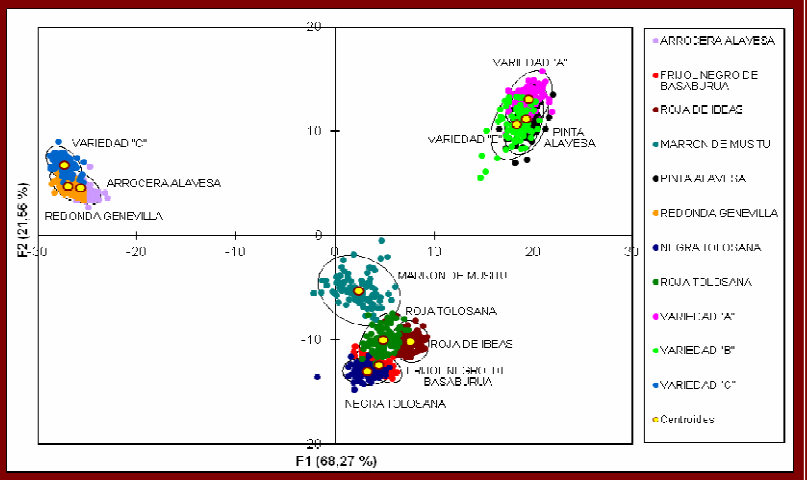


**Figura 1:** Agrupación de las variedades en función de las funciones discriminantes principales.

Arrocera alavesa

Redonda de Genevilla

Variedad C

Roja tolosana

Negra tolosana

Roja de Ibeas

Frijol negro de Basaburua

La Tabla 4 y la Tabla 5 muestran los porcentajes de clasificación correcta obtenidos por el método de validación cruzada para obtener la fiabilidad del análisis discriminante. En la Tabla 4, se observa que para las 8 variedades descritas y comúnmente reconocidas, el análisis las clasifica correctamente con una fiabilidad del 99,75%.

En la Tabla 5, aparecen los porcentajes de clasificación correcta de las 11 variedades (8 variedades conocidas y 3 variedades desconocidas), teniendo en cuenta que las variedades A, B y C se incluyen en el estudio aceptando la hipótesis de que son variedades separadas de las restantes 8 conocidas. Para el conjunto de las once variedades el porcentaje de clasificación correcta fue del 96,09%, considerando las variedades A, B y C como variedades distintas en un 83%, 86% y 98%, respectivamente (Tabla 5).

**Tabla 4**: Clasificación correcta de las 8 variedades de alubias obtenida por validación cruzada (AA: Arrocera alavesa; FNB: Frijol Negro de Basaburua; RG: Redonda de Genevilla; RI: Roja de Ibeas; MM: Marrón de Musitu; PA: Pinta alavesa; NT: Negra Tolosana; RT: Roja Tolosana).

| | AA | FNB | RG | RI | MM | PA | NT | RT | Total | % |
|---|---|---|---|---|---|---|---|---|---|---|
| AA | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 100 |
| FNB | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 100 |
| RG | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 100 | 100 |
| RI | 0 | 0 | 0 | 99 | 0 | 0 | 0 | 1 | 100 | 99 |
| MM | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100 | 100 |
| PA | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 100 | 100 |
| NT | 0 | 0 | 0 | 1 | 0 | 0 | 99 | 0 | 100 | 99 |
| RT | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 100 | 100 |
| Total | 100 | 100 | 100 | 100 | 100 | 100 | 99 | 101 | 800 | 99,75 |

**Tabla 5**: Clasificación correcta de las 11 variedades de alubias obtenida por validación cruzada (AA: Arrocera alavesa; FNB: Frijol Negro de Basaburua; RG: Redonda de Genevilla; RI: Roja de Ibeas; MM: Marrón de Musitu; PA: Pinta alavesa; NT: Negra Tolosana; RT: Roja Tolosana; "A": Variedad A; "B": Variedad B; "C": Variedad C).

| | AA | FNB | RI | MM | PA | RG | NT | RT | "A" | "B" | "C" | Total | % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AA | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 100 |
| FNB | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 100 |
| RI | 0 | 0 | 99 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 100 | 99 |
| MM | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 100 |
| PA | 0 | 0 | 0 | 0 | 92 | 0 | 0 | 0 | 3 | 5 | 0 | 100 | 92 |
| RG | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 100 | 100 |
| NT | 0 | 0 | 1 | 0 | 0 | 0 | 99 | 0 | 0 | 0 | 0 | 100 | 99 |
| RT | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100 | 100 |
| "A" | 0 | 0 | 0 | 0 | 12 | 0 | 0 | 0 | 83 | 5 | 0 | 100 | 83 |
| "B" | 0 | 0 | 0 | 0 | 14 | 0 | 0 | 0 | 0 | 86 | 0 | 100 | 86 |
| "C" | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 98 | 100 | 98 |
| Total | 101 | 100 | 100 | 100 | 118 | 101 | 99 | 101 | 86 | 96 | 98 | 1100 | 96,09 |

## 4. Discusión y conclusiones

La caracterización morfocolorimétrica conjunta de las once variedades permitió observar las relaciones taxonómicas existentes entre ellas, agrupándolas según su semejanza. Las variedades A y B están muy próximas a la variedad *Pinta alavesa*, formando una agrupación separada del resto, indicando la mayor proximidad taxonómica entre ellas. Del mismo modo, la variedad C y las variedades *Redonda de Genevilla* y *Arrocera alavesa* forman una segunda agrupación. El resto de variedades forman una tercera agrupación, quedando la variedad *Marrón de Musitu* como la más alejada de este grupo, siendo la más aislada y *a priori* la que menor semejanza presentaba con el resto.

Se puede concluir que la caracterización morfocolorimétrica de semillas presenta un alto grado de fiabilidad como clasificador estadístico y para el establecimiento de aproximaciones taxonómicas, ofreciendo una potente herramienta de clasificación taxonómica para las variedades de plantas cultivadas y las especies silvestres conservadas en bancos de germoplasma.