

## Applicazione di tecniche molecolari per l'analisi del genoma di *Cynara cardunculus* L.

Ezio Portis, Alberto Acquadro e Sergio Lanteri\*

Dipartimento di Scienze agrarie, forestali ed alimentari, settore Genetica agraria, Università di Torino

### Application of molecular techniques for genomic analyses in *Cynara cardunculus* L.

**Abstract.** We report on research activities, carried out within the MIPAAF Research Project 'CAR-VARVI', aimed at providing additional tools for molecular analyses of the *Cynara cardunculus* L. genome. Starting from EST (expressed sequence tag) sequence data acquired from the Genbank (NCBI) of the *Compositae Genome Project* (CGP, USA), an extensive set of microsatellite markers (SSRs) was developed. Sufficient flanking sequence was available to enable the design of primers to amplify 2,311 microsatellites, and a set of 300 was tested against a DNA panel derived from 28 *C. cardunculus* genotypes, including the parents of three mapping populations. A set of 172 SSR loci were integrated into the reference *C. cardunculus* genetic maps previously developed, based on segregation among the F<sub>1</sub> progeny of a cross between a globe artichoke and a cultivated cardoon genotype. The resulting maps each detected 17 major linkage groups, corresponding to the species' haploid chromosome number. A consensus map based on 66 co-dominant shared loci, which includes 694 loci was assembled, with a mean inter-marker spacing of 2.5 cM. The maps was used to elucidate the pattern of inheritance of head production earliness. The F<sub>1</sub> mapping population together with a mapping population obtained by crossing the same globe artichoke genotype with one of wild cardoon featured a wide array of phenotypes in relation to several phenotypic traits, with some individuals, the result of specific events of chromosomal segregation and recombination, displaying aspects of morphology not previously observed in either cultivated or wild types. A set of nine SSR loci, evenly dispersed across the genome, was shown to be sufficient to unambiguously identify each segregant, some of which may have high potential in the context of developing ornamental varieties of the species. NGS (next generation sequencing) technologies were applied to produce a massive set of SNPs (single nucleotide polymorphism). Based on Illumina sequencing of gDNA RAD

(restriction associated DNA) tags of three mapping parents (e.g. non spiny globe artichoke, cultivated and wild cardoon) generated ~17000 SNPs (density 1/139 bp). Side by side, the transcriptome of the same mapping parents was sequenced by using a 454 platform, and raw data *de novo* assembled and annotated to generate the first reference transcriptome of the species (38,726 unigenes, 32.7 Mbp). The 454 reads, together with Illumina paired ends (PEs) from further eight *C. cardunculus* genotypes were aligned on the reference contig set, and ~195000 SNPs were called (density 1/169bp in coding regions).

**Key words:** globe artichoke, SSR, SNPs, genetic maps, ornamentals.

### Introduzione

La disponibilità di un ampio numero di tecniche di analisi del genoma della specie *Cynara cardunculus* L. ( $2n=2x=34$ ), ottimizzate per ognuno dei tre *taxa* botanici (carciofo: var. *scolymus* L., cardo coltivato: var. *altilis* DC, cardo selvatico: var. *sylvestris* Lam.), consente di fornire un importante contributo alla:

- caratterizzazione di collezioni *ex situ* di germoplasma, con l'obiettivo di identificare i criteri per la costituzione di *core collections*,
- analisi della variabilità genetica presente nel materiale attualmente in coltivazione, consentendo di definire strategie ottimali per applicare programmi di conservazione *in situ* di germoplasma;
- tipizzazione (*fingerprinting* molecolare) di genotipi provenienti da programmi di selezione clonale, contribuendo alla loro valorizzazione e fornendo informazioni per la loro iscrizione al Registro varietale;
- quantificazione della differenziazione genetica tra genotipi selezionati, allo scopo di identificare quelli più idonei per lo sviluppo di varietà propagabili per seme.

Inoltre, i marcatori molecolari rappresentano uno strumento indispensabile per sviluppare mappe genetico-molecolari e per l'identificazione delle basi gene-

\* sergio.lanteri@unito.it

tiche di caratteri semplici e complessi (QTLs, *quantitative trait loci*), presupposto indispensabile per l'applicazione di programmi di selezione assistita (MAS=*marker assisted selection*).

Presso il DISAFA, Dipartimento di Scienze Agrarie, Forestali e Alimentari, settore Genetica Agraria, dell'Università degli Studi di Torino (ex DIVAPRA), sono state condotte attività di ricerca finalizzate ad identificare le condizioni ottimali per l'applicazione delle tecniche AFLP (*Amplified Fragment Polymorphism*), RAPD (*Random Amplified Polymorphic DNA*), ISSR (*Inter-Simple Sequence Repeats*) ed è stata sviluppata la tecnica S-SAP (*Sequence-Specific Amplification Polymorphism*) basata sull'analisi del polimorfismo inserzionale di retrotrasposoni (Acquadro *et al.* 2006). Tali tecniche sono state utilizzate per:

- effettuare la caratterizzazione molecolare di collezioni di germoplasma di carciofo (Lanteri *et al.* 2004a) e cardo coltivato (Portis *et al.* 2005a), rappresentative del materiale attualmente in coltivazione;
- quantificare la variabilità genetica nell'ambito di popolazioni di carciofo delle tipologie 'Spinoso sardo' (Lanteri *et al.* 2001), 'Spinoso di Palermo' e 'Violetto di Sicilia' (Portis *et al.* 2005b) e di popolazioni spontanee insulari (siciliane e sarde) di cardo selvatico (Portis *et al.* 2005c);
- caratterizzare cloni selezionati nell'ambito della tipologia 'Spinoso sardo' (Lanteri *et al.* 2004b);
- caratterizzare ecotipi di carciofo coltivati in orti famigliari in Sicilia, avvalorando l'ipotesi che quest'ultima possa essere considerata un possibile centro di domesticazione della specie (Mauro *et al.* 2009).

Lo stesso gruppo di ricerca ha parallelamente sviluppato marcatori microsatelliti (SSR - *Simple Sequence Repeats*) mediante diversi approcci, quali l'analisi di sequenze depositate in GeneBank e lo sviluppo di librerie arricchite (Acquadro *et al.* 2003; 2009) e la tecnica, messa a punto presso i propri laboratori, definita MAL (*Microsatellite Amplified Library*) (Acquadro *et al.*, 2005).

Grazie allo sviluppo ed applicazione di marcatori molecolari è stato possibile sviluppare, in collaborazione con il Dipartimento di Scienze delle Produzioni Agrarie e Alimentari (DISPA) dell'Università di Catania, la prima mappa genetico-molecolare della specie (Lanteri *et al.*, 2006) a partire da una progenie segregante ottenuta dall'incrocio tra due genotipi di carciofo: "Romanesco C3" (tardivo, non spinoso utilizzato come parentale femminile) e "Spinoso di Palermo" (precoce, spinoso utilizzata come parentale

maschile) in base alla strategia del *two-way pseudo-testcross*. Tale mappa è stata successivamente integrata con l'inclusione di un ampio numero di marcatori SSR (Acquadro *et al.* 2009). In tempi più recenti è stata sviluppata una ulteriore mappa dello stesso parentale femminile, successivamente allineata alla precedente sulla base di 84 loci in comune, consentendo lo sviluppo di una *reference map*, caratterizzata da 17 gruppi di associazione omologhi, corrispondenti al numero cromosomico aploide della specie (Portis *et al.*, 2009).

Di seguito sono riportati i risultati di ricerche condotte nell'ambito del progetto CARVARVI (Valorizzazione di Germoplasma di Carciofo attraverso la Costituzione varietale ed il Risanamento da Virus), finalizzate a sviluppare ulteriori marcatori microsatellite ed identificare un ampio set di SNPs (*Single Nucleotide Polymorphisms*) mediante applicazione di tecniche NGS (*Next Generation Sequencing Technologies*). Inoltre, sono riportati i risultati di attività condotte in collaborazione con il DISPA dell'Università di Catania che hanno consentito di costruire nuove mappe genetiche della specie, una mappa consenso nonché di identificare le basi genetiche per il carattere precocità nella produzione di capolini e di genotipi di potenziale utilizzo a scopo ornamentale.

### Sviluppo di marcatori SSR a partire da EST

Marcatori microsatellite sono stati sviluppati a partire da sequenze EST depositate presso *Genbank* (NCBI) dal consorzio *Compositae Genome Project* (CGP, USA). Le 36.321 EST disponibili sono state utilizzate per il disegno di primer specifici fiancheggiati motivi ripetuti (SSR). In dettaglio, le sequenze EST sono state assemblate per mezzo dello script perl TGICL (<http://compbio.dfci.harvard.edu/tgi/software>) generando 19.055 Unigene. L'identificazione di sequenze ripetute è stata ottenuta grazie allo script SSRIT (<http://www.gramene.org/db/markers/ssrtool>), e su questi loci (denominati CyEMs, *Cynara Expressed Microsatellites*) sono state disegnate 2.311 coppie di primer specifici mediante l'ausilio del software BatchPrimer3 (<http://probes.pw.usda.gov/batchprimer3>). Il profilo elettroforetico di 300 marcatori CyEMs è stato analizzato su un panel di 24 genotipi, che includevano i parentali di progenie di mappa, determinandone sia il livello di polimorfismo che la loro applicabilità in studi di associazione. Rispettivamente 178 e 198 marcatori segreganti sono stati genotipizzati sulle progenie di mappa carciofo x cardo coltivato ("Romanesco C3" x "Altilis 41") e

carciofo x cardo selvatico (“Romanesco C3” x “Creta 4”) (fig. 1). L’analisi dei dati generati dal primo incrocio ha permesso di posizionare rispettivamente 137 ed 82 EST-SSR sulle mappe molecolari dei parentali “Romanesco C3” e “Altilis 41”. Grazie all’integrazione dei nuovi marcatori microsatellite, la mappa del parentale femminile “Romanesco C3”, costituita da 17 gruppi *linkage*, ha raggiunto una copertura di 1.505 cM ed una distanza media fra i marcatori di 3,3 cM. L’annotazione delle 36k EST è stata eseguita via BlastX utilizzando come database di riferimento le sequenze proteiche di *Arabidopsis thaliana* ( $e\text{-value} < e^{-29}$ ). È stata identificata la posizione relativa di ciascun motivo microsatellite all’interno dei trascritti (CDS, 5’-UTR, 3’-UTR) (fig. 1) e sono stati valutati gli arricchimenti funzionali per specifiche categorie GO (*Gene Ontology*) utilizzando il software GoStat (Beissbarth *et al.*, 2004). I dettagli delle metodologie adottate e risultati conseguiti sono riportati in Scaglione *et al.* (2009).

### Sviluppo di una mappa genetica consenso in *Cynara cardunculus*

La genotipizzazione della progenie ottenuta dall’incrocio carciofo x cardo coltivato, con marcatori AFLP e microsatellite, ha consentito di identificare circa un migliaio di marcatori sfruttabili per lo sviluppo di mappe di entrambi i parentali. In ciascuna di esse sono stati identificati 17 gruppi di associazione (*major linkage groups*). La mappa di cardo coltivato include 273 loci distribuiti su 1.486 cM, con una distanza media tra marcatori di 5,4 cM. La mappa di carciofo include 473 loci distribuiti su 1.544 cM, con una densità di un marcatore ogni 3,4 cM. Circa il 34% dei marcatori mappati sono stati attribuiti a sequenze espresse del genoma e pertanto possono rappresentare un valido strumento per futuri studi basati sull’approccio “candidate genes”. Poiché le due mappe condividono 66 loci codominanti, ciò ha reso possibile l’allineamento dei rispettivi gruppi di associazione e lo svi-

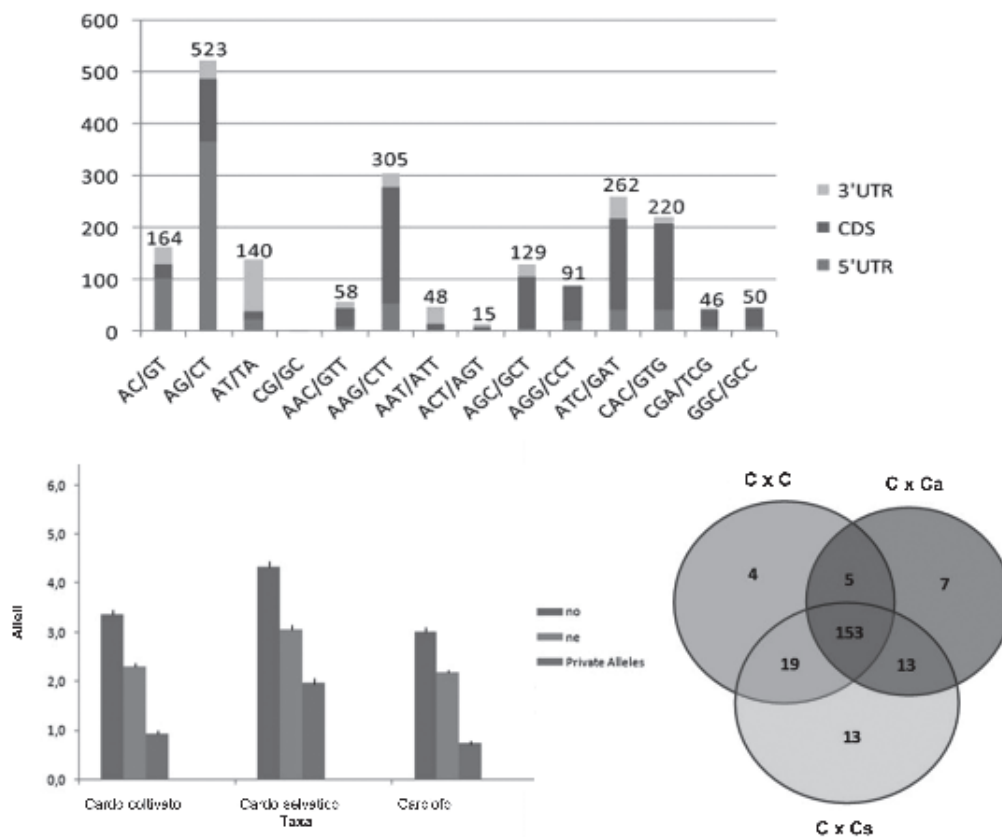


Fig. 1 - A) Distribuzione in classi dei motivi microsatellite identificati. Ripartizione dei motivi di- e tri-nucleotidici per regione genica (5’-UTR, sequenza codificante, 3-UTR). B) Statistiche relative agli alleli identificati nei tre *taxa* botanici; no = n° alleli osservati; ne = n° alleli attesi; “Private alleles” = n° alleli esclusivi. C) Marcatori che mostrano segregazione genetica all’interno delle tre popolazioni di mappaggio: CxC: entro carciofo, CxCa: carciofo x cardo coltivato, CxCs: carciofo x cardo selvatico.

Fig. 1 - A) Class distribution of the identified EST microsatellites in di- and tri-nucleotidic motifs for genic region (5’-UTR, coding sequence, 3-UTR). B) Statistics on the identified alleles in the three botanical taxa. Observed (no), effective (ne) and the number of taxon-specific alleles per marker. C) EST markers segregating within the three mapping progenies: CxC: within *scolymus*, CxCa: *scolymus* × *altilis*, CxCs: *scolymus* × *sylvestris*.

luppo di una mappa consenso esclusivamente basata su loci codominanti, che include 227 marcatori: 217 SSR e 10 SNPs presenti in geni coinvolti nella via biosintetica degli acidi caffeoilchinici (fig. 2).

La progenie di mappa ed i genotipi parentali sono stati coltivati per due anni successivi presso i campi sperimentali del DISPA - Università di Catania, e fenotipizzata per i seguenti 3 caratteri legati alla precocità: giorni intercorsi tra trapianto (o risveglio vegetativo al secondo anno) e produzione del capolino primario, produzione dei capolini di primo ordine e produzione dei capolini di secondo ordine. L'analisi della varianza per i tre caratteri ha evidenziato differenze significative tra i 2 genotipi parentali, in quanto il "Romanesco C3" ha prodotto capolini di tutti gli ordini con un anticipo di almeno tre settimane rispetto al cardo coltivato. Nell'ambito della progenie F<sub>1</sub> è stata rilevata una distribuzione continua per i tre caratteri ed un limitato numero di fenotipi trasgressivi, più tardivi rispetto al parentale cardo coltivato. I valore di ereditabilità è risultato più elevato per il carattere produzione del capolino principale ( $h^2= 0,76$ ), viceversa la produzione di capolini di primo e secondo ordine è risultata maggiormente influenzata da fattori ambientali.

Mediante analisi QTL, effettuata applicando il software *MapQTL 5.0*, sono state identificate 7 regioni genomiche in cui localizzano QTL che influiscono sulla precocità (fig. 3) e che, nel complesso, giustificano più del 74% della variabilità fenotipica osservata. In particolare un QTL, localizzato nella stessa regione omologa nelle mappe di entrambi i parentali, è risultato in grado di giustificare circa il 50% della variabilità osservata. Tale regione genomica è stretta-

mente associata a 2 marcatori microsatellite che possono rappresentare un valido strumento per una selezione indiretta del carattere precocità in futuri programmi di selezione assistita. I dettagli dell'attività sperimentale condotta sono riportati nella pubblicazione Portis *et al.* (2012).

**Sviluppo di nuovi genotipi di *Cynara cardunculus* sfruttabili a scopo ornamentale e loro *fingerprinting* molecolare**

La specie *C. cardunculus* è ampiamente sfruttata a scopo ornamentale. In essa è infatti presente una ampia variabilità per quanto riguarda il colore delle foglie (da verde scuro a grigio cenere), l'architettura e dimensione delle piante (che possono raggiungere i tre metri di altezza), la forma, dimensione e numero delle infiorescenze, il cui colore può variare dal verde al viola intenso. La specie è apprezzata come ornamentale sia nei giardini che come fiore reciso fresco e essiccato in composizioni floreali. Nelle progenie ottenute dagli incroci carciofo x cardo coltivato e carciofo x cardo selvatico sono stati identificati, a seguito di eventi di segregazione e ricombinazione cromosomica, nuovi genotipi/fenotipi di interesse ornamentale e sfruttabili a livello commerciale in un settore sempre alla ricerca di novità (fig. 4).

Nel complesso sono stati caratterizzati a livello fenotipico e molecolare 188 ibridi (94 per ciascuna delle due progenie). La progenie carciofo x cardo coltivato ha evidenziato la maggior variabilità per quanto riguarda la precocità di fioritura (da 203 a 242 giorni), l'altezza della pianta (da 56 a 127 cm) ed il numero di

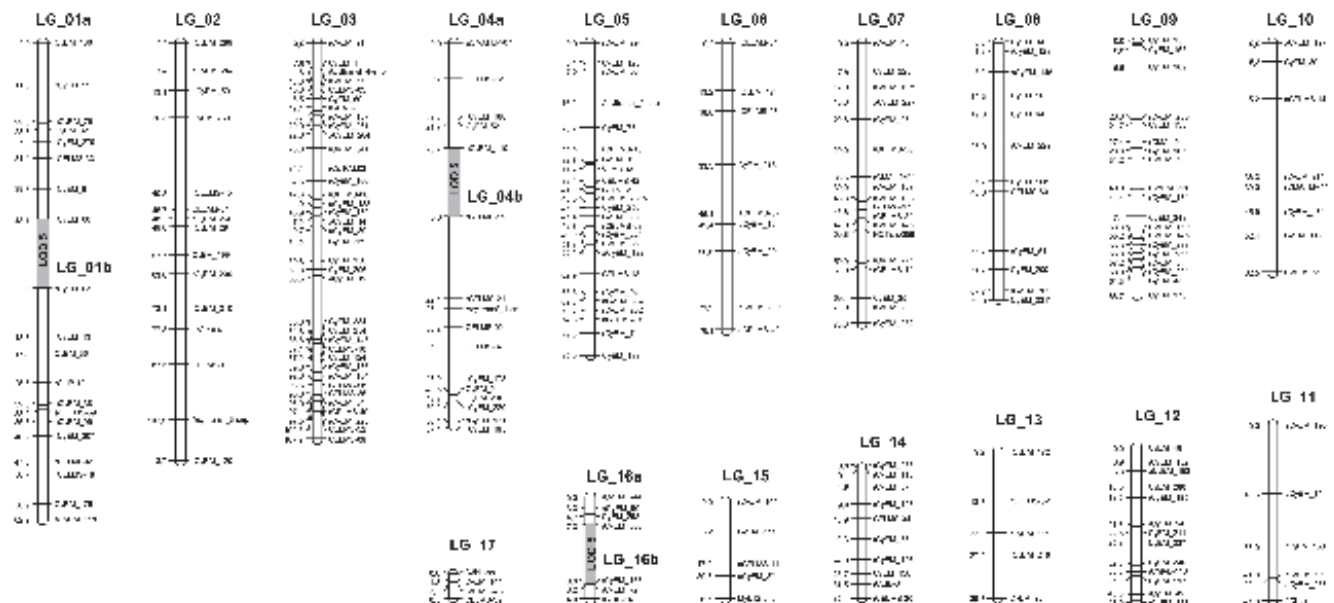


Fig. 2 - Mappa consenso di *C. cardunculus* basata su marcatori microsatellite ed SNP.  
 Fig. 2 - *C. cardunculus* consensus map based on microsatellite and SNP markers.

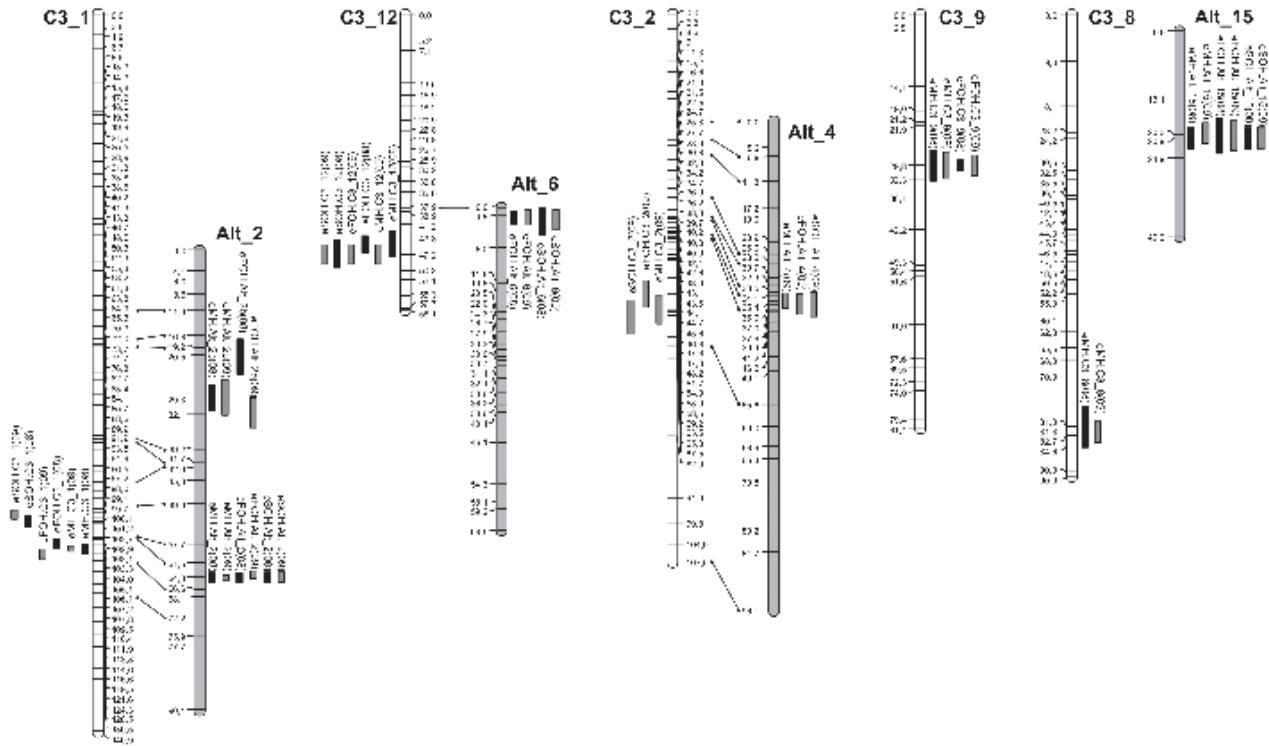


Fig. 3 - Localizzazione dei QTL identificati per la precocità nei gruppi Romanesco C3” (bianco) e “Atilis 41” (grigio). Le barre nere e grigie a lato dei gruppi di associazione rappresentano QTL evidenziati rispettivamente nel primo e nel secondo anno di caratterizzazione.  
 Fig. 3 - Location of earliness QTL in linkage groups of “Romanesco C3” (white) and “Atilis 41” (grey). Black and green bars represent QTL respectively detected the first and second year of characterization.

capolini per pianta (da 1 a 32), mentre in entrambe le progenie la durata della fioritura, carattere di interesse per lo sfruttamento come fiore reciso, è variata da 48 a 55 giorni. In entrambe le progenie è stata osservata un’ampia variabilità nella forma, da rotonda ad allungata a conica e nel colore del capolino che in alcuni casi ha assunto colorazioni screziate.

La caratterizzazione molecolare è stata effettuata mediante l’applicazione di 50 marcatori microsatellite che hanno originato 116 alleli (polimorfismi) nella progenie carciofo x cardo selvatico e 97 alleli nella progenie carciofo x cardo coltivato. Mediante l’applicazione di un set di soli 9 microsatelliti, scelti tra i più polimorfici ed uniformemente distribuiti nel genoma, è stato possibile identificare ciascun genotipo. Tale numero di marcatori può essere ulteriormente ridotto nel caso in cui si intenda effettuare la caratterizzazione molecolare di un limitato numero di genotipi per progenie a più spiccato valore ornamentale. Il *finger-printing* molecolare ha un interesse applicativo non solo per l’attribuzione del materiale propagato vegetativamente in vivaio ma anche per garantire i diritti del costituente di nuovi genotipi che potranno essere iscritti nel Registro Nazionale delle nuove costituzioni varietali. I risultati di questa linea di ricerca sono riportati in Lanteri *et al.* (2012).

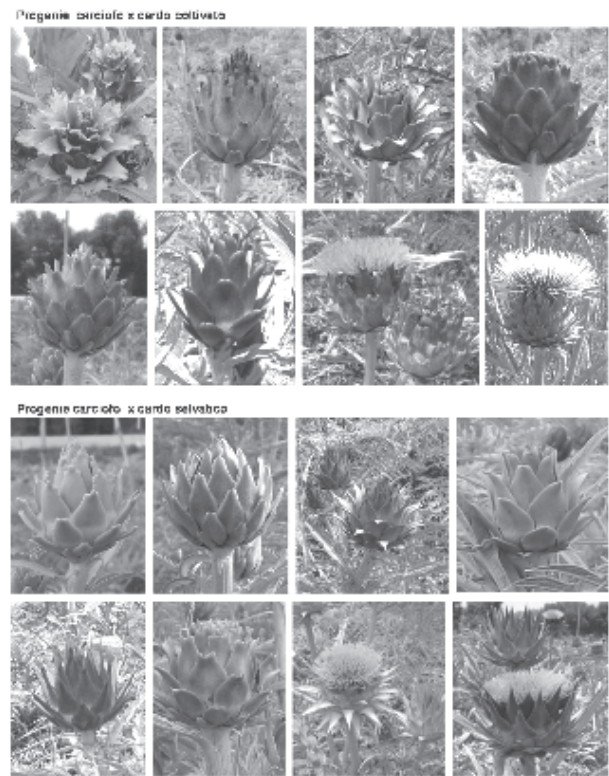


Fig. 4 - Esempi di variabilità fenotipica osservata nelle progenie da incrocio. Il genotipo a fiore bianco è caratteristico del parentale “Atilis 41” (cardo coltivato).  
 Fig. 4 - Examples of the phenotypic variation released in two inter-subspecies hybrid populations. White flower of the cultivated cardoon parent “Atilis 41” are also included.

## Sviluppo di marcatori SNP

I recenti progressi delle tecnologie di sequenziamento del DNA (*next generation sequencing* – NGS) rendono possibile identificare un elevatissimo numero di SNP (polimorfismi a carico di singoli nucleotidi) sfruttabili in piattaforme di *'high-throughput genotyping'*. In particolare sono state adottate due strategie (fig. 5):

- sequenziamento genomico mediante piattaforma Genome Analyzer II® (GAII, Illumina) di una rappresentazione genomica (*RAD - Restriction Associated DNA*) (Miller et al., 2007) dei 3 parentali utilizzati per lo sviluppo di progenie di mappa ("Romanesco C3" "Altilis 41" e "Creta 4");
- sequenziamento di librerie di cDNA a partire da RNA estratto da foglia e da radice, utilizzando la piattaforma 454® (Roche) per i tre parentali di mappa e la piattaforma GAII® (Illumina) per il sequenziamento di genotipi rappresentativi dei tre taxa botanici della specie di *C. cardunculus*: 5 varietà di carciofo, 2 di cardo coltivato ed una accessione di cardo selvatico.

Il *RAD-seq* ha prodotto 9,7 milioni di frammenti, equivalenti a circa ~1Gbp di sequenza. Il numero di frammenti è variato da 1,2 a 2,6 e 5,9 milioni rispettivamente in carciofo, cardo coltivato e cardo selvatico. L'assemblaggio *de novo* dei frammenti è stato effettuato utilizzando come campione di riferimento il cardo selvatico (l'insieme più numeroso) ed ha prodotto circa 19.000 contig di riferimento, con lunghezza media di 312 bp per un totale di 6,11 Mbp. Il contenuto in basi GC è risultato di circa 37,4%: tale stima è la prima disponibile per il genoma della specie ed è analoga a quella osservata in altre dicotiledoni. I dati

di sequenziamento hanno consentito, inoltre, di fare alcune inferenze sulla frequenza e composizione della frazione ripetuta del genoma, e di annotare, come geni putativi, il 21% circa delle sequenze ottenute. Nel complesso sono stati identificati circa 34.000 SNP e 800 *indel* (inserzioni-delezioni), con una frequenza stimata rispettivamente di 5,6 e 0,2 ogni 1.000 nucleotidi. Il numero di loci SNP eterozigoti, indice del livello di eterozigosi dei genotipi in analisi, è risultato di 1.235 in carciofo, 2.868 in cardo coltivato e di 5.069 in cardo selvatico. Gli SNP in eterozigosi, di cui un campione è stato convertito in marcatori CAPS (*cleaved amplified polymorphic sequence*) per una validazione, rappresentano marcatori ottimali per una genotipizzazione mediante piattaforma GoldenGate® (Illumina), allo scopo di ottenere mappe ad alta risoluzione della specie. I risultati di questa linea di ricerca sono dettagliatamente riportati in Scaglione et al. (2012a).

L'analisi del trascrittoma ha previsto la costruzione di undici librerie EST normalizzate, successivamente sequenziate con tecnologia 454 (per i tre parentali) o Illumina (per il germoplasma). Le sequenze 454 sono state utilizzate per produrre un trascrittoma di riferimento, mentre le sequenze Illumina sono state utilizzate per aumentare il numero totale di SNP identificati. Il sequenziamento dei tre parentali di mappa (454), ha prodotto 1,7 milioni di frammenti, equivalenti a circa 695 Mbp di sequenza, mentre il sequenziamento degli altri genotipi (Illumina) ha prodotto 46,4 milioni di frammenti (media 5,8), equivalenti a circa 6 Gbp di sequenza. L'assemblaggio delle reads 454, utilizzando l'assembler MIRA (Chevreux et al. 2004), ha generato un insieme di contig pari a 38.726 (lunghezza media: 844.3 bp; N50: 951 bp) per un

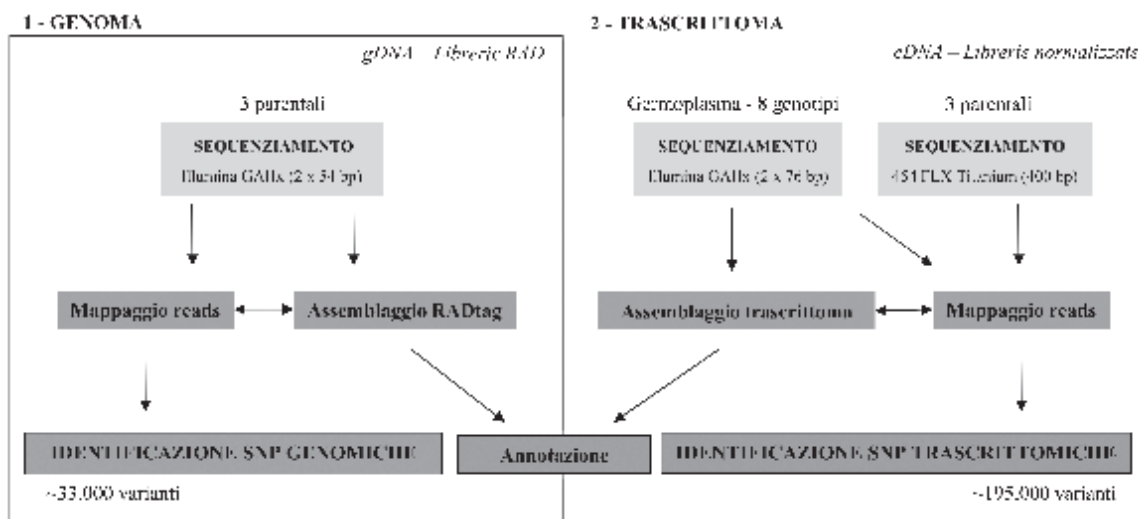


Fig. 5 - Schema relativo alla strategia adottata per l'identificazione di SNP genomiche e trascrittomiche  
Fig. 5 - Diagram of the steps followed for the identification of genomic and transcriptomic SNPs in di *C. cardunculus*.

totale di 32 Mbp (percentuale GC = 41%). Il numero di geni unici della specie è stato stimato essere di circa 30.000 (al netto delle varianti di *splicing*) e il trascrittoma di riferimento è stato annotato usando la piattaforma Blast2GO ([www.blast2go.org](http://www.blast2go.org)). In particolare, sono stati annotati con successo 32.400 geni unici. Di questi, 12.500 sono stati posizionati su pathway metabolici ([www.kegg.org](http://www.kegg.org)), 1.400 geni sono stati identificati come fattori trascrizionali (entro 67 famiglie differenti) e 316 sequenze sono state classificate entro geni di resistenza. In più di 1000 trascritti sono stati ritrovati siti di legame per 302 sequenze miRNA regolatorie, in geni appartenenti a categorie geniche quali: “risposta di difesa”, “apoptosi”, “riproduzione”, “sviluppo”, “fotosintesi”, “attività recettoriale” e “attività trascrizionale”. Nel complesso sono state identificate circa 195.400 varianti alleliche (su 11 genotipi), con una frequenza stimata di 1 SNP ogni 167 bp (in media cinque SNP per contig). La frequenza di SNP nel trascrittoma di *C. cardunculus* è apparsa comparabile con i dati relativi ad altre specie altamente eterozigoti come *Citrus* ssp. (EST; Jiang *et al.* 2010) e vite (genoma; Velasco *et al.* 2007). Gli SNP sono risultati più frequenti nella regione 3'-UTR (1 ogni 126 bp) e nelle CDS (1 ogni 169 bp) rispetto alle regioni 5'-UTR (1 ogni 265 bp). Sono state, inoltre, condotte analisi dettagliate su ogni genotipo al fine di identificare gli SNP intra-genotipiche e inter-genotipiche e sono state identificate varianti alleliche sotto pressione selettiva. I risultati di questa linea di ricerca sono dettagliatamente riportati in Scaglione *et al.* (2012b).

## Conclusioni

I risultati ottenuti a seguito delle attività di ricerca condotte nell'ambito del progetto CARVARVI hanno fornito importanti strumenti per l'analisi del genoma di *C. cardunculus*. L'identificazione di microsatelliti a partire da EST ha consentito di sviluppare un ampio set di nuovi marcatori genetici altamente polimorfici sia tra che entro i tre *taxa* botanici della specie. Tali marcatori hanno contribuito allo sviluppo di una mappa genetica ad alta risoluzione di carciofo e alla costruzione di mappa consenso di *C. cardunculus*, che costituiscono la base per futuri programmi di miglioramento genetico.

Nell'ambito delle progenie segreganti ottenute per lo sviluppo di mappe genetiche sono stati identificati genotipi segreganti/ricombinati di potenziale interesse ornamentale che, grazie alla propagazione vegetativa a cui è di norma soggetta la specie, possono essere moltiplicati e mantenuti nel tempo.

Per quanto riguarda l'identificazione di SNPs, la tecnica RAD ha dimostrato di rappresentare un approccio caratterizzato da alta efficienza anche in una specie altamente eterozigote, mentre la combinazione di due piattaforme di nuova generazione ha consentito una ampia caratterizzazione del suo trascrittoma ed un'ulteriore identificazione di un elevato numero di polimorfismi a carico di singoli nucleotidi.

Grazie alle nuove conoscenze relative all'organizzazione del genoma di *C. cardunculus* è stato possibile identificare regioni genomiche che influiscono sulla precocità nella produzione del capolino, e sono in corso attività di ricerca per identificare le basi genetiche di un ampio numero di caratteri di interesse agronomico e commerciale. Ciò consentirà l'applicazione di strategie di selezione assistita per lo sviluppo di nuovi genotipi che meglio rispondano alle richieste del mercato e nel contempo di valorizzare germoplasma autoctono italiano.

## Riassunto

Sono state condotte attività di ricerca con l'obiettivo di sviluppare in *Cynara cardunculus* L. marcatori microsatellite a partire da librerie EST (*expressed sequenced tag*) e, a seguito dell'applicazione di tecniche di sequenziamento di nuova generazione (NGS = *next generation sequencing*), marcatori SNPs (*single nucleotide polymorphisms*). In collaborazione con il DISPA dell'Università di Catania, è stata sviluppata una mappa consenso utilizzata per identificare QTLs che influiscono sulla precocità della produzione dei capolini. A seguito dell'ampia variabilità fenotipica osservata nell'ambito di due progenie di mappa sono stati identificati e caratterizzati a livello molecolare genotipi sfruttabili a scopo ornamentale.

**Parole chiave:** carciofo, marcatori microsatellite, marcatori SNP, mappe genetiche, ornamentali

## Bibliografia

- ACQUADRO A., PORTIS E., LANTERI S. 2003. *Isolation of microsatellite loci in artichoke (Cynara cardunculus L. var. scolymus L.)*. Molecular Ecology Notes 3: 37-39.
- ACQUADRO A., PORTIS E., LEE D., DONINI P., LANTERI S. 2005. *Development and characterisation of microsatellite markers in Cynara cardunculus L.* Genome 48: 217-225.
- 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: 1149-1156.
- ACQUADRO A., LANTERI S., SCAGLIONE D., ARENS P., VOSMAN B., PORTIS E. 2009. *Genetic mapping and annotation of genomic microsatellites isolated from globe artichoke.* Theoretical and Applied Genetics 118: 1573-1587.

- BEISSBARTH T, SPEED T. 2004. *GOstat: find statistically overrepresented Gene Ontologies within a group of genes*. *Bioinformatics* 20:1464-1465.
- CHEVREUX B., PFISTERER T., DRESCHER B., DRIESEL A., MULLER W., WETTER T., SUHAI, S. 2004. *Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs*. *Genome Research* 14: 1147-1159.
- JIANG D., YE Q., WANG F., CAO, L. 2010. *The mining of Citrus EST-SNP and its application in cultivar discrimination*. *Agric. Sci. China* 9: 179-190.
- LANTERI S., DI LEO I., LEDDA L., MAMELI M.G., PORTIS E. 2001. *RAPD, variation within and among populations of globe artichoke cultivar 'Spinoso sardo'*. *Plant Breeding* 120: 243-247.
- LANTERI S., SABA E., CADINU M., MALLICA G.M., BAGHINO L., PORTIS E. 2004a. *Amplified fragment length polymorphism for genetic diversity assessment in globe artichoke*. *Theoretical and Applied Genetics* 108: 1534-1544.
- LANTERI S., ACQUADRO A., SABA E., PORTIS E. 2004b. *Molecular fingerprinting and evaluation of genetic distances among selected clones of globe artichoke (Cynara cardunculus L. var. scolymus L.) 'Spinoso sardo'*. *Journal of Horticultural Science and Biotechnology* 79: 863-870.
- 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*. *Theoretical and Applied Genetics* 112: 1532-1542.
- LANTERI S., PORTIS E., ACQUADRO A., MAURO R.P., MAUROMICALE G. 2012. *Morphology and SSR fingerprinting of newly developed Cynara cardunculus genotypes exploitable as ornamentals*. *Euphytica* 184: 311-321.
- MAURO R., PORTIS E., ACQUADRO A., LOMBARDO S., MAUROMICALE G., LANTERI S. 2009. *Genetic diversity of globe artichoke landraces from Sicilian small-holdings: implications for evolution and domestication of the species*. *Conservation Genetics* 10: 431-440.
- MILLER M, DUNHAM J, AMORES A, CRESKO W, JOHNSON E. 2007. *Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers*. *Genome Res* 17: 240-248.
- PORTIS E., BARCHI L., ACQUADRO A., MACUA J.I., LANTERI S. 2005a. *Genetic diversity assessment in cultivated cardoon by AFLP (amplified fragment length polymorphism) and microsatellite markers*. *Plant Breeding* 124: 299-304.
- PORTIS E., MAUROMICALE G., BARCHI L., MAURO R., LANTERI S. 2005b. *Population structure and genetic variation in autochthonous globe artichoke germplasm from Sicily Island*. *Plant Science* 168: 1591-1598.
- PORTIS E., ACQUADRO A., COMINO C., MAUROMICALE G., SABA E., LANTERI S. 2005c. *Genetic structure of island populations of wild cardoon [Cynara cardunculus L. var. sylvestris (Lamk) Fiori] detected by AFLPs and SSRs*. *Plant Sci.* 169: 199-210.
- PORTIS E., MAUROMICALE G., MAURO R., ACQUADRO A., SCAGLIONE D., LANTERI S. 2009. *Construction of a reference molecular linkage map of globe artichoke (Cynara cardunculus var. scolymus)*. *Theoretical and Applied Genetics* 120:59-70.
- PORTIS E., SCAGLIONE D., ACQUADRO A., MAUROMICALE G., MAURO R., KNAPP S.J., LANTERI S. 2012. *Genetic mapping and identification of QTL for earliness in the globe artichoke / cultivated cardoon complex*. *BMC Research Notes* 5:252.
- SCAGLIONE D., ACQUADRO A., PORTIS E., TAYLOR C.A., LANTERI S., KNAPP S.J. 2009. *Ontology and diversity of transcript-associated microsatellites mined from a globe artichoke EST database*. *BMC Genomics* 10:454.
- SCAGLIONE D., ACQUADRO A., PORTIS E., TIRONE M., KNAPP S.J., LANTERI S. 2012a. *RAD tag sequencing as a source of SNP markers in Cynara cardunculus L.* *BMC Genomics* 13: 3.
- SCAGLIONE D., LANTERI S., ACQUADRO A., LAI Z., KNAPP S.J., RIESEBERG L., PORTIS E. 2012b. *Large-scale transcriptome characterization and mass discovery of SNPs in globe artichoke and its related taxa*. *Plant Biotechnology Journal* 10: 956-69.
- VELASCO, R., ZHARKIKH, A., TROGGIO, M., CARTWRIGHT, D., CESTARO, A. ET AL. 2007. *A high quality draft consensus sequence of the genome of a heterozygous grapevine variety*. *PLoS ONE*, 12: e1326.