Whole genome sequencing of the Greek olive cultivar Chondrolia Chalkidikis: sequence variants

Olea europaea L. – genome-wide variant data (SNPs and indels)

1. Dataset description

The genome of the Greek olive cultivar Chondrolia Chalkidikis was sequenced (whole genome sequencing, WGS) using a hybrid approach that combined short reads (Illumina NextSeq 2000) with long reads (Oxford Nanopore Technologies), and the genome was assembled de novo. Sequence variants were identified by comparing the assembled genome with the reference genome of *Olea europaea* var. *europaea*. The dataset reports the number of single nucleotide polymorphisms (SNPs), insertions and deletions, the transition/transversion ratio and the spectrum of base substitutions.

Species	*Olea europaea* L.
Cultivar	Chondrolia Chalkidikis (Greece)
Data type	Whole genome sequencing (WGS), hybrid de novo assembly
Sequencing platforms	Illumina NextSeq 2000 (short reads); Oxford Nanopore Technologies (long reads)
Reference genome	*Olea europaea* var. *europaea*
Variants	19,767,308 in total (15,715,535 SNPs; 4,051,773 indels)

2. Variant summary

Table 1. Number of sequence variants by category

Variant category	No. of variants	% of total
Single nucleotide polymorphisms (SNPs)	15,715,535	79.50
Insertions	3,319,974	16.80
Deletions	731,799	3.70
Total	**19,767,308**	**100.00**

Table 2. Transitions and transversions among SNPs

Type of change	No. of SNPs	% of SNPs
Transitions (Ts)	10,559,293	67.19
Transversions (Tv)	5,156,242	32.81
Total SNPs	**15,715,535**	**100.00**
Ts/Tv ratio	2.048	

3. Base substitution spectrum

Table 3. Number of SNPs by reference allele (rows) and alternative allele (columns)

Reference allele	→ A	→ C	→ G	→ T	Total
A	–	646,908	**2,586,883**	878,603	4,112,394
C	661,105	–	390,702	**2,688,239**	3,740,046
G	**2,691,182**	390,121	–	659,471	3,740,774
T	880,505	**2,592,989**	648,827	–	4,122,321
Total	**4,232,792**	**3,630,018**	**3,626,412**	**4,226,313**	**15,715,535**

Transitions (A↔G, C↔T) are shown in bold.

Table 4. SNPs grouped into the six substitution classes (complementary changes combined)

Substitution	Type	No. of SNPs	% of SNPs
C>T / G>A	Transition	5,379,421	34.23
A>G / T>C	Transition	5,179,872	32.96
A>T / T>A	Transversion	1,759,108	11.19
C>A / G>T	Transversion	1,320,576	8.40
A>C / T>G	Transversion	1,295,735	8.24
C>G / G>C	Transversion	780,823	4.97
Total		**15,715,535**	**100.00**