



UNIVERSIDAD NACIONAL DEL SUR

TESIS DE DOCTORADO EN AGRONOMÍA

OBTENCIÓN DE VARIEDADES INVERNALES DE CÁRTAMO (*Carthamus tinctorius* L.) Y DESARROLLO DE HERRAMIENTAS MOLECULARES PARA ASISTIR LA SELECCIÓN

Ing. Agr. Alfonso Cerrotta

Directoras:

Dra. Lilia Ivone Lindström

Dra. Viviana Echenique

Bahía Blanca

Argentina

PREFACIO

Esta Tesis se presenta como parte de los requisitos para optar al grado Académico de Doctor en Agronomía de la Universidad Nacional del Sur y no ha sido presentada previamente para la obtención de otro título en esta Universidad u otra. La misma contiene los resultados obtenidos en investigaciones llevadas a cabo en el ámbito del Departamento de Agronomía y el Laboratorio de Biotecnología Vegetal del CERZOS (CCT- Bahía Blanca) durante el período comprendido entre el 10 de marzo de 2017 (Expte. 451/17) y la fecha actual, bajo la dirección de la Dra. Lilia Ivone Lindström y la Dra. Viviana Echenique.



UNIVERSIDAD NACIONAL DEL SUR

Secretaría General de Posgrado y Educación Continua

La presente tesis ha sido aprobada el / / , mereciendo la calificación de(.....)

AGRADECIMIENTOS

Muchas personas e instituciones han contribuido a mi crecimiento personal y profesional a través de estos años y han hecho posible la realización de la presente tesis. Quisiera ofrecerles mis sinceros agradecimientos:

A mis directoras, Dras. Viviana Echenique y Lilla Ivone Lindström por dirigirme y guiarme en las distintas etapas de este trabajo. Al Dr. Diego Zappacosta por guiarme y acompañarme durante la etapa de confección del mapa genético y el mapeo de QTLs.

A mis compañeras Clara Franchini y Andrea Flemmer del Laboratorio de morfología vegetal del departamento de Agronomía de la Universidad Nacional del Sur, y a los becarios e investigadores del laboratorio de Biotecnología vegetal del CERZOS-CONICET (CCT-Bahía Blanca) por acompañarme y ayudarme en todo el proceso de realización de esta tesis.

Al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) por haber contribuido en mi formación y en la realización de esta tesis a través del otorgamiento de la beca interna doctoral, y por poner a disposición todos sus recursos para realizar las diferentes actividades. Específicamente, al Centro de Recursos Naturales Renovables de la Zona Semiárida (CERZOS) por la contribución en cuanto a recursos humanos e instalaciones. Al departamento de Agronomía de la Universidad Nacional del sur por mi formación de grado y posgrado, y por brindarme sus instalaciones, facilidades y recursos humanos para el desarrollo de este trabajo.

A la Secretaria de Ciencia y Tecnología de la Universidad Nacional del Sur y la empresa OMHSA (GLENCORE) por el otorgamiento de recursos económicos para la financiación de las actividades de esta tesis.

A la dirección de criadero de trigo de la Asociación de Cooperativas Argentinas, por permitir la realización de ensayos en su campo experimental y especialmente a los Ingenieros Armando Junquera, Mariano Beker y a la memoria de Rubén Miranda por sus aportes técnicos y predisposición para ayudar a desarrollar el presente proyecto.

A la comisión de Establecimientos Rurales del departamento de Agronomía de la UNS por poner a disposición el lote experimental del campo Napostá, al igual que las maquinarias y el personal técnico para la realización de los laboreos.

Al Proyecto H2020-MSCA-RISE-2015 (ID: 691109): "*Exploring the molecular control of seed yield in crops (EXPOSEED)*" financiado por la Unión Europea por brindar los recursos para la realización de mi estadía de entrenamiento en selección genómica en el *Research Centre for Genomics & Bioinformatics, Council for Agricultural Research and Economics (CREA)* en Fiorenzuola d'Arda (Italia).

TABLA DE CONTENIDO

Prefacio	1
Agradecimientos.....	2
Tabla de contenido.....	4
Índice de figuras	8
Índice de tablas	12
Índice del anexo	15
Resumen.....	16
Abstract.....	18
Introducción general.....	20
Cártamo: Características generales de la especie.....	20
El cártamo en Argentina.....	20
Mejoramiento en cártamo	22
Caracteres cuantitativos.....	23
Caracteres cualitativos.....	25
Optimización del manejo del cultivo.....	26
Marcadores moleculares aplicados al mejoramiento	27
Estudio genético de las características de interés para el mejoramiento... 28	
Comparación de métodos de mapeo de QTLs	29
Mapeo de QTLs en poblaciones biparentales.....	32
Hipótesis.....	34
Objetivos	35
Capítulo 1. Selección en generaciones tempranas	37
1.1. Introducción	37
1.2. Materiales y métodos.....	42
1.2.1. Selección en F ₃	42
1.2.1.1. Material experimental y sitio de evaluación.....	42
1.2.1.2. Análisis preliminar de la anatomía del pericarpio y contenido de materia grasa de los granos	42
1.2.1.3. Selección.....	43
1.2.1.3.1. Fenotipo.....	43
1.2.1.3.2. Fenología.....	43
1.2.1.3.3. Espesor y lignificación del pericarpio.....	43

1.2.2. Selección en F ₄	44
1.2.2.1. Material vegetal y sitio de ensayo	44
1.2.2.1.1. Población de prueba (F ₅)	44
1.2.2.1.2. Población de mejora (F ₄).....	45
1.2.2.2. Determinaciones	45
1.2.2.2.1. Variables medidas en antesis	45
1.2.2.2.2. Variables medidas en cosecha	46
1.2.2.3. Análisis estadísticos.....	46
1.2.2.4. Índice de selección.....	47
1.2.2.5. Validación de un marcador molecular específico para el alelo OL.....	48
1.3. Resultados	50
1.3.1. Selección en F ₃	50
1.3.1.1. Análisis preliminar de la anatomía del pericarpio	50
1.3.1.2. Selección.....	53
1.3.2. Selección en F ₄	55
1.3.2.1. Caracterización climática	55
1.3.2.2. Análisis descriptivo y ANDEVA	56
1.3.2.3. Análisis multivariados.....	59
1.3.2.4. Selección indirecta a través de un índice de selección	64
1.3.2.5. Validación del marcador molecular del alelo OL	68
1.4. Discusión	70
Capítulo 2. Evaluaciones preliminares de rendimiento de MG	80
2.1. Introducción	80
2.2. Materiales y métodos	83
2.2.1. Material vegetal y ambientes evaluados	83
2.2.2. Variables medidas	84
2.2.3. Análisis estadísticos.....	85
2.2.4. Descripción de modelos.....	85
2.2.4.1. Rendimiento de MG	85
2.2.4.2. Contenido de MG	89
2.2.4.3. Caracteres morfológicos de la planta y componentes de rendimiento	90
2.3. Resultados	90

2.3.1. Descripción de los ambientes	90
2.3.2. Modelos de efectos fijos	94
2.3.3. Modelos AMMI	97
2.3.4. Índices de estabilidad basados en el modelo AMMI	99
2.3.5. Modelos mixtos	100
2.3.6. Índice de estabilidad basados en los BLUPs	104
2.3.7. Comparación de modelos AMMI vs. BLUPs (Modelos Mixtos).....	105
2.3.8. Caracteres morfológicos de la planta y componentes de rendimiento	106
2.4. Discusión	109
Capítulo 3. Evaluación de fechas de siembra en el sudoeste bonaerense	121
3.1. Introducción	121
3.2. Materiales y métodos	122
3.3. Resultados	124
3.3.1. Clima.....	124
3.3.2. Fenología	125
3.3.3. Variables morfológicas de la planta y componentes de rendimiento	127
3.3.3.1. Variables morfológicas de planta	128
3.3.3.2. Componentes de rendimiento	130
3.3.4. Contenido de MG.....	132
3.3.5. Rendimiento de grano.....	133
3.4. Discusión	136
Capítulo 4. Mapeo de QTLs	142
4.1. Introducción	142
4.2. Materiales y métodos	145
4.2.1. Material vegetal.....	145
4.2.2. Cruzamientos.....	145
4.2.3. Fenotipado	146
4.2.4. Extracción de ADN y genotipado	148
4.2.5. Mapa de ligamiento.....	149
4.2.6. Mapeo de QTLs	150
4.3. Resultados	151

4.3.1. Fenotipado de la población de mapeo	151
4.3.2. Mapa de ligamiento.....	155
4.3.3. Mapeo de QTLs	157
4.4. Discusión	163
Discusión general.....	169
Bibliografía	175
Anexo	203

ÍNDICE DE FIGURAS

Figura 1. Metodologías utilizadas para el análisis de genes/QTLs. Adaptado de Gupta <i>et al.</i> (2013).	29
Figura 2. Relación entre el grado de parentesco y la similitud fenotípica de una población de mapeo por asociación (izquierda), y el efecto de una cruce controlada (derecha). Extraído de Myles <i>et al.</i> (2009).	31
Figura 3. Relación entre la resolución del mapeo y el tiempo de desarrollo de las poblaciones utilizadas en el mapeo de QTLs. Extraído de Gupta <i>et al.</i> (2013). NAM: <i>nested association mapping</i> , MAGIC: <i>multiple advanced generation intercrosses</i> , RIAL: <i>recombinant inbred advanced intercross lines</i> , NIL: <i>near-isogenic lines</i> , RIL: <i>recombinant inbred lines</i> , DH: <i>double haploid</i> , BC: <i>backcross</i> , AM: <i>association mapping</i>	33
Figura 4. Esquema de un corte transversal del pericarpio de cártamo adherido a la cubierta seminal. Adaptado de Lockwood (1966).	39
Figura 5. Esquema del plan de selección y las evaluaciones preliminares de rendimiento de materia grasa (MG). ACA: Asociación de Cooperativas Argentinas, Cabildo; NAP: Campo Napostá-UNS, Torquinst; AgroUNS: Universidad Nacional del Sur, Bahía Blanca.	45
Figura 6 a-k. Anatomía del pericarpio de los parentales femeninos WSRC01 (a) y WSRC02(b), el parental masculino Ole (c) y las líneas F ₃ 103 (d), 11 (e), 121 (f), 129 (g), 173 (h) sembradas en el 2015 en el campo experimental de ACA (Cabildo). a-f: Pericarpio + cubierta seminal. g y h: pericarpio.	52
Figura 7 a y b. Modelos de regresión lineal entre el porcentaje de la materia grasa (MG) y el espesor total (ET, a) y espesor esclerificado (EE, b) del pericarpio de WSRC01, WSRC02 y 10 líneas F ₃ de cártamo cultivadas en el campo experimental de ACA (Cabildo) en el 2015. a: $MG=43.08 - 0,05ET$ ($R^2=0,39$). b: $MG=39,18 - 0,05EE$ ($R^2=0,77$).	53
Figura 8. Histograma de densidad de (a) las sumas térmicas entre emergencia y antesis (E-A) en grados día de las 206 líneas F ₃ y del (b) espesor de pericarpio de las 84 líneas F ₃ seleccionadas por precocidad y adaptación. Las cruces rojas indican la parte de la población descartada y la curva de trazo negro es la normal teórica para el conjunto de datos.	55
Figura 9. Precipitación media mensual y temperaturas medias, medias mínimas y medias máximas mensuales del campo experimental de ACA (Cabildo) para el año 2017 durante el ciclo del cultivo (mayo-diciembre). Comparación con las medias mensuales históricas de precipitación y temperatura del campo experimental de ACA (Cabildo).	55
Figura 10. Contenido de materia grasa del grano (MG, %) de nueve líneas experimentales (F ₅) y dos variedades testigo de cártamo. Barras coronadas con distinta letra indican diferencias significativas ($p<0,05$). Los brazos de las columnas indican el desvío estándar.	56
Figura 11. Proporción (%) de variabilidad explicada por cada uno de los componentes del análisis de componentes principales.	63
Figura 12. Biplot del análisis de componentes principales, los vectores de las variables incluidas, y los genotipos. MG: materia grasa, AP: altura de planta, GG: grosor de grano, GC: número de granos por capítulo, L:A: relación largo ancho del grano.	64
Figura 13. Reacción KASP de la población de prueba (F ₅ + Testigos WSRC03 y Montola 2000), la población de mejora y los híbridos (heterocigotas). Los ejes X e Y reproducen la magnitud las señales de los fluorescentes FAM y HET, representando los alelos OL y o/ respectivamente.	69

- Figura 14. Genotipo/fenotipo. Composición alélica en el locus *OL* obtenido con el marcador KASP y la composición de ácidos oleico y linoleico grasos promedio de los genotipos de la población de prueba (F_5 + Testigos WSRC03 y Montola 2000). Los brazos de las columnas indican el desvío estándar..... 70
- Figura 15. Precipitación total en el ciclo del cultivo (Ppt), en el período de septiembre a noviembre (PptSepNov) y en el período octubre-noviembre (PptSep-Oct) y rendimiento medio de MG de cada localidad. ACA: Asociación de Cooperativas Argentinas, Cabildo (2018); NAP: Campo Napostá-UNS, Torquinst (2018 y 2019); AgroUNS: Universidad Nacional del Sur, Bahía Blanca (2019). 91
- Figura 16. Ensayos comparativos de rendimiento de materia grasa. Líneas en estado de roseta (a y b), preantesis (c) y antesis (d) en ACA (Cabildo) en 2018. Líneas en antesis en NAP18 (e) y en madurez de cosecha en NAP19 (f). 92
- Figura 17. Unidades experimentales del ensayo comparativo de rendimiento de materia grasa realizado en el campo experimental de ACA (Cabildo, 2018). Imágenes tomadas en preantesis (26-10-2018, a-c), inicio de antesis (30-11-18, d-f) y fin de antesis (10-12, g-i). . 93
- Figura 18. Porcentaje de la variabilidad de rendimiento de materia grasa de 12 genotipos de cártamo evaluados en cuatro ambientes del sudoeste bonaerense. 95
- Figura 19. Rendimiento de materia grasa (MG) por hectárea de 12 genotipos de cártamo evaluados en cuatro ambientes del sudoeste bonaerense, y el promedio general de cada genotipo. ACA: Asociación de Cooperativas Argentinas, Cabildo (2018); NAP: Campo Napostá-UNS, Torquinst (2018 y 2019); AgroUNS: Universidad Nacional del Sur, Bahía Blanca (2019). Los brazos de las columnas indican el desvío estándar. 95
- Figura 20. Porcentaje de la variabilidad explicado por cada componente de varianza del modelo de ANDEVA para la variable contenido de materia grasa del grano (%) de 12 genotipos de cártamo evaluados en cuatro ambientes del sudoeste bonaerense. 96
- Figura 21. Porcentaje de materia grasa de los granos (MG, %) de 12 genotipos de cártamo evaluados en cuatro ambientes del sudoeste bonaerense, y promedio general de cada genotipo. ACA: Asociación de Cooperativas Argentinas, Cabildo (2018); NAP: Campo Napostá-UNS, Torquinst (2018 y 2019); AgroUNS: Universidad Nacional del Sur, Bahía Blanca (2019). Los brazos de las columnas indican el desvío estándar. 97
- Figura 22. Biplot de los dos componentes principales del análisis AMMI de los datos de rendimiento de materia grasa (MG) por hectárea de 12 genotipos de cártamo evaluados en cuatro ambientes del sudoeste bonaerense. ACA: Asociación de Cooperativas Argentinas, Cabildo (2018); NAP: Campo Napostá-UNS, Torquinst (2018 y 2019); AgroUNS: Universidad Nacional del Sur, Bahía Blanca (2019). 98
- Figura 23. Gráfico de dispersión del componente principal 1 frente al rendimiento medio de materia grasa (MG) por hectárea de 12 genotipos de cártamo evaluados en el sudoeste bonaerense. 99
- Figura 24. BLUPts de la variable rendimiento de materia grasa del modelo multiambiental (MA) y multiambiental con datos filtrados (MAf). Las líneas punteadas indican diferencias de rendimiento de materia grasa con respecto a la media obtenido por la prueba t ($p < 0,05$). 102
- Figura 25. Índice de estabilidad WAASBY con una relación equilibrada entre rendimiento y estabilidad de la materia grasa (50/50). Los círculos de color rojo y azul indican valores inferiores e superiores al promedio, respectivamente. 104
- Figura 26. Mapa de calor de los rankings de genotipos por índice WAASBY dependiendo en la relación WAASB/MG ha⁻¹. El color de las etiquetas de los genotipos indica pertenencia a un *cluster*, diferenciándose ($p > 0.05$) de los genotipos de otros *clusters*. 105

- Figura 27. Validación cruzada. Boxplot del error cuadrático medio de cada uno de los modelos aplicados a los datos de rendimiento de materia grasa de 12 genotipos de cártamo evaluados en cuatro ambientes del sudoeste bonaerense. En celeste, se identifica el de menor error de predicción. MM: modelo mixto. 106
- Figura 28. Análisis de correlación, histograma y gráficos de dispersión del rendimiento medio de materia grasa (MG ha^{-1}) ($n=48$), las características morfológicas de planta y los componentes de rendimiento. Los coeficientes de correlación y su significancia se encuentran en el triángulo superior de la matriz, los gráficos de dispersión en el triángulo inferior de la matriz, y los histogramas en la diagonal. MG.ha: rendimiento de MG en kg.ha^{-1} , AP: altura de planta, R1: ramificaciones, C: capítulos, C5RS: capítulos en las cinco ramificaciones superiores, DC: diámetro de capítulo, GC: granos por capítulo, PGC: peso de granos por capítulo, P1000: peso de mil granos. *: $p<0,05$, **: $p<0,01$, ***: $p<0,001$ 108
- Figura 29. Temperatura (T°) media mensual, precipitaciones (Ppt.) de AgroUNS, Napostá y el promedio de los últimos 30 años para la localidad de Bahía Blanca (Estación CONICET). Las columnas verdes representan los mm de riego complementario aplicado en AgroUNS 125
- Figura 30. Días transcurridos (barras) y grados día acumulados (GDD) (líneas) entre siembra y antesis (SA) en promedio para cada fecha de siembra y genotipo (CW99 OL, Montola 2000 y L69) en las localidades de AgroUNS (izquierda) y Napostá (derecha). Para cada variable y localidad, letras diferentes indican diferencias significativas ($p<0,05$). 126
- Figura 31. Altura de planta (AP) (cm) por fecha de siembra y por localidad. Letras diferentes indican diferencias significativas ($p<0,05$). Los brazos de las columnas indican el desvío estándar. 129
- Figura 32. Diámetro de capítulo (DC) por fecha de siembra y por localidad. Letras diferentes dentro de cada gráfico indican diferencias significativas ($p<0,05$). Los brazos de las columnas indican el desvío estándar. 130
- Figura 33. Número de capítulos en las cinco ramificaciones superiores (C5RS) por fecha de siembra y por localidad. Letras diferentes indican diferencias significativas ($p<0,05$). Los brazos de las columnas indican el desvío estándar. 130
- Figura 34. Número de granos por capítulo por localidad y por genotipo. Letras diferentes indican diferencias significativas ($p<0,05$). Los brazos de las columnas indican el desvío estándar. 131
- Figura 35. Peso de 1000 granos (P1000) (g) por fecha de siembra, genotipo y localidad. Para cada Localidad, letras diferentes indican diferencias significativas ($p<0,05$). Los brazos de las columnas indican el desvío estándar. 132
- Figura 36. Materia grasa promedio (MG, %) por fecha de siembra y por genotipo dentro de cada localidad. Para cada localidad, letras diferentes indican diferencias significativas ($p<0,05$). Los brazos de las columnas indican el desvío estándar. 133
- Figura 37. Rendimiento de grano por fecha de siembra y por genotipo (CW99 OL, Montola 2000 y L69) para AgroUNS (izquierda) y Napostá (derecha). Para cada localidad, letras diferentes indican diferencias significativas ($p<0,05$). Los brazos de las columnas indican el desvío estándar. 134
- Figura 38. Análisis de regresión lineal entre el rendimiento de grano y los días transcurridos a partir de la primera fecha de siembra (14 de mayo de 2019) en tres genotipos (CW99 OL, Montola 2000 y L69) sembrados en dos localidades (Napostá y AgroUNS). Cada punto corresponde a la media general de cada fecha de siembra (FS1 a FS6). 134
- Figura 39. Histogramas de densidad junto a la curva normal de referencia de cada característica fenotípica en una población F_2 ($n=91$, WSRC03 x Montola 2000) junto a la media de los

parentales cultivados en 2018 en ACA (Cabildo). La flecha azul representa el valor fenotípico medio del parental femenino (WSRC03) y la verde del valor fenotípico medio del parental masculino (Montola 2000). 154

Figura 40. Matriz de correlaciones de Pearson de las variables fenotipadas en la población de mapeo F₂ de cártamo (n=91, WSRC03 x Montola 2000) en 2018 en ACA (Cabildo). Corr: coeficiente de correlación, AP: altura de planta, R1: número de ramificaciones, C5RS: capítulos en las 5 ramificaciones superiores, DC: Diámetro de capítulos, RENDP: rendimiento por planta, P1000: peso de mil granos, GC: granos por capítulo, PGC: peso de granos por capítulo, CC: contenido de cáscara, CLG: contenido de lignina del grano, LG: largo del grano, AG: ancho del grano, GG: grosor del grano, L:A: Relación largo-ancho del grano, L:G: relación largo grosor del grano, DIAMF: diámetro de grano, VOLG: volumen de grano, DENSG: densidad de grano, GD_S_A: grados día siembra-antesis, GDD_S_BBCH79: grados día siembra-BBCH79, GD_A_BBCH79: grados día antesis BBCH79. 155

Figura 41. Distribución de la distorsión de la segregación en el mapa de ligamiento (798 cM) realizado en la población de mapeo F₂ de cártamo (n=91, WSRC03 x Montola 2000) representado con el estadístico X² (eje y) sobre la posición en el mapa (eje x). Para los marcadores dominantes y codominantes se aplicó un umbral establecido de 1 y 2 grados de libertad, respectivamente. 157

Figura 42. Mapa de ligamiento de los 18 GL agrupados en 12 SCs según la localización de los marcadores en el genoma de referencia de Bowers *et al.* (2016). Se posicionaron los QTLs hallados en una población de mapeo F₂ de cártamo (n=91, WSRC03 x Montola 2000) fenotipada en 2018 en ACA (Cabildo). GLs del mismo color pertenecen al mismo subgrupo. A la izquierda de cada GL se encuentra la posición dentro del mismo, y a la derecha el nombre del marcador. QTLs mapeados en las regiones involucradas. R1: número de ramificaciones, C5RS: capítulos en las 5 ramificaciones superiores, DC: diámetro de capítulos, RENDP: rendimiento por planta, P1000: peso de mil granos, PGC: peso de grano por capítulo, CC: contenido de cáscara, CLG: contenido de lignina del grano, LG: largo del grano, AG: ancho del grano, GG: grosor del grano, L:A: Relación largo-ancho del grano, L:G: relación largo grosor del grano, DIAMG: diámetro de grano, VOLG: volumen de grano. Los QTLs están coloreados según el tipo de variable (amarillo: tamaño de grano, caracteres de la cáscara: azul, verde: caracteres de la planta, rojo: dimensiones del futo y el capítulo). 163

Figura 43. Perfil del LOD (línea roja) y de explicación fenotípica (EF, línea azul) del GL 7_1 para el contenido de lignina del grano de una población de mapeo F₂ de cártamo (n=91, WSRC03 x Montola 2000). 163

ÍNDICE DE TABLAS

Tabla 1. Genotipos de cártamo incluidos en la población de prueba.	44
Tabla 2. Cebadores de los marcadores asociados al carácter alto oleico. F/R: <i>forward/reverse</i>	49
Tabla 3. Media, desvío estándar (DE), coeficiente de variación (CV), y coeficiente de correlación de Pearson (r MG) del contenido de materia grasa (MG), el espesor total, el espesor esclerificado, los estratos del parénquima interno y los estratos del parénquima externo del pericarpio de WSRC01, WSRC02 y 10 líneas F ₃ de cártamo cultivadas en el 2015 en el campo experimental de ACA (Cabildo).....	53
Tabla 4. Medidas resumen de los grados día entre emergencia y antesis (E-A), espesor esclerificado (EE) y total (ET) de pericarpio. n: número de líneas, DE: desvío estándar, CV: coeficiente de variación.....	54
Tabla 5. Media, desvío estándar, coeficiente de variación, varianza, valor mínimo y máximo de las variables medidas en el ensayo de la población de prueba de cártamo (F ₅ +Testigos WSRC03 y Montola 2000) (n=33). Grados día acumulados desde la emergencia hasta la antesis (GDD E-A) y la elongación de entrenudos (GDD E-E). G: granos. R1°: ramificaciones de primer orden.....	58
Tabla 6. Estadístico de Fisher (F) de cada fuente de varianza, grados de libertad y significancia (p _{error}) para cada una de las variables analizadas en la población de prueba (F ₅ + Testigos WSRC03 y Montola 2000) (n=33). Grados día desde emergencia hasta antesis (GDD E-A) y hasta elongación de entrenudos (GDD E-E). G: grano/s. R1°: ramificaciones de primer orden.	59
Tabla 7. Matriz de correlaciones de Pearson utilizando los valores medios por parcela (n=33) de la población de prueba (F ₅ + Testigos WSRC03 y Montola 2000). En la matriz se encuentran los valores de r y la significancia del error. MG: materia grasa, AP: altura de planta, R1°: número de ramificaciones de primer orden, DC: diámetro del capítulo, CPR: capítulos por ramificación de primer orden, PGC: peso de grano por capítulo, GC: granos por capítulo, P1000: peso de 1000 granos, CC: contenido de cáscara del grano, LG: largo del grano, GG: grosor del grano, L:G: relación entre largo y grosor del grano, L:A: relación entre largo y ancho del grano, DGM: diámetro geométrico medio del grano, ESF: esfericidad del grano, GD E-A: grados día acumulados entre emergencia y antesis, GD E-E: grados día acumulados entre emergencia y elongación de entrenudos.	61
Tabla 8. Análisis de sendero de la materia grasa (MG). Efectos directos (en la diagonal) e indirectos (fuera de la diagonal). En el pie de la tabla, el coeficiente de correlación total (r total) y la probabilidad de error (P _{error}). DC: diámetro del capítulo, PGC: gramos de grano por capítulo, P1000: peso de 1000 granos, CC: contenido de cáscara del grano, GG: grosor del grano, L:G: relación entre largo y grosor del grano, L:A: relación entre largo y ancho del grano, DGM: diámetro geométrico medio del grano, ESF: esfericidad del grano.	62
Tabla 9. Representación de los vectores en el análisis de componentes principales para cada característica. CP: Componente principal, SC: suma de cuadrados.....	63
Tabla 10. Valores del índice de selección (parámetro Fi) y del factor visual fenotípico (FV) del material de mejora (F ₄). L: A: relación largo y ancho del grano, GC: granos por capítulo, CC: contenido de cáscara del grano, AP: altura de planta.	67
Tabla 11. Contenido de ácido oleico y linoleico de la población de prueba (F ₅ + Testigos WSRC03 y Montola 2000) cultivada en 2018 en ACA (Cabildo) y la composición alélica del marcador KASP.....	68

Tabla 12. Genotipos de cártamo evaluados en los ensayos multiambientales. AO: alto oleico, AL: alto linoleico.....	84
Tabla 13. Nivel de materia orgánica (MO), nitrógeno total (Nt) y fósforo extraíble (Pe) en cada localidad. ACA: Asociación de Cooperativas Argentinas, Cabildo (2018); NAP: Campo Napostá-UNS, Torquinst (2018 y 2019); AgroUNS: Universidad Nacional del Sur, Bahía Blanca (2019).	91
Tabla 14. ANDEVA de la variable rendimiento de materia grasa de 12 genotipos de cártamo evaluados en cuatro ambientes del sudoeste bonaerense. GL: grados de libertad, SC: suma de cuadrados, CM: cuadrado medio, GXA: interacción genotipo- ambiente.....	94
Tabla 15. Rendimiento medio de materia grasa (MG), scores para cada uno de los CPs (componentes principales), valor de estabilidad (ASV, AMMI <i>stability value</i>), <i>ranking</i> de rendimiento medio de materia grasa (RR), <i>ranking</i> del valor de estabilidad (RASV) y GSI (<i>genotype selection index</i>) de 12 genotipos de cártamo evaluados en el sudoeste bonaerense.	100
Tabla 16. Porcentajes de explicación de los factores aleatorios (Genotipo, GXA, Bloque y error residual) sobre varianza total. MA: modelo multiambiental, MAf: modelo multiambiental filtrado por <i>outliers</i> . ACA: Asociación de Cooperativas Argentinas, Cabildo; NAP: Campo Napostá-UNS, Torquinst; AgroUNS: Universidad Nacional del Sur, Bahía Blanca.	101
Tabla 17. BLUPs y BLUPt del rendimiento de materia grasa por hectárea de 12 genotipos de cártamo en modelos de un ambiente. El sombreado verde indica superioridad y el sombreado rojo indica inferioridad significativa con respecto a la media utilizando la prueba t. ACA: Asociación de Cooperativas Argentinas, Cabildo (2018); NAP: Campo Napostá-UNS, Torquinst (2018 y 2019); AgroUNS: Universidad Nacional del Sur, Bahía Blanca (2019). .	103
Tabla 18. Correlaciones de Pearson entre los predictores BLUP derivados de los modelos de un ambiente, del modelo multiambiental (MA) y del modelo multiambiental con filtrado de datos (MAf). ACA: Asociación de Cooperativas Argentinas, Cabildo (2018); NAP: Campo Napostá-UNS, Torquinst (2018 y 2019); AgroUNS: Universidad Nacional del Sur, Bahía Blanca (2019).	104
Tabla 19. Componentes de varianza y su participación porcentual. GXA: interacción genotipo-ambiente, C5RS: capítulos en las cinco ramificaciones superiores, P1000: peso de mil granos	107
Tabla 20. Coeficientes de correlación de Pearson del rendimiento de materia grasa con los componentes de rendimiento (Cap: capítulos por planta, GC: granos por capítulo, P1000: peso de mil granos) y variables morfológicas (AP: altura de planta, R1: ramificaciones, C5RS: capítulos en las cinco ramificaciones superiores, DC: diámetro de capítulo, PGC: peso de grano por capítulo) medidas para cada genotipo (n=12).....	109
Tabla 21. Descripción de los materiales vegetales utilizados.....	122
Tabla 22. ANDEVA de las variables fenológicas días transcurridos y grados días acumulados (GDD) entre siembra y antesis (SA) en respuesta a localidad (L), fecha de siembra (FS) y genotipo (G).	126
Tabla 23. Resultados del ANDEVA. FDV: Fuentes de varianza, AP: altura de planta (cm), R1°: número de ramificaciones primarias, DC: diámetro de capítulo principal de la ramificación primaria (mm), MG: materia grasa (%), C: número de capítulos por planta, C5RS: número de capítulos en las cinco ramificaciones superiores, GC: número de granos por capítulo, P1000: peso de 1000 granos (g), MG: materia grasa (%), y Rend: rendimiento de grano (kg ha ⁻¹).	128

- Tabla 24. Correlaciones de Pearson entre rendimiento de grano, componentes de rendimiento y caracteres morfológicos de la planta en promedio para cada fecha de siembra (n=6). Rend: rendimiento de grano (kg/ha), MG: materia grasa (%), P1000: peso de 1000 granos (g), AP: altura de planta (cm), R1°: número de ramificaciones primarias, C: número de capítulos por planta, C5RS: número de capítulos en las cinco ramificaciones superiores, DC: diámetro de capítulo principal de la ramificación primaria (mm), y F/C: número de granos por capítulo..... 136
- Tabla 25. Variables fenotipadas en los parentales y en la población de mapeo población F₂ de cártamo (WSRC03 x Montola 2000) cultivados en 2018 en ACA (Cabildo)..... 147
- Tabla 26. Estadística descriptiva de una población F₂ de cártamo (WSRC03 x Montola 2000) cultivada en 2018 en ACA (Cabildo). (n, media, desvío estándar, mínimo y máximo) y prueba de normalidad de Saphiro-Wilks (estadístico W y error). Comparación de medias de los parentales por prueba T (Valores en negrita representan diferencias significativas entre los parentales con p<0,05). AP: altura de planta, R1: número de ramificaciones, C5RS: capítulos en las cinco ramificaciones superiores, DC: Diámetro de capítulos, RENDP: rendimiento por planta, P1000: peso de mil granos, GC: granos por capítulo, PGC: peso de granos por capítulo, CC: contenido de cáscara, CLG: contenido de lignina del grano, LG: largo del grano, AG: ancho del grano, GG: grosor del grano, L:A: Relación largo-ancho del grano, L:G: relación largo grosor del grano, DIAMF: diámetro de grano, VOLG: volumen de grano, DENSG: densidad de grano, GD_S_A: grados día siembra-antesis, GD_S_BBCH79: grados día siembra-BBCH79, GD_A_BBCH79: grados día antesis BBCH79..... 152
- Tabla 27. Descripción de QTLs hallados a partir de un mapa de ligamiento generado sobre una población de mapeo F₂ de cártamo (n=91, WSRC03 x Montola 2000) fenotipada en 2018 en ACA (Cabildo). R1: número de ramificaciones, C5RS: capítulos en las 5 ramificaciones superiores, DC: Diámetro de capítulos, RENDP: rendimiento por planta, P1000: peso de mil granos, PGC: peso de granos por capítulo, CC: contenido de cáscara, CLG: contenido de lignina del grano, LG: largo del grano, AG: ancho del grano, GG: grosor del grano, L:A: Relación largo-ancho del grano, L:G: relación largo grosor del grano, DIAMG: diámetro de grano, VOLG: volumen de grano, A y B: homocigota, H: heterocigota, d/a: relación dominancia/aditividad, A: aditivo, D: dominante, PD: parcialmente dominante, PR: parcialmente recesivo, SubD: subdominante, SobD: sobredominante. QTLs de una misma característica fueron enumerados en su nomenclatura..... 158

ÍNDICE DEL ANEXO

Tablas

- Tabla anexa 1. Esperanza de los cuadrados medios de cada componente de varianza. σ^2 : componente de varianza debido al error, σ^2_{GXA} : componente de varianza debido a la interacción genotipo- ambiente, σ^2_{bloq} : componente de varianza debido al bloque, σ^2_{gen} : componente de varianza debido al genotipo, Θ^2_{loc} : efecto fijo debido a la localidad, b: número de bloques, J: número de genotipos, a: número de ambientes..... 203
- Tabla anexa 2. ANDEVA de la variable materia grasa de 12 genotipos de cártamo evaluados en cuatro ambientes del sudoeste bonaerense. GL: grados de libertad, SC: suma de cuadrados, CM: cuadrado medio, GXA: interacción genotipo- ambiente..... 203
- Tabla anexa 3. Fuentes de varianza del modelo AMMI del rendimiento de materia grasa de 12 genotipos de cártamo evaluados en cuatro ambientes del sudoeste bonaerense. GL: grados de libertad, SC: suma de cuadrados, CM: cuadrado medio, F: test de Fisher, P_{error} : probabilidad de error, GXA: interacción genotipo- ambiente, %: proporción de la interacción explicada por cada componente..... 204
- Tabla anexa 4 Test de heretocedasticidad para el modelo mixto del rendimiento de materia grasa de 12 genotipos de cártamo evaluados en cuatro ambientes del sudoeste bonaerense realizado mediante la función *ranova* () del paquete lmerTest. npar: número de parámetros, logLik: logaritmo de la verosimilitud, AIC: medida de ajuste de Akaike, LRT: test de coeficiente de verosimilitudes con distribución chi-cuadrado, GL: grados de libertad, $\text{Pr}(> \text{Chi})$: probabilidad de error. 204

Figuras

- Figura anexa 1. Gráficos de cajas con datos atípicos (círculos) del rendimiento de materia grasa (MG) por hectárea de 12 genotipos de cártamo evaluados en cuatro ambientes del sudoeste bonaerense. MG.ha: rendimiento de MG en $\text{kg} \cdot \text{ha}^{-1}$ 205
- Figura anexa 2 a-l. Correlograma de los datos de rendimiento medio de MG de cada genotipo, las características morfológicas de planta y los componentes de rendimiento. Los índices de correlación y su significancia se encuentran en el triángulo superior de la matriz, el gráfico de puntos con los datos (n=12) en el triángulo inferior de la matriz y los boxplot en la diagonal. MG.ha: rendimiento de MG en $\text{kg} \cdot \text{ha}^{-1}$, AP: altura de planta, R1: ramificaciones, C: capítulos, C5RS: capítulos en las cinco ramificaciones superiores, DC: diámetro de capítulo, GC: granos por capítulo, PGC: rendimiento por capítulo, P1000: peso de mil granos. *: $p < 0,05$, **: $p < 0,01$, ***: $p < 0,001$ 211

RESUMEN

El cártamo es un cultivo oleaginoso poco difundido en Argentina con potencial de desarrollo en regiones semiáridas de la región pampeana. Las características requeridas por la industria y los productores agropecuarios son la base de los objetivos del mejoramiento. El ideotipo para el sudoeste bonaerense presenta alto contenido de aceite alto oleico, alto rendimiento, precocidad, adaptación a siembras de otoño-invierno y al ambiente en general. El objetivo principal de la presente tesis fue el desarrollo de criterios de selección para la obtención de líneas avanzadas para el mercado o para futuros planes de mejoramiento. Además, se buscó una herramienta molecular para seleccionar por el perfil ácido de la materia grasa (MG) y generar nuevas herramientas moleculares para cártamo a través del mapeo de QTLs en una población biparental.

A partir de germoplasma F₃ proveniente de un plan de mejoramiento del USDA (Pullman, Washington), se llevó a cabo un plan de selección a campo con el fin de obtener líneas avanzadas promisorias. En generaciones tempranas (F₃-F₄) se utilizaron variables fácilmente medibles como herramientas de selección indirecta del contenido de materia grasa (MG). La selección en F₃ se realizó por adaptación, precocidad y bajo espesor de pericarpio, y en F₄ se realizó por contenido de cáscara, altura de planta, granos por capítulo y las dimensiones del grano.

En generaciones avanzadas (F₅-F₆), la evaluación a campo en ensayos comparativos en cuatro ambientes representativos del sudoeste bonaerense permitió seleccionar de forma directa por el rendimiento de MG y su estabilidad. L62, L6 y L69 mostraron un alto y estable rendimiento de MG, pero sin alcanzar el contenido de MG que rige la norma (33%). LFM3 y L3 exhibieron buena adaptación al sudoeste bonaerense, alto rendimiento de MG, alto contenido de MG, alto contenido de ácido oleico y mayor precocidad que el testigo comercial CW99-OL. El desarrollo de nuevos cultivares se acompañó con la generación de una norma de manejo en cuanto a la fecha de siembra. La evaluación de fechas de siembra en dos ambientes permitió concluir que la siembra entre fines de mayo e inicio de junio optimiza el rendimiento del cultivo.

El perfil de ácidos grasos es el carácter cualitativo más importantes del cártamo, para el cuál existe gran variabilidad genética y determina las características industriales del aceite. Las nuevas tecnologías de secuenciación generan una rápida y económica identificación de marcadores moleculares e impulsaron la selección asistida por marcadores. En este trabajo se utilizó la tecnología KASP (*kompetitive allele specific PCR*) para validar un marcador basado en la secuencia del gen responsable en la determinación del perfil acídico del aceite.

La secuenciación de última generación aplicada al genotipado por secuenciación permite descubrir millones de SNPs (polimorfismos de un solo nucleótido) en una población de mapeo. La integración de densos mapas genéticos con datos fenotípicos permite la identificación de loci de caracteres cuantitativos (QTLs). A través del genotipado de una población F₂, se identificaron 673 marcadores polimórficos, se generó un mapa de ligamiento, y sobre éste, se mapearon 31 QTLs: siete QTLs asociados a características de la cascara, cinco a arquitectura, morfología y rendimiento de planta, ocho a dimensiones de grano y capítulo, y 11 a características asociadas al tamaño de grano. Los marcadores flanqueantes a estas regiones genómicas representan un nuevo recurso para el mejoramiento, y son un punto de partida para el mapeo fino de las regiones involucradas.

Esta tesis es el primer estudio genómico de cártamo realizado en Argentina y el primer plan de selección llevado a cabo en el sudoeste de Buenos Aires con la finalidad de obtener nuevos genotipos adaptados, generando a su vez una recomendación de fecha de siembra óptima para la región en cuestión.

ABSTRACT

Safflower is an alternative oilseed crop in Argentina, potentially adapted to the semiarid “Pampa” region. Crop characteristics required by the industry and farmers like yield, oil content or oil quality are the target traits for breeding programs. Safflower ideotype for the southwest of the Buenos Aires province should be high yielding with high oil content and high oleic acid content. Earliness and adaptation to fall sowing are also required. The aims of this thesis were the development of new advanced breeding lines to be registered or as a source of initial germplasm for future breeding programs; the development of a molecular tool for selecting the high oleic allele and QTL mapping of yield and oil content related traits in a biparental population obtained from the hybridization of contrasting varieties.

Safflower F₃ germplasm from one of the USDA (Pullman, Washington) breeding programs was the initial material for the field selection plan. In early breeding lines (F₃-F₄), selection was carried out through simple traits, that were used as indirect selection tools for a complex trait like oil content. Selection was done through adaptation, earliness and hull thickness in F₃ generation, and through hull content, plant height, number of grains per capitulum and grain dimensions in F₄. After the two selection cycles, 16 promising F₅ breeding lines were obtained. In advanced generations (F₅-F₆), selection for high oil yield and stability was accomplished through field evaluation in multi environmental trials. L62, L6 and L69 showed high and stable oil yield, but an oil content below the commercial base. L3 and LFM3 were the best adapted and high yielding breeding material, with high oleic oil content and an earlier development the commercial check CW 99 OL. The development of new genotypes was integrated with the exploration of optimum sowing dates. From the evaluation of six sowing dates in two environments it was concluded that sowings during late May and early June have the highest and more stable grain yield.

The fatty acid profile is the most important qualitative trait of safflower since it determines the oil industrial quality and is genetically diverse across the germplasm. The new sequencing technologies applied to molecular marker discovery have promoted marker assisted selection procedures. In this work,

KASP (kompetitive allele specific PCR) technology was used to validate a molecular marker based on the sequence of ctFAD2-1 gene, involved in the oleic acid content of safflower oil.

Next-generation sequencing technologies applied to genotyping by sequencing allow fast and massive SNP discovery in a mapping population. The integration of dense genetic maps with phenotypic data is used in QTL (quantitative trait loci) mapping. In this study, an F₂ population was genotyped by Dartseq method and almost 700 polymorphic markers were identified and a linkage map was constructed. Thirty-one (31) QTLs were mapped: seven of them were related to hull characteristics, five to architecture and morphology of the plant, eight to grain and capitulum dimensions, and 11 to grain size. Molecular markers flanking these genomic regions represent a new genetic resource for breeding and are starting points for further studies like fine mapping of the involved regions.

This thesis represents the first genomic approach in safflower of Argentina, and the first attempt of breeding by selection in the southwest of Buenos Aires province with the objective of developing new adapted genotypes and providing a basic management tool of sowing date.