7th International Safflower Conference
WAGGA WAGGA AUSTRALIA

Genetic linkage maps of *Carthamus* species based on SSR and RFLP markers.

Reinhold Mayerhofer, Victoria Bowles, Marion Mayerhofer and Allen G. Good
Department of Biological Sciences, University of Alberta, Edmonton, Alberta, T6G 2E9 Canada
rmayerho@ualberta.ca

Abstract
Genetic linkage maps for safflower (*Carthamus tinctorius*) and its wild relative *Carthamus oxyacanthus* were constructed using SSR and RFLP markers. An F2 population of 138 progeny derived from a cross of the cultivars Centennial x NP-12 and a BC1 population of 120 progeny derived from *C. oxyacanthus* x Centennial were used to generate the two maps, respectively. As a source of SSR markers we analyzed EST sequences obtained from the Compositae Genome Project and two industrial collaborators and about 1500 clones of an SSR-enriched genomic library of safflower. RFLP markers were derived from a safflower seedling cDNA library. To date we have mapped 190 markers on the two maps and identified the 12 linkage groups of *Carthamus.* We are continuing the sequencing of more genomic clones and saturating the linkage maps. In addition, the offspring of the two mapping crosses are being evaluated in the field for agronomic traits.

Key Words: Carthamus – safflower – microsatellite - molecular markers - genetic map

Introduction

The genus *Carthamus* is native to the Mediterranean coastlines, ranging from north Africa to Egypt, Spain to Greece and into the Middle East. It is part of the *Compositae (*Asteraceae*)* family, which includes other thistle plants and sunflowers. Members of the genus have been recently introduced into Australia and South America and are now cultivated in 60 countries. All wild species of safflower are spiny weeds, some of them very serious because they occupy fields sown to other crops. Others are more prevalent in roadsides and waste places. They are characterized by yellow flowers, although white flowers with yellow pollen grains are also known. At an early stage in its evolution, *C. tinctorius* spread to Egypt, Ethiopia, southern Europe, South Asia and the Far East where distinctive types have evolved.

The genus *Carthamus* encompasses 15 species, one of which is commercially grown. The classification for the genus has varied; Vilatersana *et al.* (2000, 2005) rearranged the species within 3 groups based on chromosomal number and phylogenetic relationships. The current status of the groups were summarized by McPherson *et al.* (2004). The three sections are *Carthamus* (n=12), *Odonthagnathis* (n=10,11) and *Atractylis* (n=22,32). There are still some species that have uncertain placement within the groups, such as *C. nitidus* (McPherson *et al*, 2004). The group is made up of a variety of species with varying chromosome numbers, n=10, 12, 22, and 32. It is generally considered that species with n=12 and n=10 are related and are the early ancestors, but it is not yet possible to tell which came first. Current thought is that the aneuploids are the result of crosses between members of separate species.

The commercially grown form is *Carthamus tinctorius,* otherwise known as safflower. In North America safflower is grown for birdseed or as an oilseed, the oil of which is usually sold in the health food market due to its qualities of high linoleic acid content and low saturate qualities. Other possible uses for safflower include industrial uses, dairy cattle feed, medical uses and cosmetics. Recently, safflower has been of interest for genetic engineering as a protein factory and is now being developed as a production platform for insulin (http://www.sembiosys.com/).

Safflower has not been the subject of many genetic studies and to date there are no reported genetic maps for this species. Genetic maps for species complexes of this nature can be developed using either RFLP or microsatellite/SSR (Simple Sequence Repeat) markers. SSR

markers have been found to show relatively large amounts of variability between individuals, and as such are good candidates for both mapping and for analyzing the genetic diversity within and between species (Hearne et al, 1992). The current model of is that variation within these repeats results from slippage of the polymerase during replication, resulting in the duplication or deletion of repeat units and thus altering the length of the overall microsatellite. This difference in length can be visualized by designing PCR primers on either side of the microsatellite and separating the PCR products on sequencing gels or capillaries.

The objects of this study were to develop a set of microsatellite markers for the genus *Carthamus* and to use these to develop molecular maps of *Carthamus tinctorius* and related species. These genetic maps can then be used to analyse the genetic basis of agronomic traits and traits that affect the suitability of this species as a protein production platform.

Materials and Methods

Plant material

Individual seeds were germinated from different accessions which were received from several international germplasm centers. For the development of genetically defined populations, where possible, single seed descent was performed for a minimum of two generations to reduce the level of genetic variability in the parental genotypes used in crossing. Plants were then emasculated and hand pollination was carried out. The flowers were then bagged, and the plants allowed to mature and fully senesce. Once dried, seeds were harvested and stored for 4 to 6 months to allow a break of dormancy. F1 seeds were then germinated in ddH_20 and sand and F1 plants were either selfed (Centennial x NP-12) or backcrossed with *C. tinctorius* as the recurrent parent (Centennial x *C. oxyacanthus*).

DNA resources and molecular markers

DNA markers were selected from several sources (Table 1). These included cDNA clones that were developed by the Compositae Genome Project (http://compgenomics.ucdavis.edu/), and a number of cDNA libraries that were developed by Syngenta Biotechnology or at the University of Alberta. Screening of the cDNAs for microsatellites and design of flanking SSR primers was carried out using the program PrimerPro (http://www.cs.ualberta.ca/~yifeng/primerpro/). A 20 bp M13(-21) tail was added to the 5' end of the forward primer for subsequent fluorescent labeling.

Table 1: Genomic resources evaluated as genetic markers in Carthamus. A total of 55,224 clones were analyzed of which 190 were mapped on to a *Carthamus tinctorius* map or a *C. tinctorius* x *C. oxycanthus* map.

Species	Genotype	Source	Tissue	Clones	Markers	Mapped
C. tinctorius	AC Sunset	Composite Genome Project	diverse tissues	40,000 cDNA	975 SSR	85
C. tinctorius	S-317	Syngenta	seed and seedling	7,500 cDNA	34 SSR	3
C. tinctorius	S-317	Syngenta	seed and seedling	1,224 genomic	62 SSR	n/a
C. tinctorius	S-317	U of Alberta	seedlings	5,000 cDNA	190 RFLP	90
C. tinctorius	NP-12	U of Alberta	seedlings	5,000 cDNA	n/a	n/a
C. tinctorius	Centennial	U of Alberta	seedlings	5,000 cDNA	n/a	n/a
C. tinctorius	S-317	U of Alberta	young leaves	1,500 genomic	576 SSR	12

A second set of microsatellite markers, developed from a small insert library of safflower as described by Bowles et al. (2008) were also used as a source of markers. All cDNAs used as

RFLP probes in Southern analyses were derived from a cDNA library developed at the University of Alberta and were analysed as described by Mayerhofer et al. (2005).

Microsatelite amplification and data analysis

The microsatellite loci were amplified and simultaneously labeled with fluorescent dyes using a protocol adapted from Schuelke (2000). PCR reactions were performed in a Gene Amp 9700 thermocycler (Perkin-Elmer,Norwalk, CT) in 15-ml reaction volumes containing 10 mm Tris-HCl (pH 8.0); 50 mm KCl; 0.01% gelatin; 2.5 mM MgCl; 0.2 mM each of dATP, dCTP, dGTP, and dTTP; 0.04 µM of the M13-tagged forward primer, 0,16 µM each of the reverse primer and the fluorescently labeled M13 primer; 1 unit/reaction Taq DNA polymerase; and 25 ng of template DNA. The cycle parameters were 94°C/5 min; 30 cycles of 94°C/30 sec, 56°C/45 sec, 72°C/45 sec; followed by 8 cycles of 94°C/30 sec, 53°C/45 sec, 72°C/45 sec; and a final extension of 72°C/10 min. The M13 primers were labeled with the fluorescent dyes FAM, VIC, NED or PET (Applied Biosystems, Foster City, CA), allowing for multiplexing of the PCR products before fragment analysis. The PCR products were diluted 10x in water and 2 µl were added to 8 µl formamide and 0.2 µl LIZ600 standard. The samples were subsequently run on a 3730 DNA Analyzer (Applied Biosystems). After scoring of the RFLP and microsatellite markers, MapDisto (http://mapdisto.free.fr/) was used to develop the genetic maps.

Results

Genomic resources

A variety of genomic resources were evaluated as potential genetic markers in Carthamus, including 5 cDNA libraries and 2 genomic libraries, one of which was a small insert library enriched for microsatellite markers. A total of 55,224 clones were analyzed of which 190 mapped on to a *Carthamus tinctorius* map or a *C. oxyacanthus* map. Of the cDNAs that were analysed in-silico, only a small percentage contained suitable microsatellite repeats, and of these only about 10% could be mapped. This was based in part on the failure to identify unique sequences from the 5' or 3' end of the SSR and from the inability to detect variation between the two parents used in the cross.

Molecular maps

Two genetic maps have been developed, the first being a intra-specific map which included all 12 of the Carthamus linkage groups (LGs) and a second, inter-specific map based on a *C. tinctorius* x *C. oxycanthus* cross (Table 2). For the intra-specific cross, a total of 12 LGs could be identified for a total of 896 cM or an average of 81.4 cM per LG. LG 12 was not included in this analysis. The LGs varied in size from 379 cM to 9 cM in length. The inter-specific cross has only 43 RFLP markers that have been mapped to date and each of the LGs was fairly small, which is not surprising, given the small number of markers that we were able to map.

Table 2: Genetic maps of *Carthamus tinctorius* (Centennial x NP-12) and *C. tinctorius* x *C. oxycanthus*. (Note: The Linkage Groups (LG1-12) cannot be assumed to be the same in each cross.

Centennial X NP12			C. oxyacanthus x Centennial		
Linkage Groups	**SSR Markers**	**Length in cM**	**Linkage Groups**	**RFLP Markers**	**Length in cM**
LG1	24	379	LG1	3	11
LG2	18	116	LG2	2	14
LG3	15	89	LG3	4	29
LG4	9	48	LG4	2	n/a
LG5	6	67	LG5	8	32
LG6	4	61	LG6	2	17
LG7	4	38	LG7	3	17
LG8	3	9	LG8	4	42
LG9	4	31	LG9	2	7
LG10	3	19	LG10	3	3
LG11	4	39	LG11	7	44
LG12	2	n/a	LG12	3	27

Discussion

The genomic resources analysed in this research were used to develop a number of chromosome specific markers. Several of these SSR based DNA markers also showed good amplification across species (Bowles et al. 2008) and the use of SSR and RFLP markers will allow us to identify specific LGs within both intra and inter-specific crosses.

While we have been able to identify all of the 12 LGs in *C. tinctorius*, further markers will be required to expand the linkage map to make it more reliable. Currently, some of the LGs have significant numbers of markers available (LG1) while others (LG7 to LG12) do not. Currently, additional markers from the genomic DNA libraries (both small insert SSR libraries and random genomic libraries), are being used to attempt to saturate the LGs. The resulting markers will be used to study the genetic basis of a number of key agronomic traits including flowering time, seed yield, and other traits of interest to safflower researchers. In addition, this data will allow us to continue our analysis of the relationships between species in the genus.

Acknowledgements

We would like to thank Corey Davis for help with the genomic library construction and microsatellite analysis and Laila Muzammal for technical assistance. The research was funded by SemBioSys Genetics and a NSERC CRD grant to AGG.

References

Bowles, V., Hall, J. and Good, A.G. 2008 Creation of Microsatellite Markers for investigation of relationships among closely related *Carthamus* species. 7th International Safflower Conference, Wagga Wagga, NSW, Australia, Nov 3-6.

Hearne, C.M., Ghosh, S. and Todd, J.A. 1992 Microsatellites for linkage analysis of genetic traits. Trends in Genetics **8**:288-294

Mayerhofer, R., Wilde, K., Mayerhofer, M., Lydiate, D., Bansal, V.K., Good, A.G. and Parkin, I.A.P. 2005 Complexities of chromosome landing in a highly duplicated genome: Toward map-based cloning of a gene controlling blackleg resistance in Brassica napus Genetics **171**:1977–1988

McPherson, M.A., Good, A.G., Topinka, A.K.C. and Hall, L.M. 2004 Theoretical hybridization potential of transgenic safflower (*Carthamus tinctorius* L.) with weedy relatives in the New World. Canadian Journal of Plant Science **84**:923-934

Scheulke, M. (2000) An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology **18**:233-234

Vilatersana, R., Garnatje, T., Susanna, A. and Garcia-Jacas, N. 2005 Taxonomic problems in *Carthamus* (Asteraceae): RAPD markers and sectional classification. Botanical Journal of the Linnean Society **147**:375-383

Vilatersana, R., Susanna, A., Garcia-Jacas, N. and Garnatje, T. 2000. Generic delimitation and phylogeny of the *Carduncellus-Carthamus* complex (Asteraceae) based on ITS sequences. *Plant Syst. Evol*. 221:89-105