

Construction of a reference molecular linkage map of globe artichoke (*Cynara cardunculus* var. *scolymus*)

E. Portis · G. Mauromicale · R. Mauro ·
A. Acquadro · D. Scaglione · S. Lanteri

Received: 6 April 2009 / Accepted: 11 September 2009 / Published online: 29 September 2009
© Springer-Verlag 2009

Abstract The genome organization of globe artichoke (*Cynara cardunculus* var. *scolymus*), unlike other species belonging to Asteraceae (=Compositae) family (i.e. sunflower, lettuce and chicory), remains largely unexplored. The species is highly heterozygous and suffers marked inbreeding depression when forced to self-fertilize. Thus a two-way pseudo-testcross represents the optimal strategy for linkage analysis. Here, we report linkage maps based on the progeny of a cross between globe artichoke (*C. cardunculus* var. *scolymus*) and cultivated cardoon (*C. cardunculus* var. *altilis*). The population was genotyped using a variety of PCR-based marker platforms, resulting in the identification of 708 testcross markers suitable for map construction. The male map consisted of 177 loci arranged in 17 major linkage groups, spanning 1,015.5 cM, while female map was built with 326 loci arranged into 20 major linkage groups, spanning 1,486.8 cM. The presence of 84 loci shared between these maps and those previously developed from a cross within globe artichoke allowed for

map alignment and the definition of 17 homologous linkage groups, corresponding to the haploid number of the species. This will provide a favourable property for QTL scanning; furthermore, as 25 mapped markers (8%) correspond to coding regions, it has an additional value as functional map and might represent an important genetic tool for candidate gene studies in globe artichoke.

Introduction

Genetic maps provide a powerful means of analysing the inheritance of agronomic traits, many of which are under polygenic or oligogenic control. The establishment of linkage relationships between molecular marker loci represents the initial step in the identification of which chromosomal regions carry genes relevant for marker-assisted breeding applications. The Asteraceae (Compositae) species *Cynara cardunculus* L. is native to the Mediterranean basin, and incorporates the taxa globe artichoke (*C. cardunculus* var. *scolymus*), cultivated cardoon (var. *altilis*) and wild cardoon (var. *sylvestris*); the latter is by far the most widely distributed of these, thriving in warm, dry and low altitude environments. The globe artichoke makes a significant contribution to the agricultural economy of southern Europe, and is also cultivated across North Africa, the Middle East, South America, the USA and China (<http://faostat.fao.org/>). The cultivated cardoon is of some regional importance to local cuisine in Italy, Spain and southern France. The three taxa are sexually compatible with one another, producing fertile inter-taxon F₁ hybrids. Thus, the wild and cultivated cardoons together represent a natural genetic resource for globe artichoke improvement. The species is an out-breeder, and is characteristically highly heterozygous. Its marked level of inbreeding

Communicated by M. Havey.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-009-1159-2) contains supplementary material, which is available to authorized users.

E. Portis · A. Acquadro · D. Scaglione · S. Lanteri (✉)
DIVAPRA Plant Genetics and Breeding,
University of Torino, via L. da Vinci 44,
10095 Grugliasco (Turin), Italy
e-mail: sergio.lanteri@unito.it

G. Mauromicale · R. Mauro
Dipartimento di Scienze Agronomiche,
Agrochimiche e delle Produzioni Animali,
sez. Scienze Agronomiche, University of Catania,
via Valdisavoia 5, 95123 Catania, Italy

depression inhibits the use of backcross, F_2 or recombinant inbred populations for mapping purposes. As haploid induction—via either andro- or gynogenesis—has not been achievable till now (Chatelet et al. 2005; Motzo and Deidda 1993; Stamigna et al. 2005), no possibility is presently available to generate doubled haploid populations. Thus, genetic mapping in globe artichoke has had to rely on a double pseudo-testcross approach, in which segregating F_1 progeny is derived from a cross between two heterozygous individuals.

The first genetic maps of *C. cardunculus* were provided by Lanteri et al. (2006), based on a cross between ‘Romanesco C3’ (a late-maturing non-spiny type as female) and ‘Spinoso di Palermo’ (an early-maturing spiny type as male). This population was genotyped using a number of PCR-based marker platforms, resulting in a $\sim 1,300$ cM female map consisting of 204 loci, divided into 18 linkage groups (LGs) and a $\sim 1,200$ cM male map comprising 180 loci and 17 LGs. The two maps shared 78 loci, which allowed for the alignment of 16 of the LGs. The maps have since been extended by the inclusion of a number of microsatellite loci, of which 19 were represented in both maps (Acquadro et al. 2009). Three genes involved in the synthesis of caffeoylquinic acid have also been positioned on the maps (Comino et al. 2007, 2009; Moglia et al. 2009).

The aims of the research described in this paper were to develop a genetic map based on a cross between the same female parent as previously (‘Romanesco C3’) and an accession of cultivated cardoon, and to align this map with the one already established.

Materials and methods

Plant material and DNA isolation

The mapping population was a set of F_1 progeny derived from a controlled cross between the globe artichoke ‘Romanesco C3’ and the cultivated cardoon ‘Altalis 41’ genotypes. Hybrid seeds were germinated in lightly moistened potting mix at room temperature. Emergence occurred within 10–12 days, and healthy seedlings were transferred to the field after about 30 days, at which time the seedlings had developed three true leaves. Two weeks after transplanting in the field, DNA was extracted following Lanteri et al. (2001) and the hybridity of each progeny was confirmed by genotyping with two microsatellite markers. Of the 154 true hybrids, a random selection of 94 was used for marker linkage analysis and genetic map construction, and the presence/absence of spines was scored on well-developed leaves over two seasons (2006 and 2007).

Marker analysis

AFLP profiling followed Vos et al. (1995) with minor modifications as described by Lanteri et al. (2004). Parental and progeny DNA was digested with either *EcoRI/MseI* (E/M), *PstI/MseI* (P/M) or *EcoRI/TaqI* (E/T), and ligated to standard adaptors. The ligation reaction was used as a template for pre-amplification, using primers complementary to the adaptor sequences plus one selective nucleotide (*EcoRI* + A, *PstI* + A, *MseI* + C, *TaqI* + T). Selective amplification was performed using primers with two or three selective nucleotides. The same E/M and P/M AFLP primer combinations (PCs) were used as those chosen by Lanteri et al. (2006). A further 20 E/T PCs (five *EcoRI* primers \times four *TaqI* primers) were applied to the parental and six of the progeny DNA templates in order to extend the number of informative PCs to 48 (18 E/M, 19 P/M and 11 E/T; see Table 1). Amplicons were electrophoretically resolved through 5% denaturing polyacrylamide gels and silver stained as described by Bassam et al. (1991).

The S-SAP fingerprinting method followed Waugh et al. (1997). The selective amplification used a fluorescence-labelled (IRD-700) cyre5 primer, designed to anneal to the long terminal repeat of a retroelement (Acquadro et al. 2006), in combination with an unlabelled AFLP primer with three selective nucleotides (Table 1). PCR products were separated on a DNA analyser Gene ReadIR 4200 (LI-COR) through 6.5% polyacrylamide gels (Sigma), as described by Jackson and Matthews (2000). The S-SAP primer pairs used by Lanteri et al. (2006) were implemented by the three most polymorphic E/T-based PCs (Table 1).

A set of 114 microsatellite primer pairs was used to amplify template of the two parents and six progeny. Of these, 93 have been developed by our laboratory (Acquadro et al. 2003, 2005a, b, 2009). Eight SSRs were developed from the genic sequences SST (sucrose-sucrose 1-fructosyltransferase), cyprosin (an aspartic proteinase), and phenylalanine ammonia-lyase (PAL) (Acquadro et al. 2003; Sonnante et al. 2008). Of the 61 genomic CELMS microsatellites (Acquadro et al. 2009), 27 contained at least one open reading frame and have been assumed to target genic sequence. Primer pairs identifying polymorphism in this screen were applied to the full mapping population. The PCR conditions were as detailed by Acquadro et al. (2005a), and the amplicons were separated as described above for the S-SAP markers.

Single nucleotide polymorphism (SNP) assays were directed at the two acyltransferase genes HCT and HQT (Comino et al. 2007, 2009) and the hydroxylase gene C3/H (Moglia et al. 2009), assayed by a tetra-primer ARMS-PCR method (Ye et al. 2001; Chiapparino et al. 2004) and separated by 2% agarose gel electrophoresis.

Table 1 AFLP and SSAP primer combinations used for linkage analysis

<i>EcoRI/MseI</i> template		<i>EcoRI/TaqI</i> template		<i>PstI/MseI</i> template	
PC	Code	PC	Code	PC	Code
AFLP					
E + AAT/M + CAG	e34/m49	E + AAC/T + TAC	e32/t80	P + AC/M + CAA	p12/m47
E + AAT/M + CAT	e34/m50	E + AAC/T + TAG	e32/t81	P + AC/M + CAT	p12/m50
E + ACA/M + CAA	e35/m47	E + AAC/T + TAT	e32/t82	P + AC/M + CTA	p12/m59
E + ACA/M + CAC	e35/m48	E + AAG/T + TAC	e33/t80	P + AC/M + CTC	p12/m60
E + ACA/M + CAG	e35/m49	E + AAG/T + TGG	e33/t89	P + AC/M + CTG	p12/m61
E + ACA/M + CAT	e35/m50	E + ACA/T + TAC	e35/t80	P + AC/M + CTT	p12/m62
E + ACA/M + CTT	e35/m62	E + ACA/T + TAG	e35/t81	P + AG/M + CAA	p13/m47
E + ACC/M + CAA	e36/m47	E + ACA/T + TGG	e35/t89	P + AG/M + CAT	p13/m50
E + ACC/M + CAC	e36/m48	E + ACT/T + TAC	e38/t80	P + AG/M + CTA	p13/m59
E + ACC/M + CTA	e36/m59	E + ACT/T + TAT	e38/t82	P + AG/M + CTC	p13/m60
E + ACG/M + CAG	e37/m49	E + AGA/T + TAC	e39/t80	P + AG/M + CTG	p13/m61
E + ACG/M + CAT	e37/m50			P + AG/M + CTT	p13/m62
E + ACG/M + CTG	e37/m61			P + ATG/M + CAA	p45/m47
E + ACT/M + CAA	e38/m47			P + ATG/M + CAT	p45/m50
E + ACT/M + CAT	e38/m50			P + ATG/M + CCA	p45/m51
E + ACT/M + CTA	e38/m59			P + ATG/M + CTA	p45/m59
E + AGA/M + CAT	e39/m50			P + ATG/M + CTC	p45/m60
E + AGT/M + CAT	e42/m50			P + ATG/M + CTG	p45/m61
				P + ATG/M + CTT	p45/m62
S-SAP					
Cyre5/E + AAG	cyre5/e33	Cyre5/T + TGA	cyre5/t87	Cyre5/P + AGT	cyre5/p42
Cyre5/E + ACA	cyre5/e35	Cyre5/T + TGG	cyre5/t89		
Cyre5/M + CAA	cyre5/m47	Cyre5/T + TGT	cyre5/t90		
Cyre5/M + CAC	cyre5/m48				
Cyre5/M + CAG	cyre5/m49				
Cyre5/M + CAT	cyre5/m50				

All electrophoretic profiles were documented using the Quantity One Programme gel documentation system (BioRad). Each PCR was replicated once, and only unambiguous fragments were considered. AFLP and S-SAP profiles were treated as sets of dominant markers (presence/absence of fragment), and each marker was named according to the PC used to generate it, followed by the estimated molecular weight of the fragment, e.g. e33/t89-166 is an AFLP fragment of 166-bp length amplified by primer pair Eco + AAG/Taq + TGG. Microsatellite and SNP loci were identified by the primer pair used for their assay.

Linkage analysis and map construction

Markers were grouped as either maternal testcross markers (segregating only within ‘Romanesco C3’) with an expected monogenic segregation ratio of 1:1; paternal testcross markers (segregating only within ‘Altisil 41’); or

intercross markers (segregating in both parents), with an expected segregation ratio of 3:1 for dominant markers, and 1:2:1 for co-dominant ones. Co-dominant markers showing three or four alleles (one parent *ab*, the other either *ac* or *cd*), giving an expected segregation ratio of 1:1:1:1, were initially converted into 1:1 markers, according to the parental origin of the segregating alleles, and so were included within either the maternal or paternal data sets.

Independent framework linkage maps were constructed for each parent on the basis of the double pseudo-testcross mapping strategy (Weeden 1994), with only those markers in the testcross configuration being considered. One linkage map was generated for ‘Romanesco C3’, and a second for ‘Altisil 41’, by applying JoinMap v4.0 (Van Ooijen 2006) and treating the populations as a backcross. Goodness-of fit between observed and expected segregations was assessed using the χ^2 test. Markers fitting a Mendelian pattern closely ($\chi^2 \leq \chi^2_{\alpha=0.1}$) or with only a minor deviation

($\chi^2_{\alpha=0.1} < \chi^2 \leq \chi^2_{\alpha=0.01}$) were used for map construction and for the estimation of genetic distances, when their presence did not alter surrounding marker order in the LG. Those for which the deviation was highly significant ($\chi^2 > \chi^2_{\alpha=0.01}$) were not immediately excluded but were handled with caution in subsequent analyses (see further). Markers with missing data for more than 30 of the 94 F_1 individuals were excluded. For both maps, LGs were established on the basis of an initial threshold logarithm of odds ratio (LOD) of 6.0, with parameters set as follows: Rec = 0.40, LOD = 1.0, Jump = 5.

Once the framework maps had been established, the intercross markers were added as accessory markers, by treating the population as a cross-pollination type. Intercross markers were associated with particular LGs, but were not used for the estimation of genetic distances. Additional accessory markers were subsequently added by lowering the LOD threshold to 4.0, with the inclusion of the ones departing from the Mendelian ratio at the $\alpha > 0.001$ level of confidence. They were not forced in the map to avoid potential artefact (major variation in marker orders and relative distances), but checked one by one and placed in their most likely position within LGs.

Markers deviating in their segregation only marginally from the expected Mendelian ratio were identified with one ($\chi^2_{\alpha=0.1} < \chi^2 \leq \chi^2_{\alpha=0.05}$), two ($\chi^2_{\alpha=0.05} < \chi^2 \leq \chi^2_{\alpha=0.01}$), or three ($\chi^2 > \chi^2_{\alpha=0.01}$) asterisks (Fig. 1). LGs on the female map were named LG_C3, and those on the male map LG_Alt, with both numbered serially in descending order of genetic length (Fig. 1; Table 2).

Results

Genotyping

The 48 AFLP PCs produced 719 informative markers, of which 638 (89%) were testcross markers, and the remainder intercross markers. Per AFLP PC, the number of informative fragments ranged from 5 to 34, with a mean of 15. The most informative enzyme combination was E/T and the least P/M. The number of S-SAP markers ranged from two to eight per PC (mean 4.2), producing 42 mappable markers (38 testcross, four intercross).

Of the combined 676 AFLP and S-SAP testcross markers, 400 (59%) segregated in 'Romanesco C3' and 276 (41%) in 'Altilis 41'. About 13% produced distorted segregation ratios ($\chi^2 > \chi^2_{\alpha=0.1}$), resulting in the discarding of 51 markers [44 testcross (22 for each parent) and seven intercross].

Of the 114 microsatellite primer pairs, 56 segregated in the F_1 population for at least one parent with major differences in their informativeness (70% of the CELMS

markers, only 7% of the CsLib and none of the CsEST loci). Of the 22 which segregated in both parents, 21 segregated consistently with a 1:1:1:1 and the remaining one with a 1:2:1 ratio. The other 34 loci segregated within only one of the parents (29 female and 5 male only). Five loci suffered from minor segregation distortion ($\chi^2_{\alpha=0.1} < \chi^2 \leq \chi^2_{\alpha=0.05}$).

Only the female parent was heterozygous for the SNPs HCTsnp97, HQTsnp359 and C3H'snp447, all of which segregated consistently with a 1:1 ratio.

Phenotyping

The presence/absence of spines is controlled by a single gene with the two alleles *Sp* (dominant non-spiny) and *sp* (recessive spiny) (Lanteri et al. 2006). Spines were absent from both parents. As the character segregated consistently among the F_1 progeny with a 3:1 ratio ($\chi^2 = 0.35$), both parents were clearly of genotype *Spsp*. To locate this gene within an LG, the data were treated as resulting from the segregation of a dominant intercross marker.

Map construction

From the 431 loci (359 AFLP, 19 S-SAP, 50 microsatellite and three SNP) segregating in the female parent, a map was generated which consisted of 326 loci grouped into 20 LGs (each with four or more loci), for a total genetic length of 1,486.8 cM (Table 2; Fig. 1a). A further nine loci formed one triplet and three doublets. LG length varied from 35.0 to 138.7 cM, with the largest containing 49 loci. The bulk (77%) of map intervals were less than 7 cM, with only eight gaps of >15 cM. To evaluate the reliability of the LGs, for each marker pairs with gaps >10 cM, the specific LOD in support of the two-point placement was checked and reported in the Electronic supplementary Table S1. The E/M, E/T and P/M AFLP loci were evenly distributed, without any noticeable clustering of loci generated by any one PC. Only two LGs (C3_6 and _18) were composed solely of AFLP loci. The 46 microsatellite loci which mapped within an LG were distributed over 17 LGs. The 15 loci for which segregation was marginally distorted, mapped to nine LGs (eight at $\alpha = 0.05$ and seven at $\alpha = 0.01$). None of these distorted loci were linked to one another.

The male map was based on segregation at 280 loci (246 AFLP, eight S-SAP and 26 microsatellite). Here, 176 loci were distributed over 17 LGs, covering 1,015.5 cM, with 4 triplets and 11 doublets (Table 2; Fig. 1b). LG length varied from 17.4 to 117.7 cM and the highest number of loci per LG was 26. Most of the genetic intervals (79%) were <10 cM, with 11 gaps of >15 cM, the largest of which was 20.1 cM (on LG Alt_11, see Electronic

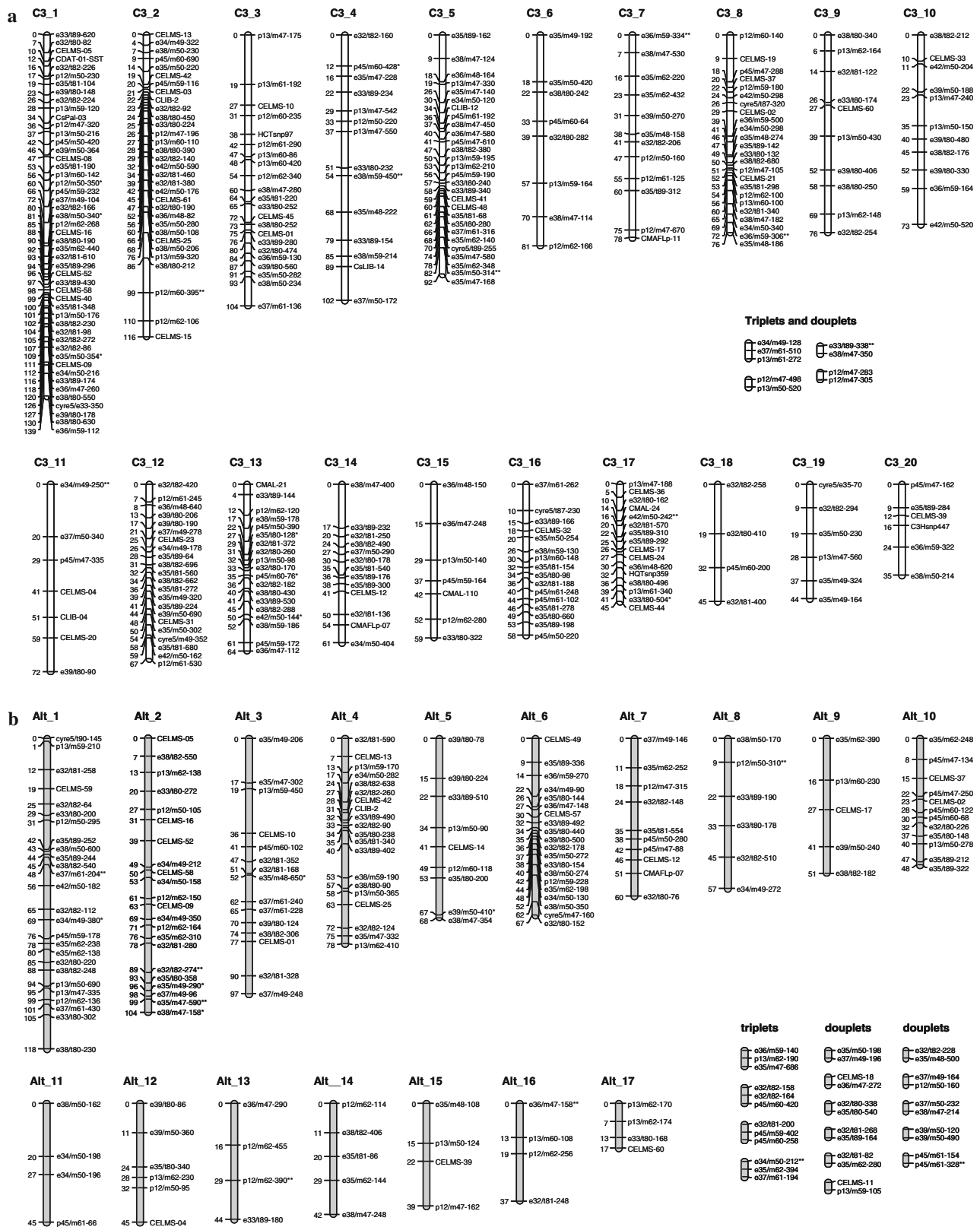


Fig. 1 Genetic maps of globe artichoke ‘Romanesco C3’ (a) and cultivated cardoon ‘Altalis 41’ (b). Marker names are shown to the right of each LG, with map distances (in cM) to the left. LGs with

fewer than four markers are shown as ‘minor groups’. Markers showing significant levels of segregation distortion are indicated by asterisks ($0.1 > P \geq 0.05$; $0.05 > P \geq 0.01$)

Table 2 Characteristics of the globe artichoke ('Romanesco C3') and cultivated cardoon ('Altilis 41') linkage maps

Romanesco C3					Altilis 41				
Linkage group	Size (cM)	No. of markers	Marker density	Gaps (>15 cM)	Linkage group	Size (cM)	No. of markers	Marker density	Gaps (>15 cM)
C3_1	138.7	49	2.9	0	Alt_1	117.7	26	4.7	0
C3_2	116.3	32	3.8	0	Alt_2	103.9	22	4.7	0
C3_3	103.5	22	4.9	1	Alt_3	96.6	15	6.9	2
C3_4	102.0	14	7.8	0	Alt_4	78.2	20	4.1	0
C3_5	92.0	28	3.4	0	Alt_5	67.7	9	8.5	1
C3_6	80.6	8	11.5	2	Alt_6	66.7	20	3.5	0
C3_7	78.4	12	7.1	1	Alt_7	60.4	10	6.7	0
C3_8	76.2	24	3.3	0	Alt_8	56.5	6	11.3	0
C3_9	75.9	10	8.4	0	Alt_9	50.9	5	12.7	1
C3_10	72.8	11	7.3	0	Alt_10	47.9	12	4.4	0
C3_11	71.9	7	12.0	1	Alt_11	45.3	4	15.1	2
C3_12	67.3	22	3.2	0	Alt_12	44.6	6	8.9	0
C3_13	64.1	19	3.6	0	Alt_13	44.0	4	14.7	2
C3_14	61.0	13	5.1	1	Alt_14	41.5	5	10.4	0
C3_15	58.5	7	9.8	1	Alt_15	39.3	4	13.1	2
C3_16	58.0	16	3.9	0	Alt_16	36.9	4	12.3	1
C3_17	45.1	16	3.0	0	Alt_17	17.4	4	5.8	0
C3_18	45.4	4	15.1	1					
C3_19	44.1	6	8.8	0					
C3_20	35.0	6	7.0	0					
Average	74.3	16.3	4.6	0.4		59.7	10.4	5.7	0.6
Total	1,486.8	326		8		1,015.5	176		11.0

supplementary Table S1). The AFLP loci were evenly distributed over the 17 LGs, with no apparent clustering of markers generated by any one PC. Five LGs (Alt_8, _11, _13, _14 and _16) included only AFLP markers, and the 23 mapped microsatellite loci were distributed over 12 LGs. The ten loci showing minor segregation distortion mapped to six LGs (four at $\alpha = 0.05$ and six at $\alpha = 0.01$), and a group of linked distorted loci was present on LG_Alt_2.

Of 79 intercross markers, 25 (22 AFLP, one microsatellite, one S-SAP, *Sp/sp*; Table 3) were assigned to a specific LG, but were not ordered within their LG. These, together with 18 microsatellite loci, were used to identify 11 homologous LGs (Table 3). The number of common markers per homologous LG varied between one and nine. *Sp/sp* was assigned to LG C3_14 and to LG Alt_7 (Table 3).

Accessory markers not placed in LGs at the initial stringency ($\text{LOD} \geq 6$), together with the highly distorted loci, were subsequently added to the maps when showing linkage only to a single LG and with a LOD score ≥ 4 . Table 4 reports the most likely positions within LGs for the accessory testcross markers. As a result of this second stage of analysis, two doublets of the female map (Fig. 1a) were linked to LG C3_2 and C3_12 (Table 4). In the globe

artichoke map, the most likely positioning of the highly distorted markers did not evidence clustering of distorted loci, while in the cultivated cardoon map a highly distorted locus was added to the cluster present at the end of LG Alt_2 (Table 4). A complete list of the accessory markers is reported in the Electronic supplementary Table S2.

Map alignment

The female map was compared with the 'Romanesco C3' map produced by Lanteri et al. (2006) using 32 common microsatellite loci, three common SNP loci and 49 common AFLP/S-SAP loci. These bridge markers identified 17 LGs (I to XVII in Fig. 2). Between 1 and 12 loci were present on any one of these aligned LGs, with microsatellite loci on 11 LGs and at least one AFLP locus on all 17. Eleven of the LGs aligned readily. In LGs III, IV and VI, two groups (major or minor) of the previously developed map joined and aligned with one group of the present map. On the other hand, in LGs XI and XII, two groups of the present map joined and aligned with one group of the previous ones. LG XIII is formed by two LGs from the previous and the present maps, which might be merged on the basis of microsatellite CMAL-21 (Fig. 2). Although the mean genetic separation between

Table 3 Intercross and shared testcross markers assigned to LGs in both the globe artichoke ('Romanesco C3') and cultivated cardoon ('Altalis 41') genetic maps

Romanesco C3 linkage group	Altalis 41 linkage group	Shared testcross markers	Intercross markers
C3_1	Alt_2	CELMS-05	e35/t80-364
		CELMS-16	e36/m59-236
		CELMS-52	CELMS-26
		CELMS-58	
C3_2	Alt_4	CELMS-13	p12/m47-266
		CELMS-42	e33/t89-500
		CELMS-25	e36/m59-216
		CLIB-02	e33/t80-288 e32/t82-88
C3_3	Alt_3	CELMS-10 CELMS-01	e38/t80-154
C3_5	Alt_1	–	e32/t80-304 e35/t89-116
C3_7	Alt_5	–	e33/t80-206
C3_8	Alt_10	CELMS-37	p12/m62-180
		CELMS-02	e35/t80-180 cyre5/m47-250
C3_9	Alt_17	CELMS-60	e33/t80-172 e33/t80-168
C3_11	Alt_12	CELMS-04	p45/m50-318
C3_14	Alt_7	CELMS-12	p45/m47-142
		CMAFLP-07	e35/m62-102 Spines
C3_17	Alt_9	CELMS-17	
C3_20	Alt_15	CELMS-39	p12/m50-450 p13/m60-260 e35/t80-220 e33/t89-450

pairs of loci and their relative order were mostly conserved across the two maps, 12 variation in marker order, affecting ten LGs, were detected.

Discussion

Map construction

Here, we have applied the double pseudo-testcross mapping strategy to construct linkage maps of a globe artichoke and a cultivated cardoon genotypes. The efficiency of this strategy depends both on the level of heterozygosity present in the parents, and on the level of detectable polymorphism between the parents (Cervera et al. 2001; Kenis and Keulemans 2005). *C. cardunculus* is an out-crossing species, and thus is expected to be highly heterozygous. By intermating two different taxa, it has been possible to create a population segregating for a number of significant agronomic traits (such as the size, shape, weight and form

of the head, and biomass production), as well as for the content of a number of secondary metabolism products of nutraceutical and pharmaceutical interest (Comino et al. 2007; Lanteri and Portis 2008).

For the linkage analysis of populations in cross-pollinating species, dominant intercross markers can be highly non-informative in certain configurations, and very often generate zero estimates for recombination frequency (Maliepaard et al. 1997); a further problem concerning co-dominant markers is that 1:2:1 segregations do not allow for the deduction of the parental origin of the segregating alleles (Maliepaard et al. 1998). The inclusion of markers segregating within both parents produces an estimate for recombination frequency which is the average outcome of both male and female meiosis, so may differ from estimated testcross frequencies, which are based on either male or female meiosis. Conflicts can thus arise between marker orders. The usefulness of a map clearly depends on its faithful reflection of actual locus order, so a decision was made to build the framework maps by

Table 4 Accessory testcross markers and specific LOD values in support of the two-point analysis with their linked loci

LG	Accessory marker	Linked locus 1	LOD	Linked locus 2	LOD
C3_1	p45/m59-460	e33/t89-430	5.29	e35/t81-348	4.47
C3_2	e35/m47-340***	e38/t80-390	9.20	p13/m60-110	8.69
	p12/m60-116*	p12/m62-106	4.54	e38/t80-212	4.14
	e35/t89-496	e35/m50-220	4.87	CELMS-13	3.73
	e33/t89-338**	e38/m47-350	8.71	CELMS-15	4.12
	e38/m47-350	e33/t89-338**	8.71		
C3_3	e38/t82-214	p12/m62-340	5.11	e35/t80-252	3.71
C3_4	cyre5/m47-180***	P13/m47-542	5.44	p12/m50-220	4.97
	e39/t80-556***	e33/t80-232	4.13	e35/m48-222	4.03
C3_5	e38/t82-124	p13/m47-542	5.69	e35/m47-168	4.09
	e38/t80-504	e35/t89-162	5.21	p13/m47-330	4.14
	e35/m49-610	e36/m47-580	4.20	p13/m59-195	3.80
C3_6	p13/m60-162***	e35/m50-420	4.12	e38/t80-242	4.01
C3_8	e32/t80-230	p12/m60-140	4.02		
C3_10	e32/t82-118***	e39/t80-330	10.4	e38/t82-176	6.22
	e38/m47-144	e39/t80-330	4.99	e39/t80-480	3.47
C3_12	p12/m50-320***	e35/m50-302	4.65	e38/t82-696	4.15
	p12/m47-283	p12/m47-305	9.65	p12/m61-530	4.02
	p12/m47-305	p12/m47-283	9.65		
C3_13	e35/m47-682***	e32/t80-170	8.91	e32/t82-182	8.03
C3_15	e38/t82-600	p45/m59-164	5.03	e33/t80-322	4.25
Alt_1	e39/m50-186	p13/m47-335	4.35	p13/m50-690	3.88
Alt_2	e33/t80-340***	e32/t81-280	7.41	e35/t80-358	6.64
	e35/t89-144	e35/m47-590**	5.02	e38/m47-158*	4.66
Alt_5	cyre5/t89-110***	e33/t89-510	4.71	e39/t80-224	4.27
Alt_8	e33/t89-512***	e33/t89-190	7.39	e33/t80-178	4.74
Alt_9	cyre5/m49-170***	p13/m60-230	5.13	CELMS-17	4.06
	e38/m50-500	p13/m60-230	4.12	CELMS-17	3.34
Alt_13	e38/t82-78	e36/m47-290	5.19	p12/m62-455	4.66
	e32/t81-160	e33/t89-180	4.88	p12/m62-455	4.03
Alt_15	e37/m49-140	e35/m48-108	4.09		

Markers showing significant levels of segregation distortion are indicated by one ($\chi^2_{x=0.1} < \chi^2 \leq \chi^2_{x=0.05}$), two ($\chi^2_{x=0.05} < \chi^2 \leq \chi^2_{x=0.01}$), or three ($\chi^2 > \chi^2_{x=0.01}$) asterisks. Markers reported as minor groups in Fig. 1 are shown in bold

excluding intercross markers which segregated consistently with a 3:1 or a 1:2:1 ratio. The cross used to generate these maps was wider than the one used to develop the first *C. cardunculus* maps (Lanteri et al. 2006), so there was a decreased number of common alleles between the parents and fewer markers segregating either 3:1 or 1:2:1.

The number of informative loci was lower for the cultivated cardoon parent than for the globe artichoke one, suggesting that the latter parent was more heterozygous than the former. Of the 114 microsatellite loci, 26 were heterozygous in the male parent and 50 in the female parent, giving estimates for the respective levels of heterozygosity of ~23 and ~44%, which is in broad agreement with the assessment of genetic variation previously detected in globe artichoke (Portis et al. 2005). ‘Romanesco C3’

is vegetatively propagated and has maintained its level of heterozygosity over time, whereas cultivated cardoon is seed-propagated and has probably been subjected to a limited degree of purifying selection aimed at stabilizing its production.

Framework maps were constructed using a high grouping threshold (LOD 6) along with an interval support LOD threshold of 1. The male parent framework map length (1,015.5 cM, 17 major LGs) was rather smaller than the female one (1,486.8 cM, 20 major LGs), and a higher frequency of unlinked triplets and duplets was generated, which suggests that certain genomic regions remain under-represented. Indeed, an higher efficiency in map construction of the most heterozygous parent has been previously reported in rubber tree (Lespinasse et al. 2000) and apricot

(Hurtado et al. 2002). Despite the large number of markers used to construct the female map, the number of LGs was greater than the known haploid number of 17. This may be because the mapping parameters were overly stringent, but may also be due a paucity of polymorphism in certain chromosomal regions. Although 43 intercross markers were assignable to an LG (Table 3), this number was insufficient to generate a genome-wide map, as only 11 LGs could be identified in this way.

Marker distribution and segregation ratio distortion

AFLPs were generated from restriction fragments produced by digestion with *Pst*I and *Mse*I (P/M), or *Eco*RI and either *Mse*I or *Taq*I (E/M, E/T). Although fewer E/M markers were obtained than E/T markers, their genomic distribution was more uniform. A globe artichoke map based solely on E/M PCs would give an higher coverage (862 cM) than the one based on E/T PCs (757 cM), making the former more suitable for mapping saturation. However, their combination improves the density of the linkage map. Some gaps remain in the female map, such as the ~20 cM ones on LGs C3_3, C3_11 and C3_18, and some of the LGs have a high mean inter-marker distance (>10 cM; LGs C3_6, C3_11 and C3_18). These may reflect localized low recombination rates, such as characterize the centromeric chromosome regions (Tanksley et al. 1992) or regions containing an excess of repetitive DNA (Jeuken et al. 2001; Vuylsteke et al. 1999).

Microsatellite loci are represented on 17 of the 20 female LGs, and 12 of the 17 male LGs. Two of the mapped microsatellites lie within the known genes SST and PAL, and others within genes whose function has been identified in other species (Electronic supplementary Table S3). The present results have confirmed the assignment of the HCT, HQT and C3'H genes (Comino et al. 2009; Moglia et al. 2009), involved in caffeoyl quinic acids biosynthesis, to specific globe artichoke LGs. The placement on the globe artichoke map of functionally annotated gene-derived markers may help to clarify the role of the genes which influence traits of interest, and is an essential step in the process of identifying candidate genes underlying the mechanisms of important morphological and physiological traits (Marino et al. 2009; Sargent et al. 2006; Vezzulli et al. 2008).

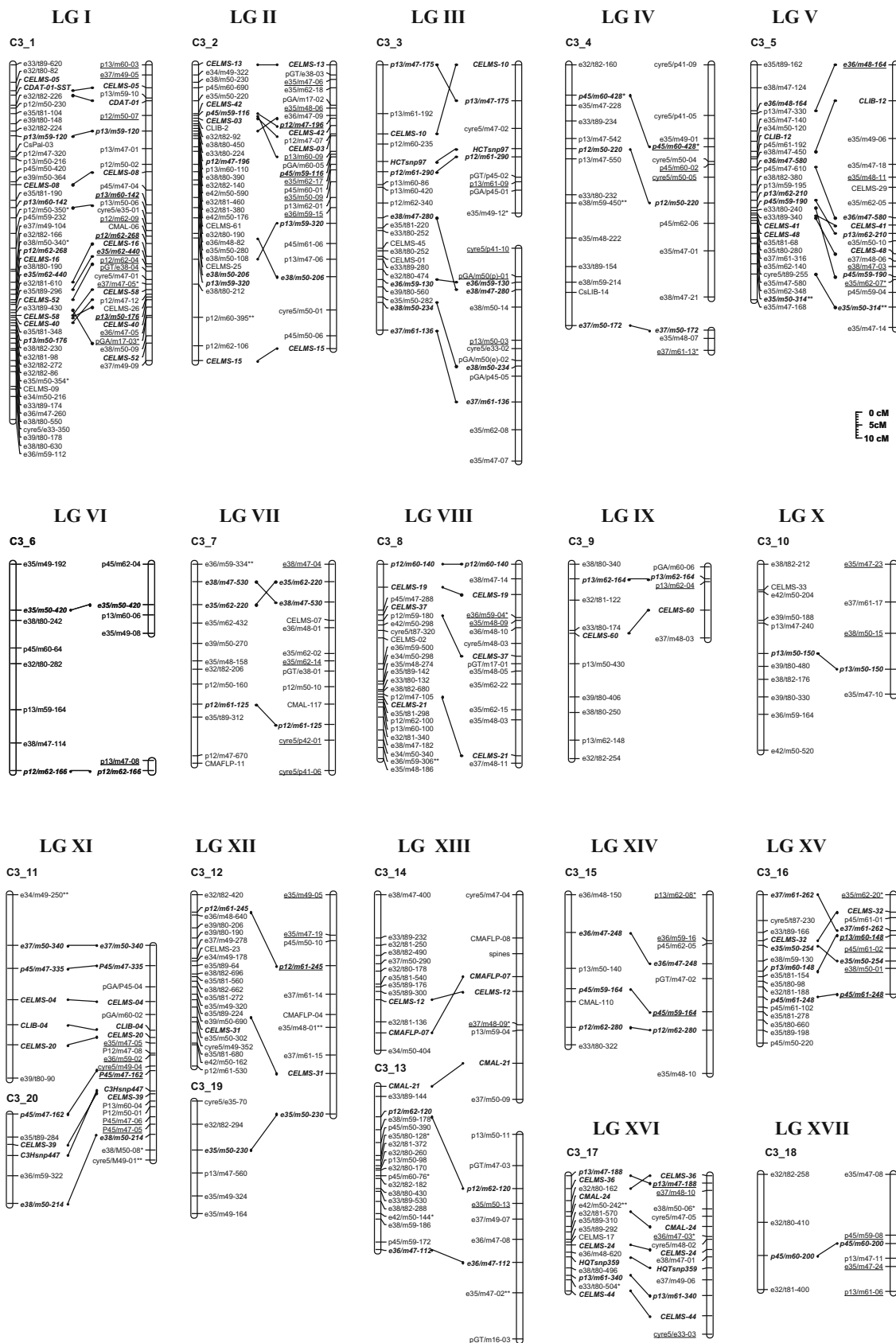
In the population derived from an intra-taxa (globe artichoke) cross (Lanteri et al. 2006), segregation distortion affected ~10% of loci, while in the present population, ~13% were affected. A greater degree of distorted segregation was not unexpected since it tends to increase with the genetic distance between the parents (Grandillo and Tanksley 1996; Verde et al. 2005). The basis of segregation distortion is not clearly understood and has been attributed

to a number of causes including aneuploidy, chromosomal translocations, competition among gametes, and the inheritance of alleles affecting the viability of the embryo (Gonzalo et al. 2005). Scoring and sampling errors can also influence the assessment of the level of distortion (Echt and Nelson 1997; Hackett and Broadfoot 2003; Nikaido et al. 1999). Some studies reported that where distorted markers are ignored, a significant part of a linkage group can be excluded (Cervera et al. 2001; Lorieux et al. 2000), we have only included those deviating up to 1%, to reduce the probability of false linkage. Notwithstanding, highly distorted markers have also been included in a second stage of analysis and their most likely positioning established. In the previous 'Romanesco C3' map, distorted markers clustered on LGs 1, 5 and 10 (Lanteri et al. 2006), but in the present map they were scattered across different LGs. This difference may be cross-specific (Igarashi et al. 2008) and/or may reflect the stringent LOD threshold applied.

Alignment and comparison of globe artichoke maps

Map alignment and merging of LGs are commonly achieved by analysing additional populations derived from the same parents or, as here, populations with one parent in common (Pelgas et al. 2005). We have established a skeletal globe artichoke map on the basis of 84 loci, defining 17 LGs (Fig. 2) corresponding to the haploid number of *C. cardunculus*. The three SNP loci and 32 of the microsatellite loci are present on the 'Romanesco C3' map (Lanteri et al. 2006), while a further 14 microsatellite loci have been mapped here for the first time. A set of 49 AFLP fragments common to both populations was also identified and, as previously reported (Costa et al. 2000; Krutovskii et al. 1998; Laucou et al. 1998; Zraidi et al. 2007) they have been successfully used for map alignment both here and in other studies. In general, marker order was conserved between the two maps, although some inconsistencies were noted (Fig. 2). Re-ordering of closely linked markers is relatively commonplace (Cervera et al. 2001; Jeuken et al. 2001; Lepinasse et al. 2000; Lombard and Delourme 2001; Sebastian et al. 2000) and reflects the statistical nature of the estimation of map order. Variation in stringency (LOD thresholds) is a major cause of mapping inconsistency, as is the use of many intercross dominant markers (Fig. 2). Other potential sources of variation are genotyping errors, an excess of missing values and the mapping of distorted markers (Hackett and Broadfoot 2003).

The extensive polymorphism between the mapping population parents reflects their taxonomical distance and helps to explain the reduced mean inter-marker map distances. The total length of the original 'Romanesco C3' map is about 93% of the present one, despite an ~33%



◀ **Fig. 2** A comparative genetic linkage map of globe artichoke ‘Romanesco C3’. LGs derived from the current population (on the left) aligned with the corresponding LGs defined by Lanteri et al. 2006 (on the right) on the basis of common microsatellite and AFLP loci. Common marker loci are shown in **bold**, and their positions are connected with a *line*. Intercross markers positioned in the LGs of the first map are *underlined*

increase in the number of loci mapped. The increased coverage of the globe artichoke genome obtained here will enable the exploration of distal portions of LGs VI, IX, X, XII and XV which were previously unmapped, as well as linking previously identified minor LGs on LGIV and VI (Fig. 2).

Conclusions

The 1C value of *C. cardunculus* has been estimated to be ~1,078 Mb (Marie and Brown 1993), thus the global relationship between physical and genetic distance is of the order of 1 cM = 725 kb. At present, none of the LGs can be assigned to a particular chromosome, a process which awaits the application of cytogenetic analysis. The current maps represent the most likely, but possibly not the actual arrangement of loci. An evenly spaced framework of markers based on the present set of LGs should facilitate genome-wide QTL scanning, and since 25 of the mapped markers (8%, Electronic supplementary Table S3) identify coding regions, the present map supplies the basis for candidate gene studies within the species.

Acknowledgments We thank Dr. Aniko Stigel, Dr. Lorenzo Barchi and Dr. Andrea Moglia for their technical assistance. This research was supported by MIPAAF (Ministero delle Politiche Agricole, Alimentari e Forestali - Italy) through the CarVarVi (“Valorizzazione di germoplasma di carciofo attraverso la costituzione varietale ed il risanamento da virus”) project.

References

- Acquadro A, Portis E, Lanteri S (2003) Isolation of microsatellite loci in artichoke (*Cynara cardunculus* L. var. *scolymus*). *Mol Ecol Notes* 3:37–39
- Acquadro A, Portis E, Lee D, Donini P, Lanteri S (2005a) Development and characterization of microsatellite markers in *Cynara cardunculus* L. *Genome* 48:217–225
- Acquadro A, Portis E, Albertini E, Lanteri S (2005b) M-AFLP-based protocol for microsatellite loci isolation in *Cynara cardunculus* L. (Asteraceae). *Mol Ecol Notes* 5:272–274
- Acquadro A, Portis E, Moglia A, Magurno F, Lanteri S (2006) Retrotransposon-based S-SAP as a platform for the analysis of genetic variation and linkage in globe artichoke. *Genome* 49:1149–1159
- Acquadro A, Lanteri S, Scaglione D, Arens P, Vosman B, Portis E (2009) Genetic mapping and annotation of genomic microsatellites isolated from globe artichoke. *Theor Appl Genet* 118:1573–1587
- Bassam BJ, Caetanoanollés G, Gresshoff PM (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. *Anal Biochem* 196:80–83
- Cervera MT, Storme V, Ivens B, Gusmao J, Liu BH, Hostyn V, Van Slycken J, Van Montagu M, Boerjan W (2001) Dense genetic linkage maps of three *Populus* species (*Populus deltoides*, *P. nigra* and *P. trichocarpa*) based on AFLP and microsatellite markers. *Genetics* 158:787–809
- Chatelet P, Stamigna C, Thomas G (2005) Early development from isolated microspores of *Cynara cardunculus* var *scolymus* (L.) Fiori. *Acta Hort* 681:375–380
- Chiapparino E, Lee D, Donini P (2004) Genotyping single nucleotide polymorphisms in barley by tetra-primer ARMS-PCR. *Genome* 47:414–420
- Comino C, Lanteri S, Portis E, Acquadro A, Romani A, Hehn A, Larbat R, Bourgaud F (2007) Isolation and functional characterization of a cDNA coding a hydroxycinnamoyltransferase involved in phenylpropanoid biosynthesis in *Cynara cardunculus* L. *BMC Plant Biol* 7:14
- Comino C, Hehn A, Moglia A, Menin B, Bourgaud F, Lanteri S, Portis E (2009) The isolation and mapping of a novel hydroxycinnamoyltransferase in the globe artichoke chlorogenic acid pathway. *BMC Plant Biol* 9:30
- Costa P, Pot D, Dubos C, Frigerio JM, Pionneau C, Bodenes C, Bertocchi E, Cervera MT, Remington DL, Plomion C (2000) A genetic map of Maritime pine based on AFLP, RAPD and protein markers. *Theor Appl Genet* 100:39–48
- Echt CS, Nelson CD (1997) Linkage mapping and genome length in eastern white pine (*Pinus strobus* L.). *Theor Appl Genet* 94:1031–1037
- Gonzalo MJ, Oliver M, Garcia-Mas J, Monfort A, Dolcet-Sanjuan R, Katzir N, Arus P, Monforte AJ (2005) Simple-sequence repeat markers used in merging linkage maps of melon (*Cucumis melo* L.). *Theor Appl Genet* 110:802–811
- Grandillo S, Tanksley SD (1996) Genetic analysis of RFLPs, GATA microsatellites and RAPDs in a cross between *L. esculentum* and *L. pimpinellifolium*. *Theor Appl Genet* 92:957–965
- Hackett CA, Broadfoot LB (2003) Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. *Heredity* 90:33–38
- Hurtado MA, Romero C, Vilanova S, Abbott AG, Llacer G, Badenes ML (2002) Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.), and mapping of PPV (sharka) resistance. *Theor Appl Genet* 105:182–191
- Igarashi M, Abe Y, Hatsuyama Y, Ueda T, Fukasawa-Akada T, Kon T, Kudo T, Sato T, Suzuki M (2008) Linkage maps of the apple (*Malus × domestica* Borkh.) cvs ‘Ralls Janet’ and ‘Delicious’ include newly developed EST markers. *Mol Breed* 22:95–118
- Jackson JA, Matthews D (2000) Modified inter-simple sequence repeat PCR protocol for use in conjunction with the LI-COR gene ImagIR(2) DNA analyzer. *Biotechniques* 28:914–917
- Jeuken M, van Wijk R, Peleman J, Lindhout P (2001) An integrated interspecific AFLP map of lettuce (*Lactuca*) based on two *L. sativa* × *L. saligna* F2 populations. *Theor Appl Genet* 103:638–647
- Kenis K, Keulemans J (2005) Genetic linkage maps of two apple cultivars (*Malus × domestica* Borkh.) based on AFLP and microsatellite markers. *Mol Breed* 15:205–219
- Krutovskii KV, Vollmer SS, Sorensen FC, Adams WT, Knapp SJ, Strauss SH (1998) RAPD genome maps of Douglas-fir. *J Hered* 89:197–205
- Lanteri S, Portis E (2008) Globe artichoke and cardoon. In: Prohens J, Nuez F (eds) *Vegetables I: Asteraceae, Brassicaceae, Chenopodiaceae, and Cucurbitaceae*. Springer, New York, pp 49–74
- Lanteri S, Di Leo I, Ledda L, Mameli M, Portis E (2001) RAPD variation within and among populations of globe artichoke cultivar ‘Spinoso sardo’. *Plant Breed* 120:243–246

- Lanteri S, Saba E, Cadinu M, Mallica G, Baghino L, Portis E (2004) Amplified fragment length polymorphism for genetic diversity assessment in globe artichoke. *Theor Appl Genet* 108:1534–1544
- Lanteri S, Acquadro A, Comino C, Mauro R, Mauromicale G, Portis E (2006) A first linkage map of globe artichoke (*Cynara cardunculus* var. *scolymus* L.) based on AFLP, S-SAP, M-AFLP and microsatellite markers. *Theor Appl Genet* 112:1532–1542
- Lauou V, Haurogne K, Ellis N, Rameau C (1998) Genetic mapping in pea. 1. RAPD-based genetic linkage map of *Pisum sativum*. *Theor Appl Genet* 97:905–915
- Lespinasse D, Rodier-Goud M, Grivet L, Leconte A, Legnate H, Seguin M (2000) A saturated genetic linkage map of rubber tree (*Hevea* spp.) based on RFLP, AFLP, microsatellite, and isozyme markers. *Theor Appl Genet* 100:127–138
- Lombard V, Delourme R (2001) A consensus linkage map for rapeseed (*Brassica napus* L.): construction and integration of three individual maps from DH populations. *Theor Appl Genet* 103:491–507
- Lorieux M, Ndjondjop MN, Ghesquiere A (2000) A first interspecific *Oryza sativa* × *Oryza glaberrima* microsatellite-based genetic linkage map. *Theor Appl Genet* 100:593–601
- Maliepaard C, Jansen J, Van Ooijen JW (1997) Linkage analysis in a full-sib family of an outbreeding plant species: overview and consequences for applications. *Genet Res* 70:237–250
- Maliepaard C, Alston FH, van Arkel G, Brown LM, Chevreau E, Dunemann F, Evans KM, Gardiner S, Guilford P, van Heusden AW, Janse J, Laurens F, Lynn JR, Manganaris AG, den Nijs APM, Periam N, Rikkerink E, Roche P, Ryder C, Sansavini S, Schmidt H, Tartarini S, Verhaegh JJ, Vrielink-van Ginkel M, King GJ (1998) Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. *Theor Appl Genet* 97:60–73
- Marie D, Brown S (1993) A cytometric exercise in plant DNA histograms, with 2C-values for 70 species. *Biol Cell* 78:41–51
- Marino R, Ponnaiah M, Krajewski P, Frova C, Gianfranceschi L, Pe ME, Sari-Gorla M (2009) Addressing drought tolerance in maize by transcriptional profiling and mapping. *Mol Gen Genomics* 281:163–179
- Moglia A, Comino C, Portis E, Acquadro A, De Vos R, Beekwilder J, Lanteri S (2009) Isolation and mapping of a C3'H gene (CYP98A49) from globe artichoke, and its expression upon UV-C stress. *Plant Cell Rep* 28:963–974
- Motzo R, Deidda M (1993) Anther and ovule culture in globe artichoke. *J Genet Breed* 47:263–266
- Nikaido A, Yoshimaru H, Tsumura Y, Suyama Y, Murai M, Nagasaka K (1999) Segregation distortion for AFLP markers in *Cryptomeria japonica*. *Genes Genet Syst* 74:55–59
- Pelgas B, Bousquet J, Beauseigle S, Isabel N (2005) A composite linkage map from two crosses for the species complex *Picea mariana* × *Picea rubens* and analysis of synteny with other Pinaceae. *Theor Appl Genet* 111:1466–1488
- Portis E, Barchi L, Acquadro A, Macua JJ, Lanteri S (2005) Genetic diversity assessment in cultivated cardoon by AFLP (amplified fragment length polymorphism) and microsatellite markers. *Plant Breed* 124:299–304
- Sargent DJ, Clarke J, Simpson DW, Tobutt KR, Arus P, Monfort A, Vilanova S, Denoyes-Rothan B, Rousseau M, Foltá KM, Bassil NV, Battey NH (2006) An enhanced microsatellite map of diploid *Fragaria*. *Theor Appl Genet* 112:1349–1359
- Sebastian RL, Howell EC, King GJ, Marshall DF, Kearsey MJ (2000) An integrated AFLP and RFLP *Brassica oleracea* linkage map from two morphologically distinct doubled-haploid mapping populations. *Theor Appl Genet* 100:75–81
- Sonnante G, Carluccio A, De Paolis A, Pignone D (2008) Identification of artichoke SSR markers: molecular variation and patterns of diversity in genetically cohesive taxa and wild allies. *Genet Resour Crop Evol* 55:1029–1046
- Stamigna C, Saccardo F, Pandozy G, Ancora G, Crinò P (2005) In vitro mutagenesis of globe artichoke (cv Romanesco). *Acta Hort* 681:403–410
- Tanksley S, Ganai M, Prince J, Devicente M, Bonierbale M, Broun P, Fulton T, Giovannoni J, Grandillo S, Martin G, Messeguer R, Miller J, Miller L, Paterson A, Pineda O, Roder M, Wing R, Wu W, Young N (1992) High-density molecular linkage maps of the tomato and potato genomes. *Genetics* 132:1141–1160
- Van Ooijen J (2006) JoinMap® 4: software for the calculation of genetic linkage maps in experimental populations. Kyazma B.V., Wageningen
- Verde I, Lauria M, Dettori M, Vendramin E, Balconi C, Micali S, Wang Y, Marrazzo MT, Cipriani G, Hartings H, Testolin R, Abbott AG, Motto M, Quarta R (2005) Microsatellite and AFLP markers in the *Prunus persica* [L. (Batsch)] × *P. ferganensis* BC(1) linkage map: saturation and coverage improvement. *Theor Appl Genet* 111:1013–1021
- Vezzulli S, Troggio M, Coppola G, Jermakow A, Cartwright D, Zharkikh A, Stefanini M, Grando MS, Viola R, Adam-Blondon AF, Thomas M, This P, Velasco R (2008) A reference integrated map for cultivated grapevine (*Vitis vinifera* L.) from three crosses, based on 283 SSR and 501 SNP-based markers. *Theor Appl Genet* 117:499–511
- Vos P, Hogers R, Bleeker M, Reijmans M, Vandelee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA-fingerprinting. *Nucleic Acids Res* 23:4407–4414
- Vuylsteke M, Mank R, Antonise R, Bastiaans E, Senior ML, Stuber CW, Melchinger AE, Lubberstedt T, Xia XC, Stam P, Zabeau M, Kuiper M (1999) Two high-density AFLP (R) linkage maps of *Zea mays* L.: analysis of distribution of AFLP markers. *Theor Appl Genet* 99:921–935
- Wagh R, McLean K, Flavell AJ, Pearce SR, Kumar A, Thomas BBT, Powell W (1997) Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). *Mol Gen Genet* 253:687–694
- Weeden N (1994) Approaches to mapping in horticultural crops. In: Gressho P (ed) *Plant genome analysis*. CRC Press, Boca Raton, pp 57–68
- Ye S, Dhillon S, Ke X, Collins A, Day I (2001) An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Res* 29:e88
- Zraidi A, Stift G, Pachner M, Shojaeiyan A, Gong L, Lelley T (2007) A consensus map for *Cucurbita pepo*. *Mol Breed* 20:375–388