# A Genetic Linkage Map of Olive Based on Amplified Fragment Length Polymorphism, Intersimple Sequence Repeat and Simple Sequence Repeat Markers

# Bouchaib Khadari<sup>1</sup>, Amal Zine El Aabidine, and Cinderella Grout

Institut National de la Recherche Agronomique (INRA), Unité Mixte de Recherche (UMR) 1098, Développement et Amélioration des Plantes (DAP), Bât. 3, Campus du Centre International de la Recherche Agronomique pour le Développement (CIRAD) TA A 96/03, Av. Agropolis, 34398 Montpellier Cedex 5, France

## Inès Ben Sadok

INRA, UMR 1098 DAP, Bât. 3, Campus du CIRAD TA A 96/03, Av. Agropolis, 34398 Montpellier Cedex 5, France; and Montpellier SupAgro, UMR 1098 DAP, Bât. 3, Campus CIRAD TA A 96/03, Av. Agropolis, 34398 Montpellier Cedex 5, France

#### **Agnès Doligez**

INRA, UMR 1097, Diversité et Adaptation des Plantes Cultivées (DiA-PC), Bât. 33, 2 place Viala, 34060 Montpellier Cedex 1, France

#### **Nathalie Moutier**

INRA, UMR 1098 DAP, Bât. 3, Campus du CIRAD TA A 96/03, Av. Agropolis, 34398 Montpellier Cedex 5, France

#### Sylvain Santoni

INRA, UMR 1097 DiA-PC, Bât. 33, 2 place Viala, 34060 Montpellier Cedex 1, France

#### **Evelyne Costes**

INRA, UMR 1098 DAP, Bât. 3, Campus du CIRAD TA A 96/03, Av. Agropolis, 34398 Montpellier Cedex 5, France

ADDITIONAL INDEX WORDS. Olea europaea, F1 pseudo-test cross, consensus map, 'Olivière', 'Arbequina', microsatellite loci, dominant markers

ABSTRACT. A detailed genomic linkage map of the olive [*Olea europaea* L. ssp. *europaea* (2x = 2n = 46)] was constructed with a 147 F1 full-sib 'Olivière' × 'Arbequina' progeny in a two-way pseudo-test cross-mapping configuration. Based on a logarithm of odds threshold of 6 and a maximum recombination fraction of 0.4, maternal and paternal maps were constructed using 222 makers [178 amplified fragment length polymorphism (AFLP), 37 simple sequence repeat (SSR), seven intersimple sequence repeat (ISSR)] and 219 markers (174 AFLP, 39 SSR, 6 ISSR) markers, respectively. The female map regrouped 36 linkage groups (LGs) defining 2210.2 cM of total map length with an average marker spacing 11.2 cM and a maximum gap of 48.5 cM between adjacent markers. The male map contained 31 LGs and covered a distance of 1966.2 cM with an average and a maximum distance between two adjacent markers of 10.3 and 40.4 cM, respectively. Mean LG size was 61.3 and 63.4 cM in the maternal and paternal maps, respectively. The LGs consisted of two to 17 loci (up to 21 loci in the paternal map) and ranged in length from 2.7 to 182 cM (female map) or from 4.1 to 218.1 cM (paternal map). Markers were distributed throughout the maps without any clustering. The total length of the consensus map was 3823.2 cM containing 436 markers distributed into 42 LGs with a mean distance between two adjacent loci of 8.7 cM. Both parental maps and the consensus maps were compared with previously published olive maps. Although not saturated yet, the present maps offer a promising tool for quantitative trait loci mapping because phenotypic characterization of the cross is currently carried out.

Olive is one of the most important Mediterranean fruit species cultivated for oil and canned fruit consumption. With over 2000 cultivars, the olive tree is highly diverse and adapted to Mediterranean agroecological conditions. It plays a major socioeconomical role in traditional agroecosystems, especially in southern Mediterranean areas with an expanding use in oil and cosmetic industries (Aburjai and Natsheh, 2003).

Olive is a diploid species [2n = 46 (Green and Wickens, 1989)] with a nuclear DNA content ranging from  $2.90 \pm 0.020$ 

548

Received for publication 21 July 2010. Accepted for publication 4 Oct. 2010. We are grateful to Prof. P. Villemur who made the 'Olivière' × 'Arbequina' cross and to Prof F. Dosba for her support. This work was performed in the laboratory of "Atelier de Marquage Moléculaire"; we acknowledge all team members for their technical support. We also ackowledge J. Charafi (INRA Morocco) for his help in SSR analysis. It was financed by the Department of Genetics and Plant Breeding INRA and by PRAD 03-06. Comments on an earlier version of the manuscript from A. Bervillé, Ph. Chatelet, and J.P. Perros were greatly appreciated. <sup>1</sup>Corresponding author. E-mail: khadari@supagro.inra.fr.

# to $3.07 \pm 0.018$ pg/2C (Loureiro et al., 2007). These values correspond to $\approx 3120$ Mbp because 1 pg is $\approx 978$ Mbp based on the conversion factor proposed by Dolezel et al. (2003). Despite its high economical value, few genomic studies have been performed on the olive compared with other Mediterranean crops such as grapevine [*Vitis vinifera* L. (Troggio et al., 2007)].

Most Mediterranean olive cultivars have been selected over several centuries and rarely have been subjected to deep genetic characterization. Breeders have selected cultivars integrating various improved agronomic traits such as fruit yield or fruit and oil quality. However, selection for elite cultivars in olive remains a long (at least 10 years) and expensive procedure. The use of molecular breeding will allow saving time and increasing selection efficiency. Linkage maps of progenies segregating for important economic traits such as fruit quality and yield are required to develop marker-assisted selection in olive-breeding programs.

Most of the established linkage maps in plants have been achieved using segregating populations resulting from crosses between inbred lines. Obtaining such populations in forest and fruit trees is limited as a result of significant genetic load and time constraints (Venkateswarlu et al., 2006). For this reason, an F1 population in woody perennial plants may be compared with the F2 or BC1 generation of annual self-fertilized crops (Dondini et al., 2004). As a result of the longevity of woody perennials plants, it is possible to maintain segregating populations in various environments for a long period and to score them for multiple phenotypic traits over time (Venkateswarlu et al., 2006).

The strategy frequently used to map F1 populations in tree species is the two-way pseudo-test cross-mapping strategy (Grattapaglia and Sederoff, 1994) because it was efficient for mapping eucalyptus species (Eucalyptus grandis W. Hill ex Maiden and E. urophylla S.T. Blake) and several heterozygous species such as maritime pine (Pinus pinaster Aiton) (Chagné et al., 2002; Scott et al., 2005). In this strategy, dominant molecular markers are considered because in highly heterozygous species, numerous alleles are present in one parent only. Both ISSR and AFLP have been reported to be reliable dominant markers in many species. AFLP markers are highly polymorphic, robust, and tend to be distributed throughout plant genomes (Casasoli et al., 2001). SSR markers are ubiquitous, locus-specific, codominants, highly polymorphic, and well distributed throughout the genome. They are efficient markers for genetic mapping and for genetic map comparisons (De La Rosa et al., 2003).

In olive, only two genetic maps have been published to date: 'Leccino'  $\times$  'Dolce Agogia' by De La Rosa et al. (2003) and 'Frantoio'  $\times$  'Kalamata' by Wu et al. (2004). These maps are mostly based on randomly amplified polymorphism DNA and AFLP markers including only five and six codominant SSR markers, respectively.

Construction of a linkage map with consistent molecular markers constitutes the key starting point for the analysis of agronomically important traits. The objective of our investigation was to construct a genetic linkage map of the olive using a large set of AFLP and SSR markers in a pseudo-test crossmapping configuration. The population derived from a cross between the cultivars Olivière and Arbequina was chosen for establishing a genetic linkage map because they are genetically distant (Khadari et al., 2003).

#### **Materials and Methods**

**PLANT MATERIAL AND DNA ISOLATION.** A population of 147 F1 individuals, derived from a 1981 cross between 'Olivière' [OLV (female parent)] and 'Arbequina' [ARB (male parent)], grown at the experimental station of Institut National de la Recherche Agronomique at Montpellier, France, was used to establish the genetic linkage map. OLV is a French cultivar displaying cytoplasmic male sterility (Besnard et al., 2000) and shows very high vigor. OLV is highly productive with medium fruit size and low oil content (Moutier et al., 2004). ARB is a Spanish cultivar with very low vigor. ARB is very productive with very small fruit and a high oil content (Rallo et al., 2005). In previous studies using SSR markers, Khadari et al. (2003) observed that these cultivars were genetically distinct.

DNA was isolated from 100 mg of fresh leaf tissue from each offspring and parents, according to the procedure previously described by Khadari et al. (2008).

**AMPLIFIED FRAGMENT LENGTH POLYMORPHISM MARKERS.** The AFLP procedure was performed as outlined by Vos et al. (1995) with the minor modifications proposed in the AFLP Core Reagent Kit (Invitrogen, Cergy Pontoise, France). Four EcoRI primers and seven MseI primers were used to amplify AFLP markers among 23 primer combinations of EcoRI/MseI. The pre-amplification polymerase chain reaction (PCR) was done using primers carrying one selective nucleotide (*MseI*-A, EcoRI-C). The selective amplification was carried out using primers carrying three overhanging nucleotides of *MseI/EcoRI* (+3/+3). The *EcoRI*+3 primers were 5'-labeled with one of the three fluorophores, 6FAM, NED, or HEX.

The preamplification reaction was cycled 20 times for 1 min at 94 °C, at 56 °C for 1 min, and at 72 °C for 1 min. Preamplification products were diluted 10 times and stored at -20 °C. The selective amplification consisted in a touchdown-PCR in which the annealing temperature was 65 °C for the first cycle and then reduced by 0.7 °C for each of the next 12 cycles. The denaturing and extension stages for each cycle were 94 °C for 10 s and 72 °C for 90 s, respectively. This procedure was followed by 25 cycles of 94 °C for 10 s, 56 °C for 40 + 1 s per cycle, and 72 °C for 90 s.

Amplified fragments were visualized on an ABI Prism 3130 XL and scored semiautomatically with the GeneMapper 3.7 software (Applied Biosystems, Carlsbad, CA).

SIMPLE SEQUENCE REPEAT MARKERS. The primer sequences for the SSR markers used in this study were developed in olive by Carriero et al. (2002), Cipriani et al. (2002), De La Rosa et al. (2002), Díaz et al. (2006), Rallo et al. (2000), Sabino et al. (2006), and Sefc et al. (2000). The SSR markers developed in the related species Phillyrea angustifolia L. by Saumitou-Laprade et al. (2000) and in common ash (Fraxinus excelsior L.) by Brachet et al. (1999) were also tested. Among the 90 SSRs initially screened, 44 were used in this study because of their polymorphism between the two parent cultivars and six progenies. SSR primers providing a clear polymorphism were 5'-labeled with one of three fluorophores (6FAM, NED, or HEX) and were analyzed on an ABI Prism 3130 XL Genetic Analyser. The PCR amplification procedure was described by Khadari et al. (2008). Allele sizes were assigned using the GeneMapper 3.7 software (Applied Biosystems).

INTERSIMPLE SEQUENCE REPEAT MARKERS. We used three ISSR primers: the IMA855 primer previously used by Vargas and Kadereit (2001) and two primers, UBC818 and UBC841Y,

commercialized by the University of British Columbia (Vancouver, BC, Canada). PCR amplifications were carried out according to Khadari et al. (2005).

NAMING MARKERS. The AFLP markers generated were identified by a primer pair combination, the three selective EcoRI nucleotides followed by a slash, the three selective *MseI* nucleotides, and the fragment size. Only unambiguous AFLP polymorphic bands were used and scored for presence or absence in the segregating population. SSR markers were designated by their published names and ISSR markers by the name identifying the primers followed by estimated fragment size.

SEGREGATION ANALYSIS AND LINKAGE MAPPING. Polymorphic markers were scored as present or absent on the mapping population. Chi squares were calculated for individual markers to detect any deviation ( $P \le 0.05$ ) of marker segregation from the expected Mendelian ratio. Both distorted and undistorted markers were used for map construction, except distorted ones that affected linkage group definition.

and number of markers obtained for each amplified fragment length polymorphism primer combination used for olive genetic mapping based on the 'Olivière' (OLV) ×

According to the two-way pseudo-test cross-mapping strategy, markers were first separated into three data matrices: 1) with markers showing segregation only in the female parent OLV; 2) with markers segregating only in the male parent ARB; and 3) with markers segregating 1:1 and 3:1. The two first matrices were used to construct separate genetic linkage maps for OLV and ARB using CarthaGene software (De Givry et al., 2005). Initially, the construction of separate parental maps was tested and computed using different combinations of criteria to define LGs with a logarithm of odds (LOD) threshold ranging from 3 to 6 and a maximum recombination fraction varying from 0.3 to 0.5. Within each LG, the best order of markers and the genetic distances were found by using the following functions of Cartha-Gene: Build, Annealing, Taboo, Genetic, Flips, and Bestprint. The Kosambi mapping function was applied to convert recombination rates into map distances in centiMorgans. MapChart 2.1 software (Voorrips, 2002) was used to draw the linkage map.

The third data matrix, containing all markers, was used to construct the integrated linkage map. The marker phase was deduced from the parental maps for most markers segregating 1:1; whereas for markers segregating in the 3:1 configuration and those with unknown phases, all possible phases, were tested and the best fitting one was chosen. Average marker spacing was calculated by dividing the total length of each map by the number of intervals.

Table 1. Prime 'Arbequina	r code and nui ' (ARB) cross	mber of marke	rs obtained for eac	ch amplified fragi	ment length poly	morphism prime	r combination u	sed for olive ge	netic mapping	based on the 'Oliv	/ière' (OLV) ×
		-		Segregating	Segregating	Segregating	Distorted	Distorted	Distorted	Total of	
Primer	3' selective EcoRI	nucleotides Msel	Polymorphic bands (no.)	markers in OLV (no.)	markers in ARB (no.)	markers 3:1 (no.)	markers in OLV (no.)	markers in ARB (no.)	markers 3:1 (no.)	distorted markers (no.)	Distorted markers (%)
AGA/CAG	AGA	CAG	6	5	4	0	0	0	0	0	0.0
AGA/CTT	AGA	CTT	12	С	С	9	1	0	С	4	33.3
AGA/CTG	AGA	CTG	25	7	8	10	2	2	5	6	36.0
ACT/CTT	ACT	CTT	12	4	ŝ	5	0	0	0	0	0.0
ACG/CAT	ACG	CAT	26	7	10	6	1	2	1	4	15.3
ACG/CAA	ACG	CAA	28	6	10	6	б	9	2	11	39.2
ACG/CAC	ACG	CAC	20	9	6	5	2	1	4	7	35.0
ACG/CTT	ACG	CTT	23	10	9	7	1	2	б	9	26.0
ACG/CAG	ACG	CAG	22	9	6	7	2	С	2	L	31.8
ACT/CAC	ACT	CAC	14	7	С	4	0	1	0	1	7.1
AGA/CAA	AGA	CAA	10	4	1	5	1	0	1	2	20.0
AGA/CAC	AGA	CAC	13	9	4	ŝ	1	2	1	4	30.7
AGA/CAT	AGA	CAT	16	5	7	4	2	2	0	4	25.0
AGA/CTC	AGA	CTC	14	4	5	5	0	0	б	ю	21.4
ACG/CTC	ACG	CTC	25	8	8	6	1	4	б	8	32.0
ACT/CAT	ACT	CAT	6	ю	С	б	1	1	1	ю	33.3
ACT/CTC	ACT	CTC	12	4	б	5	1	0	0	1	8.3
ACA/CAC	ACA	CAC	27	11	8	8	0	0	2	2	7.4
ACA/CAT	ACA	CAT	44	6	20	15	1	4	1	9	13.6
ACA/CTC	ACA	CTC	28	10	12	9	0	1	0	1	3.5
ACA/CTG	ACA	CTG	34	13	15	9	2	2	б	7	20.5
ACA/CAA	ACA	CAA	39	18	15	9	1	0	1	2	5.1
ACA/CAC	ACA	CAC	32	15	12	5	2	2	1	5	15.6
		Total	494	174	178	142	25	35	37	97	

ANALYSIS OF DISTRIBUTION MARKERS. The distribution of markers throughout the linkage map was analyzed using the method of Remington et al. (1999). Marker clustering was checked by testing each linkage group for the expected number of markers ( $\lambda$ i) using the formula:  $\lambda i = mGi/\sum Gi$  where m is the total number of markers and Gi is the map distance of each linkage group adjusted for chromosome ends as Gi = Mi+2s where Mi is the map distance between terminal markers and s is the average marker spacing for the linkage map.

**OBSERVED AND EXPECTED GENOME LENGTHS AND MAP COVERAGE.** Average marker spacing for each LG was calculated by dividing the total length of the LG by the number of intervals. The expected genome length (centiMorgans) was estimated following Method 4 of Chakravarti et al. (1991): the length of each linkage group was multiplied by (m + 1)/(m - 1), where *m* is the number of markers on each linkage group. Observed genome coverage was assessed by dividing the observed genome length by the expected genome length.

#### **Results and Discussion**

**AMPLIFIED FRAGMENT LENGTH POLYMORPHISM MARKER GENOTYPING.** Based on the number of polymorphic markers revealed on a subset of six individuals in the progeny and on both parents, 23 highly polymorphic primer pairs were chosen among all tested combinations (Table 1). The AFLP genotyping analysis was carried out on the entire progeny and the parents, generating a total of 494 polymorphic markers, of which 352 (71%) were segregating 1:1 and 142 (28%) were segregating 3:1. Of the 352 markers scored as heterozygous (present in one parent only), 174 markers were heterozygous for the female parent (35%) and 178 markers for the male parent (36%; Table 1).

AFLP appeared particularly suitable for constructing our linkage map. The number of polymorphic loci distinguished by a given primer combination varied from nine to 44 with an average value of eight for both the female and male maps (Table 1). These results are partly the result of marker type and the level of genetic polymorphism noted within the OLV  $\times$ ARB mapped population. Similar results were obtained in the 'Picholine Marocaine' × 'Picholine du Languedoc' for segregating olive population (Charafi et al., 2009). In the present study, 29% of the polymorphic AFLP markers generated were found in both parents. The high percentage of markers segregating 3:1 was similar to those found for other intraspecific crosses in olive (Charafi et al., 2009). The level of polymorphism detected by AFLP markers in olive is similar to other woody perennial species such as peach [Prunus persica (L.) Batsch (Dirlewanger et al., 1998)], sweet cherry [Prunus avium L. (Olmstead et al., 2008)], or grapevine (Doligez et al., 2002) in which an average of 6.1, nine, and 16 markers per primer combinations were identified, respectively.

SIMPLE SEQUENCE REPEAT AND INTERSIMPLE SEQUENCE REPEAT MARKER GENOTYPING. Parental screening revealed 44 polymorphic SSR markers out of the 90 loci tested. In the entire full-sib progeny, 26 loci segregated 1:1:1:1, seven loci segregated 1:2:1 (of which four loci with a null allele), and 11 segregated 1:1 (Table 2). Despite the small number of SSR markers, their high degree of polymorphism and their transferability within or even across species guarantees their extensive applicability in olive genome mapping studies. The ISSR analysis generated 15 markers: seven markers (47%) segregating in the female parent OLV, six markers (40%) in the male parent ARB, and two markers (13%) in both parents.

SEGREGATION DISTORTION. For AFLPs, a chi-square test was performed to verify the null hypothesis of a 1:1 or 3:1

Table 2. Simple sequence repeat loci name and parental alleles used for genetic mapping in 'Olivière' × 'Arbequina' olive progeny.

genetie mapping	in Onviere × moequina or	ive progeny.
Locus	Parent genotype	Segregation type
IAS-oli11 <sup>z,y</sup>	133/167, 151/155	1:1:1:1
IAS-oli12 <sup>z</sup>	127/null allele, 136/144	1:1:1:1
IAS-oli17 <sup>z</sup>	171/189, 183/189	1:1:1:1
ssrOeUA-DCA11 <sup>y, x</sup>	141/179, 131/179	1:1:1:1
ssrOeUA-DCA14 <sup>x</sup>	198/209, 209/209	1:1
ssrOeUA-DCA15 <sup>x</sup>	242/262, 242/242	1:1
ssrOeUA-DCA16 <sup>x</sup>	121/144, 144/172	1:1:1:1
ssrOeUA-DCA17 <sup>x</sup>	111/111, 111/113	1:1
ssrOeUA-DCA18 <sup>x</sup>	166/176, 174/176	1:1:1:1
ssrOeUA-DCA3 <sup>x</sup>	229/241, 229/247	1:1:1:1
ssrOeUA-DCA5 <sup>x</sup>	199/203, 203/203	1:1
ssrOeUA-DCA7 <sup>x</sup>	128/null allele, 126/145	1:1:1:1
ssrOeUA-DCA9 <sup>x</sup>	183/205, 185/193	1:1:1:1
EMO13 <sup>w</sup>	120/139, 120/122	1:1:1:1
EMO2 <sup>y,w</sup>	203/311, 211/223	1:1:1:1
EMO88 <sup>w</sup>	187/null allele, 187/191	1:1:1:1
EMO90 <sup>w</sup>	185/192, 185/192	1:1:1:1
GAPU11 <sup>v</sup>	184/206, 206/217	1:1:1:1
GAPU12 <sup>v</sup>	124/null allele, 128/null allele	e 1:2:1
GAPU103A <sup>v</sup>	134/150, 150/160	1:1:1:1
GAPU47 <sup>v</sup>	181/22, 195/22	1:1:1:1
GAPU59 <sup>v</sup>	213/213, 223/null allele	1:2:1
GAPU71B <sup>v</sup>	122/141, 119/122	1:1:1:1
GAPU89 <sup>v</sup>	160/160, 160/175	1:1
IAS-oli22 <sup>y,u</sup>	133/135, 133/135	1:2:1
IAS-oli23 <sup>u</sup>	220/235, 222/235	1:1:1:1
IAS-oli26 <sup>u</sup>	186/200, 189/ 190	1:1:1:1
IAS-oli28 <sup>u</sup>	178/178, 178/182	1:1
ME30MS <sup>t</sup>	216/233, 212/ 233	1:1:1:1
PA(ATT)2 <sup>y,s</sup>	109/121, 112/112	1:1
PA(GA)2 <sup>s</sup>	113/null allele, 113/115	1:2:1
UDO-014 <sup>y, r</sup>	121/null allele, 102/null allele	e 1:2:1
UDO-015 <sup>r</sup>	102/105, 105/105	1:1
UDO-017 <sup>r</sup>	156/168, 152/154	1:1:1:1
UDO-025 <sup>r</sup>	171/189, 177/187	1:1:1:1
UDO-027 <sup>y, r</sup>	119/196, 191/191	1:1
UDO-042 <sup>r</sup>	144/147, 141/144	1:1
UDO-012 <sup>y, r</sup>	156/166, 158/166	1:1:1:1
UDO-024 <sup>r</sup>	193/209, 225/null allele	1:1:1:1
UDO-031 <sup>r</sup>	110/142, 117/110	1:1:1:1
UDO-034 <sup>r</sup>	90/null allele, 79/90	1:2:1
UDO-036 <sup>s</sup>	147/152, 144/152	1:1:1:1
UDO-005 <sup>s</sup>	157/159, 159/159	1:1
UDO-006 <sup>y, s</sup>	162/null allele, 182/null allele	e 1:2:1

<sup>z</sup>Rallo et al. (2000).

<sup>y</sup>Distorted markers at the 5% level.

<sup>x</sup>Sefc et al. (2000).

"De La Rosa et al. (2002).

- <sup>v</sup>Carriero et al. (2002).
- <sup>u</sup>Díaz et al. (2006).

<sup>t</sup>Brachet al. (1999).

<sup>s</sup>Saumitou-Laprade et al. (2000).

<sup>r</sup>Cipriani et al. (2002).

segregation ratio for marker alleles. At the 5% level, 25 (26%) and 35 (36%) AFLP loci were distorted in the female and the male parents, respectively. A total of 37 AFLP markers segregating in both parents presented a distorted segregation (Table 1). Among the 44 SSR loci used, nine loci presented distorted segregation (Table 2). Concerning the ISSR markers, no segregation distortion was noticed.

Segregation distortion is recurrent in outcrossing species (Ky et al., 2000; Nikaido et al., 1999) and was previously observed in olive (De La Rosa et al., 2003; Wu et al., 2004). In the present study, because the parents of our cross are either male sterile, OLV, or self-incompatible, ARB (Díaz et al., 2006; Moutier, 2002), the mechanisms preventing self-fertilization may have generated a stress during the prezygotic phase resulting in an unbalance in fertilized gametes as already suggested in olive by Wu et al. (2004).

Because marker distortion may have biological origins, markers displaying significant deviation from Mendelian segregation were not automatically excluded. We choose to eliminate only distorted markers affecting linkage group definition. In our study, the proportion of distorted dominant markers (P < 0.05) is 19%. A similar proportion of distorted markers was reported by De La Rosa et al. (2003) and Wu et al. (2004) in other mapping olive hybrid populations. Marker distortion was also noted in other perennial fruit species such as grapevine [1.67%] (Doligez et al., 2002)] and apricot [Prunus armeniaca L., 13% (Hurtado et al., 2002)]. Generally, a lower level of segregation distortion is expected in pseudo-test cross populations compared with F2 or BC1 progenies (Tavoletti et al., 1996; Venkateswarlu et al., 2006) and, accordingly, the map construction in such a population is efficient (Tavoletti et al., 1996).

**LINKAGE ANALYSIS.** To select the best LOD and recombination fraction thresholds, we analyzed data varying LOD from 3 to 6 and recombination fraction from 0.3 to 0.5. In both parental maps, under the recombination fraction of 0.5, the number of LGs increased from 3 to 6 LOD. It did not depend on LOD value with a recombination fraction (rec) of 0.3 or 0.4, but it was lower for rec = 0.4 than for rec = 0.3 (data not shown, Supplementary Fig. S1 online). Therefore, a minimum LOD score of 6 and a maximum recombination fraction of 0.4 were chosen as a good compromise between reliable statistical criteria and the number of LGs obtained.

The linkage analysis in the female parent OLV was based on 222 markers, including 178 AFLPs, 37 SSRs, and seven ISSRs. Based on the chosen threshold and recombination values, the map was defined by 36 linkage groups composed of 197 markers. The mean distance between two markers was 11.2 cM with a maximum distance of 48.5 cM between ACG/CAA-172 and IMA888-398 in the LG OLV35 and an average number of 5.4 markers per LG (data not shown, Supplementary Fig. S5 online). The linkage distance between the markers on our map varied greatly across the different LGs. LG size ranged from 2.7 to 182 cM with an average of 61.3 cM and did not necessarily reflect the number of linked markers. For instance, in the LG OLV40, a total linkage distance of 42.8 cM corresponded to seven markers, whereas in LG OLV50, a 134-cM distance (approximately three times more) was covered by only 10 markers (data not shown, Supplementary Table S1 online).

The male parent ARB map was based on a total of 219 markers, including 174 AFLPs, 39 SSRs, and six ISSRs. Using the LOD threshold of 6, 191 markers were assigned to 31 LGs. The mean distance between two markers was 10.3 cM with a maximum of 40.4 cM between UDO34 and ACA/CTG-297 in the LG *ARB30* and an average number of 6.1 markers per LG. Linkage group size ranged from 4.1 to 218.1 cM with an average of 63.4 cM (data not shown, Supplementary Fig. S6 and Table S2 online).

The markers are well scattered in the obtained maps. The olive genome length, based on total map length, was 1577.6 and 1529.1 cM for the female and male maps, respectively (Table 3). Wu et al. (2004) reported 759 cM for 'Kalamata' and 798 cM for 'Frantoio', whereas De La Rosa et al. (2003) observed 2765 cM for the 'Leccino' map and 2445 cM for the 'Dolce Agogia' map. These large differences in genome length between studies do not seem to be mainly the result of differences in the average distance between adjacent markers, because average distances are all within a quite narrow range [10.2 to 13.2 cM (Table 3)]. The observed discrepancies between our parents or with parents of other populations may be partly the result of the percentage of genome coverage, marker type, recombination rates, distorted loci proportion, and the heterozygosity level between the parents (Remington et al., 1999).

Table 3. Comparison between the	different available	maps developed on olive.
---------------------------------	---------------------	--------------------------

			Maximum					Mean distance	Mean size
	Population	LOD	recombination	Linkage	Map	Mapped	Mapped	between	of linkage
Crossed cultivars	size	$threshold^z$	fraction	group (no.)	length (cM)	loci (no.)	SSR (no.)	loci (cM)	group (cM)
Frantoio <sup>y</sup>	104	3	0.49	27	798	92	3	12.3	29.5
Kalamata <sup>y</sup>	104	3	0.49	23	759	89	1	11.5	33.0
Integrated mapy	104	3	0.49	15	879	101	5	10.2	58.6
Leccino <sup>x</sup>	100	4	0.30	39	2765	249	3	13.2	70.9
Dolce agogia <sup>x</sup>	100	4	0.30	30	2445	236	4	11.9	81.5
Integrated map <sup>x</sup>	—	—		—		—	—		—
Olivière <sup>w</sup>	147	6	0.40	34	2210.20	197	34	11.2	61.3
Arabequina <sup>w</sup>	147	6	0.40	31	1966.20	191	30	10.3	63.4
Integrated map <sup>w</sup>	147	6	0.40	42	3823.20	436	26	8.7	91.0

<sup>z</sup>Logarithm of odds.

<sup>y</sup>Wu et al. (2004).

<sup>x</sup>De La Rosa et al. (2003).

<sup>w</sup>Present study.

Compared with our parental maps, the number of LGs obtained by Wu et al. (2004) is closer to the number of haploid chromosomes. This result does not indicate, however, the saturated state of their maps because this may result from the limited number of mapped markers (less than 101 markers) and the relaxed criteria for LG definition (maximum recombination fraction = 0.49, LOD threshold = 3). Using strict grouping criteria, we obtained a number of LGs similar than that reported by De La Rosa et al. (2003). This number was higher than the 23 haploid chromosomes even when mapping an at least twofold number of markers used by Wu et al. [2004 (Table 3)]. Few SSR markers were common to our maps and those obtained by Wu et al. [2004 (DCA3, UDO44, UDO11, and UDO06 common SSR loci)] and by De La Rosa et al. [2003 (ME3OMS, PAATT2, and PAGA2 common SSR loci)]. This limited number markers shared among the three studies does not allow comparative mapping.

The observed number of markers per group was tested using a two-tailed cumulative Poisson distribution with parameter  $\lambda i$ [P < 0.025 are significant at  $\alpha = 0.05$  (Ukrainetz et al., 2008; data not shown; Supplementary Fig. S2, S3 and S4 online)]. Analysis of marker distribution in the parental and the integrated maps showed a significant excess of markers in 21 LGs displaying high significance for the two-tailed probability value (P > 0.90), for example, OLV1, OLV3, ARB1, and C34. In general, fewer markers than expected were noticed in the smaller groups, whereas more markers than expected were in the larger groups. Similar results were reported by Ukrainetz et al. (2008) in a mapped douglas fir [*Pseudotsuga menziesii* (Mirb.) Franco] population.

Distorted markers were dispersed over nine, seven, and 18 linkage groups in maternal, paternal, and integrated maps, respectively. Interestingly, 37% of distorted mapped markers in the consensus map was concentrated on three LGs [C26, C38, and C40 (Fig. 1 and data not shown, Supplementary Table S3 online)]. Similar results were reported by Thoquet et al. (2002) in barrel medic (*Medicago truncatula* Gaertn.) in which distorted markers were limited to three LGs only. This tendency of distorted markers to be positioned on the same group suggests either structural origin for this distortion (Thoquet et al., 2002) or selection acting on some genes (Wu et al., 2004).

For the construction of the consensus map, among the 545 available markers, a total of 436 markers were mapped (26 SSR, 403 AFLP, and seven ISSR). The markers were assigned to 42 groups and covered 3823.2 cM. The mean distance between two adjacent markers was 8.7 cM with a maximum of 81 cM between UDO36 and DCA16 in the LG C40 and an average number of 10.3 markers per LG. The LG size ranged from 2.7 to 270.7 cM with an average of 91.0 cM (Fig. 1). Markers segregating 3:1 were not sufficient to regroup small LGs found in parental maps. In addition, we noted that ARB31 and C42 LG length were considerable [218.1 and 270.7 cM, respectively (data not shown, Supplementary Tables S2 and S3 online)]. These LGs were validated by reanalyzing the 2-point LOD (mrklod2p) for all pairs of loci and eliminating loci one by one to verify linkage consistency.

Marker order in the consensus map was consistent with marker order in each parental map testifying of consensus map good reliability. Markers segregating as 1:1 or 1:3 allowed us to connect groups in the consensus map. Two direct consequences of such connections are 1) the number of linkage groups identified in the consensus map (42 LGs) was inferior to the sum of the number of linkage group of both parental maps (77); and 2) the



Fig. 1. An integrated genetic linkage map for the 'Olivière' (OLV)  $\times$  'Arbequina' (ARB) olive cross (see Fig. 1 in color online). Map distances were derived using the Kosambi mapping function. Markers in red and blue indicate maternal and paternal markers, respectively. Bold characters indicate simple sequence repeat (SSR) markers. SSR markers mapped in both maternal and paternal maps are indicated in black. Distorted markers are underlined.

map distance of the consensus map was smaller than the sum of map distances of both parental maps. However, in the consensus map, the smallest linkage groups did not originate from the merging of the parental maps but already existed in the maternal or paternal map. The number of intercross markers remains insufficient to establish connection between all parental linkage groups, mainly the smallest ones, and which probably correspond to a portion of linkage groups. These observations indicated that our maps are not saturated and the necessity of additional markers to reach saturation. Based on the formula developed by Beckman and Soller (1983) and according to Wu et al. (2004),  $\approx$ 360 and 500 markers would be required to saturate our genetic linkage map at 80% and 95%, respectively, with an average distance between markers of 10 cM and a map length of 3000 cM.

**MAP DISTANCE.** The estimated map distance corresponding to the total length map was 2210.2 cM for the female parent and 1966.2 cM for the male parent and 3823.2 cM for the consensus map. Differences in the estimated distances of both parental maps may reflect differences in the recombination frequencies of both parents. Putative causes for the difference between the two estimated parental maps genome are: 1) marker distribution along the chromosome varies between parents; and 2) male and female gametes probably display different recombination frequencies.

The consensus map described in the present study is a substantial result offering an efficient tool for genetic analysis even if it is only partly saturated. Its robustness results from the use of a large population of 147 individuals and a high number of markers (436), whereas the previously published consensus map in olive has been constructed using 104 individuals and 101 markers only (Wu et al., 2004). Furthermore, the map construction criteria we used were more constraining (LOD = 6, rec = 0.40) than in this previous study [LOD = 3, rec = 0.49 (Table 3)].

### Conclusion

Additional markers are obviously required to expand the coverage of parental and consensus maps of the studied segregating population, because the number of LGs is still much higher than the haploid number of chromosomes. However, the present maps already offer a promising tool for quantitative trait loci (QTLs) mapping in olive. The mapped segregating population  $OLV \times ARB$  is currently phenotyped for growth and tree architecture to detect QTLs related to these traits.

#### Literature Cited

- Aburjai, T. and F.M. Natsheh. 2003. Review article: Plants used in cosmetics. Phytother. Res. 17:987–1000.
- Beckman, J. and M. Soller. 1983. Restriction fragment length polymorphism in genetic improvement: Methodologies, mapping and costs. Theor. Appl. Genet. 67:35–43.
- Besnard, G., B. Khadari, P. Villemur, and A. Bervillé. 2000. Cytoplasmic male sterility in the olive (*Olea europaea* L.). Theor. Appl. Genet. 100:1018–1024.
- Brachet, S., M.F. Jubier, M. Richard, B. Jung-Muller, and N. Frascaria-Lacoste. 1999. Rapid identification of microsatellite loci using 5' anchored PCR in the common ash *Fraxinus excelsior*. Mol. Ecol. 8: 160–163.
- Carriero, F., G. Fontanazza, F. Cellini, and G. Giorio. 2002. Identification of simple sequence repeats (SSRs) in olive (*Olea europaea* L.). Theor. Appl. Genet. 104:301–307.
- Casasoli, M., C. Mattioni, M. Cherubini, and F. Villani. 2001. A genetic linkage map of european chestnut (*Castanea sativa* Mill.)

based on RAPD ISSR and isozyme markers. Theor. Appl. Genet. 102:1190–1199.

- Chagné, D., C. Lalannea, D. Madura, S. Kumarb, J.M. Frigérioa, C. Kriera, S. Decroocqa, A. Savouréc, M. Bou-Dagher-Kharratc, E. Bertocchia, J. Bracha, and C. Plomion. 2002. A high density genetic map of maritime pine based on AFLPs. Ann. For. Sci. 59: 627–636.
- Chakravarti, A., L.K. Lasher, and J.E. Reefer. 1991. A maximum likelihood method for estimating genome length using genetic linkage data. Genetics 128:175–182.
- Charafi, J., A.Zine El Aabidine, C. Grout, B. Rahioui, A. El Meziane, A. Moukhli, B. Boulouha, C. El Modafar, and B. Khadari. 2009. A genetic linkage map of *Olea europaea* L. using a pseudo-test-cross mapping strategy based on SSR, AFLP, ISSR, RAPD and SCAR markers. Acta Hort. 814:609–614.
- Cipriani, G., M.T. Marrazzo, R. Marconi, A. Cimato, and R. Testolin. 2002. Microsatellite markers isolated in olive (*Olea europaea* L) are suitable for individual fingerprinting and reveal polymorphism within ancient cultivars. Theor. Appl. Genet. 104:223–228.
- De Givry, S., M. Bouchez, P. Chabrier, D. Milan, and T. Schiex. 2005. CarthaGene: Multipopulation integrated genetic and radiation hybrid mapping. Bioinformatics 21:1703–1704.
- De La Rosa, R., A. Angiolillo, C. Guerrero, M. Pellegrini, L. Rallo, G. Besnard, A. Bervillé, A. Martin, and L. Baldoni. 2003. A first linkage map of olive (*Olea europaea* L.) cultivars using RAPD AFLP and SSR markers. Theor. Appl. Genet. 106:1273–1282.
- De La Rosa, R., C.M. James, and K.R. Tobutt. 2002. Isolation and characterization of polymorphic microsatellites in olive (*Olea europaea* L.) and their transferability to other genera in the Oleaceae. Mol. Ecol. Notes 2:265–267.
- Díaz, A., R. De la Rosa, A. Martín, and P. Rallo. 2006. Development, characterization and inheritance of new microsatellites in olive (*Olea europaea* L.) and evaluation of their usefulness in cultivar identification and genetic relationship studies. Theor. Appl. Genet. 2:165–175.
- Dirlewanger, E., V. Pronier, C. Parvery, C. Rothan, A. Guye, and R. Monet. 1998. Genetic linkage map of peach [*Prunus persica* (L.) *Batsch*] using morphological and molecular markers. Theor. Appl. Genet. 97:888–895.
- Dolezel, J., J. Bartos, H. Voglmay, and J. Greilhuber. 2003. Nuclear DNA content and genome size of trout and human. Cytometry 51: 127–128.
- Doligez, A., A. Bouquet, Y. Danglot, F. Lahogue, S. Riaz, C.P. Meredith, K.J. Edwards, and P. This. 2002. Genetic mapping of grapevine (*Vitis vinifera L.*) applied to the detection of QTLs for seedlessness and berry weight. Theor. Appl. Genet. 105:780–795.
- Dondini, L., L. Pierantoni, F. Gaiotti, R. Chiodini, S. Tartarini, C. Bazzi, and S. Sansavini. 2004. Identifying QTLs for fire-blight resistance via a european pear (*Pyrus communis* L.) genetic linkage map. Mol. Breed. 14:407–418.
- Grattapaglia, D. and R.R. Sederoff. 1994. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross mapping strategy and RAPD markers. Genetics 137:1121– 1137.
- Green, P.S. and G.E. Wickens. 1989. The *Olea europaea* complex, p. 287–299. In: K. Tan, R.R. Milland, and T.S. Elias (eds.). Edinburgh University Press, Edinburgh, UK.
- Hurtado, M.A., C. Romero, S. Vilanova, A.G. Abbott, G. Llácer, and M.L. Badenes. 2002. Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.), and mapping of PPV (sharka) resistance. Theor. Appl. Genet. 105:182–191.
- Khadari, B., C. Breton, N. Moutier, J.P. Roger, G. Besnard, A. Bervillé, and F. Dosba. 2003. The use of molecular markers for germplasm management in a French olive collection. Theor. Appl. Genet. 106:521–529.
- Khadari, B., J. Charafi, A. Moukhli, and M. Ater. 2008. Substantial genetic diversity in cultivated Moroccan olive despite a single major cultivar: A paradoxical situation evidenced by the use of SSR loci. Tree Genet. Genomes 4:213–221.

- Khadari, B., A. Oukabli, M. Ater, A. Mamouni, J.P. Roger, and F. Kjellberg. 2005. Molecular characterization of Moroccan fig germplasm using intersimple sequence repeat and simple sequence repeat markers to establish a reference collection. HortScience 40: 29–32.
- Ky, C.L., P. Barre, M. Lorieux, P. Trouslot, S. Akaffou, J. Louarn, A. Charrier, S. Hamon, and M. Noirot. 2000. Interspecific genetic linkage map segregation distortion and genetic conversion in coffee (*Coffea* sp.). Theor. Appl. Genet. 101:669–676.
- Loureiro, J., E. Rodriguez, A. Costa, and C. Santos. 2007. Nuclear DNA content estimations in wild olive (*Olea europaea L ssp europaea var sylvestris Brot*) and Portuguese cultivars of *O. europaea* using flow cytometry. Genet. Resources Crop Evol. 54:21–25.
- Moutier, N. 2002. Self-fertility and inter-compatibilities of sixteen olive varieties. Acta Hort. 586:209–212.
- Moutier, N., C. Pinatel, A. Martre, J.P. Roger, B. Khadari, J.F. Burgevin, D. Ollivier, and J. Artaud. 2004. Identification et caractérisation des variétés d'olivier cultivées en France–Tome 1. Naturalia, Turriers, France.
- Nikaido, A., H. Yoshimaru, Y. Tsumura, Y. Suyama, M. Murai, and K. Nagasaka. 1999. Segregation distortion for AFLP markers in *Cryptomeria japonica*. Genes Genet. Syst. 74:55–59.
- Olmstead, J.W., A.M. Sebolt, A. Cabrera, S.S. Sooriyapathirana, S. Hammar, G. Iriarte, D. Wang, C.Y. Chen, E. van der Knaap, and A.F. Iezzoni. 2008. Construction of an intra-specific sweet cherry (*Prunus avium* L.) genetic linkage map and synteny analysis with the *Prunus* reference map. Tree Genet. Genomes 4:897–910.
- Rallo, L., D. Barranco, J.M. Caballero, C. Del Rio, A. Martin, J. Tous, and I. Trujillo. 2005. Las variedades de olivo cultivadas en España. Consejería de Agricultura y Pesca, Ministerio de Agricultura, Pesca y Alimentación. Mundi-Prensa, Madrid, Spain.
- Rallo, P., G. Dorado, and A. Martin. 2000. Development of simple sequence repeats (SSRs) in olive tree (*Olea europaea* L.). Theor. Appl. Genet. 101:984–989.
- Remington, D.L., R.W. Whetten, B.H. Liu, and D.M. O'Malley. 1999. Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. Theor. Appl. Genet. 98:1279–1292.
- Sabino, G.F., M. Busconi, A. Da Câmara Machad, and C. Fogher. 2006. Development and characterization of microsatellite loci from *Olea europaea*. Mol. Ecol. 6:1275–1277.
- Saumitou–Laprade, P., C. Vassiliadis, J.T. Epplen, and C. Hardt. 2000. Isolation of microsatellites loci for paternity testing in *Phillyrea angustifolia* L (Oleaceae). Mol. Ecol. Notes 9:112–114.

- Scott, L.J., M. Shepherd, D.G. Nikles, and R.J. Henry. 2005. Low efficiency of pseudo-test-cross mapping design was consistent with limited genetic diversity and low heterozygosity in hoop pine (*Araucaria cunninghamii* Araucariaceae). Tree Genet. Genomes 1:124–134.
- Sefc, K.M., M.S. Lopes, D. Mendonça, M. Rodrigues Dos Santos, M. Laimer Da Câmara Machado, and A. Da Câmara Machado. 2000. Identification of microsatellite loci in olive (*Olea europaea*) and their characterization in Italian and Iberian olive trees. Mol. Ecol. Notes 9:1171–1173.
- Tavoletti, S., F. Veronesi, and T.C. Osborn. 1996. RFLP linkage map of an alfalfa meiotic mutant based on an F1 population. J. Hered. 87: 167–170.
- Thoquet, P., M. Ghérardi, E.P. Journet, A. Kereszt, J.M. Ané, J.M. Prosperi, and T. Huguet. 2002. The molecular genetic linkage map of the model *Medicago trunctula*: An essential tool for comparative legume genetics and isolation of agronomically important genes. BMC Plant Biol. 2:1.
- Troggio, M., G. Malacarne, G. Coppola, C. Segala, A.D. Cartwright, M. Pindo, M. Stefanini, R. Mank, M. Moroldo, M. Morgante, M.S. Grando, and R. Velasco. 2007. A dense SNP-based genetic linkage map of grapevine (*Vitis vinifera* L) anchoring Pinot Noir BAC contigs. Genetics 176:2637–2650.
- Ukrainetz, N.K., K. Ritland, and S.D. Mansfield. 2008. An AFLP linkage map for douglas-fir based upon multiple full-sib families. Tree Genet. Genomes 4:181–191.
- Vargas, P. and J.W. Kadereit. 2001. Molecular fingerprinting evidence (ISSR inter-simple sequence repeats) for a wild status of *Olea europaea* L (Oleaceae) in the Eurosiberian North of the Iberian Peninsula. Flora 196:142–152.
- Venkateswarlu, M., S. Raje Urs, B. Surendra Nath, H.E. Shashidhar, M. Maheswaran, T.M. Veeraiah, and M.G. Sabitha. 2006. A first genetic linkage map of mulberry (*Morus* spp.) using RAPD, ISSR, and SSR markers and pseudotestcross mapping strategy. Tree Genet. Genomes 3:15–24.
- Voorrips, R. 2002. MapChart: Software for the graphical presentation of linkage maps and QTLs. J. Hered. 93:77–78.
- Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Fritjers, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: A new concept for DNA fingerprinting. Nucleic Acids Res. 23: 4407–4414.
- Wu, S.B., G. Collins, and M. Sedgley. 2004. A molecular linkage map of olive (*Olea europaea* L) based on RAPD microsatellite and SCAR markers. Genome 47:26–35.