

# Evaluación de la diversidad genética de la Monastrell, una variedad antigua en la provincia de Alicante (España) mediante Genotipado por Secuenciación (GBS) / Genotyping by Sequencing (GBS) to evaluate the genetic diversity of Monastrell, an old grapevine variety in the province of Alicante (Spain)

C. Esteras<sup>1</sup>, D. López-Lluch<sup>2</sup>, S. Derdak<sup>3,4</sup>, B. Picó<sup>1</sup>, y J.J. Ruiz<sup>5</sup>

<sup>1</sup> Universitat Politècnica de Valencia, UPV-COMAV, Spain

<sup>2</sup> Dpto. Economía Agroambiental, Ing. Cartográfica y Expresión Gráfica en la Ingeniería, Escuela Politécnica Superior de Orihuela, UMH, Spain

<sup>3</sup> CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Baldori i Reixac 4, 08028 Barcelona, Spain

<sup>4</sup> Universitat Pompeu Fabra (UPF), Barcelona, Spain

<sup>5</sup> Dpto. Biología Aplicada, Escuela Politécnica Superior de Orihuela, UMH, Spain

**Abstract.** Monastrell variety, also known as Mourvèdre, is cultivated mainly in the southeast of Spain in more than 45,000 ha. It is an ancient variety, whose origins are not known with certainty, highly adapted to the dry and warm climate of the southeast Spanish. The existence of a great clonal diversity is well known but, although some clonal selection programs have been carried out (Murcia, Bandol, etc.), the genetic diversity has not been evaluated to date. The measure of this diversity is urgent, since very old vines have been plucked in recent years. We have initiated a project to estimate the genetic diversity of the conserved clones of this variety through GBS (Genotyping by Sequencing). To carry out this analysis, a survey was carried out covering the whole area of Monastrell cultivation in the Alicante Denomination of Origin (D.O.). We selected 46 vines from 46 different vineyards, always prioritizing the oldest vines at each sampling site. The GBS has allowed us to obtain between 0.54 and 1.17 million of reads per sample, of which between 73 and 78% have been mapped to the reference genome developed by the French-Italian consortium (vvinifera.IGGP 12x) with a mapping quality above threshold. More than 100,000 SNPs have been identified, ranging from 14,000 to 43,000 per sample, with a coverage greater than 10X. A high percentage of these variants are polymorphic among Monastrell clones, confirming the existence of significant levels of genetic diversity. In addition, this analysis provides a high number of high quality SNPs well distributed across the genome, suitable for clone genotyping, which will allow the design of strategies to optimize its conservation and use.

## 1. Introducción

La variedad Monastrell se cultiva mayoritariamente en el sureste de España, en más de 45.000 ha (Fig. 1). Se trata de una variedad antigua, cuyos orígenes no se conocen con certeza, muy adaptada al clima seco y cálido del sureste español.

Se tiene conocimiento de la existencia de una gran diversidad clonal, pero, aunque se han llevado a cabo algunos programas de selección clonal (Murcia, Bandol, etc.), no se ha evaluado de forma adecuada la diversidad genética existente. La medida de esta diversidad es urgente, por cuanto se han venido arrancando en los últimos años cepas muy antiguas. En la Denominación de Origen Protegida Vinos de Alicante (DOP Alicante) se cultivan 6.284 ha de Monastrell, que suponen un 66,1% de la superficie cultivada. Hemos iniciado recientemente un proyecto para estimar la diversidad genética existente en los clones conservados de esta variedad mediante

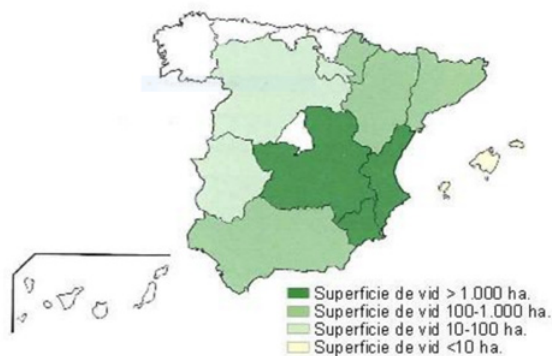
Genotipado por secuenciación (GBS, Genotyping by Sequencing). En este trabajo se presentan los resultados preliminares obtenidos de un muestreo llevado a cabo en las comarcas del Alto y Medio Vinalopó, en Alicante.

## 2. Material y métodos

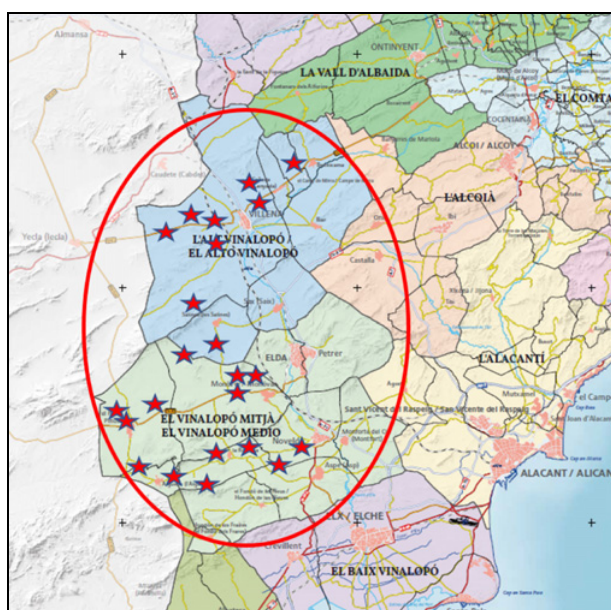
### 2.1. Muestras evaluadas

Para llevar a cabo este análisis, durante el mes de junio de 2016 se realizó una toma de muestras, recorriendo las principales zonas de cultivo de la variedad Monastrell en la DOP Alicante, fundamentalmente campos de cultivo de las comarcas del Alto y Medio Vinalopó (Fig. 2).

Procedentes del muestreo anteriormente citado, se han seleccionado 46 cepas individuales, procedentes de 46 viñedos diferentes, priorizando siempre las cepas más antiguas en cada uno de los sitios de muestreo. No se ha tratado de obtener muestras representativas de todo el



**Figura 1.** Distribución geográfica de la variedad Monastrell en España. (Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente, Gobierno de España.)



**Figura 2.** Distribución aproximada de las muestras de Monastrell analizadas. En cada localidad representada se ha muestreado uno o más viñedos, obteniéndose hasta un total de 46 muestras.

territorio cultivado de la variedad Monastrell, sino que, basándose en el conocimiento de los técnicos de la D.O. y de bodegueros y viticultores locales, se ha buscado siempre obtener muestras de las cepas más antiguas y que pudieran tener posibles características diferenciales.

Cada una de las 46 cepas seleccionadas fue posicionada en un mapa físico además de ser georeferenciada. Se obtuvieron de dos a tres muestras de brotes jóvenes para el posterior análisis de ADN.

## 2.2. Genotipado

## 2.3. Extracción de ADN

El ADN se extrajo de hojas jóvenes de cada una de los genotipos evaluados. La extracción se llevó a cabo empleando el kit DNeasy Plant Mini Kit (Qiagen, Hilden, Germany). Las concentraciones de ADN se midieron mediante un espectrofotómetro Nanodrop ND-1000 v.3.5, seleccionando las muestras con al menos una concentración de 50 ng/ $\mu$ l. Finalmente, 46 muestras

con ADN de alta calidad fueron elegidas para la secuenciación, incluyendo además 6 genotipos duplicados como control. Las muestras se enviaron para su genotipado al CNAG (Centro Nacional de Análisis Genómico, Centre de Regulació Genòmica, Barcelona, Spain). Se empleó un mínimo de 500 ng de ADN para el genotipado de SNPs (Polimorfismos de nucleótido simple).

## 2.4. Genotipado por secuenciación (GBS)

Para la identificación de SNPs entre los genotipos evaluados se empleó la estrategia de Genotipado por secuenciación (Genotyping-by-sequencing, GBS) [1]). Se ensayaron en una prueba piloto con 6 muestras dos endonucleasas (ApeKI y PstI). La enzima ApeKI fue finalmente seleccionada, ya que produjo más fragmentos genómicos secuenciables, que fueron más reproducibles en las muestras evaluadas. Además, los fragmentos obtenidos con ApeKI fueron más diversos y representaron una mayor parte del genoma, dando por tanto lugar a un proceso de identificación de variantes más robusto que el obtenido con la enzima PstI.

Las librerías GBS de las 46 muestras se prepararon empleando la endonucleasa ApeKI (sitio de reconocimiento: G/CWGC) y se secuenciaron mediante la plataforma Illumina HiSeq 2000 (Illumina Inc, San Diego, CA, USA). Las lecturas obtenidas se procesaron, eliminando las secuencias de los adaptadores y se analizaron por su calidad. Las lecturas con menos de 30 pares de bases y que no superaron los umbrales de calidad fueron eliminadas antes de proceder al mapeo e identificación de SNPs. El mapeo se realizó empleando como genoma de referencia la secuencia de *Vitis vinifera* producida por el consorcio público franco-italiano (secuencia de Pinot Noir PN40024, recubrimiento 12X) [2] empleando el programa Bowtie 2 [3]). La identificación de SNPs se realizó empleando el programa GATK (<https://software.broadinstitute.org/gatk/>). Se empleó un umbral de calidad de mapeo de 30 y una cobertura mínima de 5. Los SNPs resultantes se seleccionarán, además, descartando aquellos con más de un 70% de datos faltantes, los no bialélicos, y exigiendo una frecuencia alélica mínima superior al 10%.

## 3. Resultados y discusión

### 3.1. Pérdida de diversidad en la variedad Monastrell

Durante la toma de muestras en campo, recorriendo las distintas zonas de cultivo del Alto y Medio Vinalopó, se constató que en los últimos años está teniendo lugar un acelerado proceso de pérdida de cepas antiguas de la variedad Monastrell. Muchos de los viñedos más antiguos habían sido arrancados muy recientemente, siendo ya relativamente infrecuente encontrar cepas como las de la Fig. 3, con una edad superior a 60–70 años. Algunos de los viticultores consultados refieren la existencia de viñedos con cepas muy antiguas que no han podido ser mantenidas en producción, pero que hubieran sido de interés para su caracterización genética y para su posible inclusión en programas de selección clonal de la variedad o en bancos de germoplasma.



**Figura 3.** Cepa de Monastrell, con una edad estimada de aproximadamente 80 años (cerca de Pinoso, Alicante).

### 3.2. Genotipado por secuenciación

En el presente ensayo, la técnica de genotipado por secuenciación (GBS) ha permitido obtener entre 0.54 y 1.17 millones de lecturas por muestra, de las cuales entre un 73 y un 78% han podido maparse al genoma de referencia desarrollado por el consorcio franco-italiano (vvinifera.IGGP\_12x) [2] con una calidad de mapado superior al umbral.

Entre los resultados preliminares del ensayo destaca el número total de SNPs identificados, que ha sido superior a 100.000. Se obtuvieron entre 14.000 y 43.000 SNPs por muestra, con una cobertura mayor a 10X. Un elevado porcentaje de las variantes detectadas han sido polimórficas entre las cepas de Monastrell evaluadas, lo que confirma la existencia de importantes niveles de diversidad genética. El estudio detallado de la diversidad y estructura genética de todas las muestras analizadas nos proporcionará en breve plazo información de interés para

determinar la importancia de estos niveles de diversidad, la necesidad de ampliar las muestras y zonas de cultivo a analizar, así como posibilitará la comparación con la diversidad genética existente en otras áreas de cultivo de la variedad Monastrell.

Además, este análisis preliminar nos ha proporcionado un elevado número de SNPs de alta calidad y distribuidos por el genoma. Una vez depurados y seleccionados teniendo en cuenta las dificultades asociadas a las especies altamente heterocigotas como la vid [4,5], serán adecuados para el genotipado de clones de esta variedad, lo que permitirá en el futuro diseñar estrategias para optimizar su conservación y uso, por ejemplo, en programas de selección clonal.

La presente investigación está siendo financiada por la Cátedra Cajamar de Economía y Desarrollo Agroambiental de la Universidad Miguel Hernández de Elche. Queremos agradecer la colaboración prestada por la Denominación de Origen Protegida Vinos de Alicante, y en especial la colaboración de Omar Amorós, técnico de la D.O.

### References

- [1] R.J. Elshire, J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler et al. PLoS ONE 6, 5 (2011)
- [2] URGI. 12X.2 version of the grapevine reference genome sequence from The French-Italian Public Consortium (PN40024) 2014 [cited 2016]. Available from: <https://urgi.versailles.inra.fr/Species/Vitis/DataSequences/Genome-sequences>
- [3] B. Langmead, S. Salzberg, Nat. Methods 9, 357 (2012)
- [4] K.E. Hyma, P. Barba, M. Wang, J.P. Londo, C.B. Acharya, S.E. Mitchell, L. Cadle-Davidson, PLoS ONE 10, 8 (2015)
- [5] C. Kim, H. Guo, W. Kong, R. Chandnani, L.S. Shuang, A.H. Paterson, Plant Sci. 242, 14 (2016)