

**Aus dem Institut für Pflanzenbau und Pflanzenzüchtung I
der Justus-Liebig-Universität Gießen
Lehrstuhl für Pflanzenzüchtung
Prof. Dr. Dr. h.c. Wolfgang Friedt**

**Identifizierung und Charakterisierung
samenfaserkorrelierter Gene in Raps (*Brassica
napus* L.)**

Dissertation zur Erlangung des akademischen Grades
Doctor agriculturae
– Dr. agr. –

im Fachbereich
Agrarwissenschaften, Ökotropologie und Umweltmanagement der
Justus-Liebig-Universität Gießen

vorgelegt von

Anna Stein
aus Oberursel, Hessen

Gießen, im Mai 2013

Inhaltsverzeichnis

Einleitung.....	9
Raps und seine wirtschaftliche Bedeutung	9
Rapsöl: Zusammensetzung, Verwendung, züchterische Verbesserung.....	10
Rapsschrot: Zusammensetzung, Verwendung, züchterische Verbesserung.....	11
Rohfaser.....	13
Samenfarbe: Bedeutung und genetische Determinierung.....	16
Allopolyploidie und ihre Bedeutung für die Züchtung von Raps	18
Molekulargenetik als Werkzeugkasten der Züchtung: Markergestützte Selektion und <i>Next Generation Sequencing</i>	20
Kandidatengenansatz: Ligninsynthesegene <i>Cinnamylalkoholdehydrogenase (CAD)</i> und <i>Cinnamoyl-CoA-Reduktase (CCR)</i>	22
Zielsetzung der Arbeit.....	23
Publikationen.....	25
Diskussion	49
Mögliche funktionelle Bedeutung der QTL	49
Bewertung der genetischen Kartierung und QTL-Bestimmung: Möglichkeiten und Grenzen der Methoden.....	51
Chromosomale Umstrukturierungen von Kandidatengen-Regionen und ihre phänotypischen Auswirkungen	53
Aminosäuresubstitutionen in BnCAD2/CAD3 und BnCCR1	56
Chromosomenmutationen und ihre Bedeutung für die Züchtung	57
Neue DNA-Sequenzierungstechniken in Forschung und Züchtung	59
Schlußbetrachtung	62
Zusammenfassung.....	63
Summary.....	65
Literatur	67
Anhang.....	77

Danksagung	81
-------------------------	-----------

Eidesstattliche Erklärung	83
--	-----------

Verzeichnisse

Abbildungen

Abb. 1: Weltweite Produktionsmengen der wichtigsten Ölpflanzen.....	9
Abb. 2: Weltweite Anbauflächen der wichtigsten Ölpflanzen.....	9
Abb. 3: Ligninbiosynthese in <i>Arabidopsis thaliana</i>	14
Abb. 4: Samenschalenaufbau und Phänotyp in <i>Arabidopsis thaliana</i>	17

Tabellen

Tab.1: Inhaltsstoffe von Raps- und Sojaextraktionsschrot	12
--	----

Abkürzungen

ADF	<i>acid detergent fibre</i> , säuredetergente Faserkomponenten
ADL	<i>acid detergent lignin</i> , säuredetergentes Lignin
At	<i>Arabidopsis thaliana</i>
BAC	<i>bacterial artificial chromosome</i> , künstliches Bakterienchromosom
ban	banyuls
Bn	<i>Brassica napus</i>
Bo	<i>Brassica oleracea</i>
Br	<i>Brassica rapa</i>
Ca	Kalzium
CAD	Cinnamylalkoholdehydrogenase
CanOLA	<i>Canadian Oil Low Acid</i>
CCR	Cinnamoyl-CoenzymA-Reduktase
cM	centimorgan
Cu	Kupfer
DNA	<i>deoxyribonucleic acid</i> , Desoxyribonukleinsäure
dt	Dezitonne
ExV8-DH	Doppelhaploidpopulation Express 617*V8
g	Gramm
ha	Hektar
HEAR	<i>High Erucic Acid Rape</i>
Heynh.	Gustav Heinhold, Botaniker
HNRT	homöologe nichtreziproke Translokation
HOLLI	<i>High Oleic Low Linolenic</i>
HRT	homöologe reziproke Translokation
K	Kalium
L.	Carl von Linné, Botaniker
MAS	<i>marker-assisted selection</i> , marker-gestützte Selektion
Mb	Megabase
ME	<i>metabolizable energy</i> , verwertbare Energie, Einheit MJ
mg	Milligramm
Mio	Millionen
MJ	Megajoule
mm	Millimeter
mmol	Millimol

μm	Mikrometer
μmol	Mikromol
Na	Natrium
NDF	<i>neutral detergent fibre</i> , neutraldetergente Faserkomponenten
NIRS	Nahinfrarotspektrometrie
NRT	nichtreziproke Translokation
P	Phosphor
QTL	<i>quantitative trait locus</i>
r	Korrelationskoeffizient
RNAi	RNA-Interferenz, RNA = <i>ribonucleic acid</i>
Rp.	Rohprotein
SNP	<i>single nucleotide polymorphism</i> , Einzelnukleotidpolymorphismus
spp.	<i>species pluralis</i> , mehrere Species einer Gattung
syn.	synonym
t	Tonne
tds	tannin deficient seed
tt	transparent testa
ttg	transparent testa glabra
YE2-DH	Doppelhaploidpopulation Express 617*1012-98
Zn	Zink
χ ²	Chi-Quadrat

Einleitung

Hier wird zunächst die Kulturpflanze Raps vorgestellt. Raps wird hierzulande primär als Winterform zur Ölgewinnung angebaut. Das Öl findet je nach Fettsäuremuster verschiedene Anwendungen in Chemie und Technik sowie im Futter- und Nahrungsmittelsektor. Die Preß- und Extraktionsrückstände stellen ein wertvolles Eiweißfuttermittel dar. Daher werden nach einem kurzen Abriß der wirtschaftlichen Bedeutung und der Entwicklung moderner Rapssorten die Inhaltsstoffe mit einem besonderen Fokus auf die antinutritiven Komponenten dargestellt. Hier wird der Schwerpunkt auf die Qualitätsparameter Samenfarbe und Lignin gelegt. Daraus leitet sich die Wahl der in dieser Arbeit untersuchten Kandidatengene ab. Ergänzend beleuchtet ein Kapitel die genetische Konstitution verbunden mit den entsprechenden Besonderheiten von Raps.

Raps und seine wirtschaftliche Bedeutung

Raps ist heute eine der bedeutendsten Ölpflanzen im weltweiten Anbau. Dabei beinhaltet „Rapssaat“ neben dem Raps im engeren Sinne (*Brassia napus*) noch die verwandten Arten *B. rapa* (Rübsen), *B. carinata* (Abessinischer Senf) und *B. juncea* (Indischer Senf). Rapssaat nimmt nach Sojabohne, Ölpalme und Baumwolle global den vierten Rang in der Herstellung von Ölsaaten und Schroten ein. Die weltweite Produktion von Rapssaat lag 2011 bei 62 Mio. t auf einer Anbaufläche von 34 Mio. ha (Abbildungen 1 und 2).

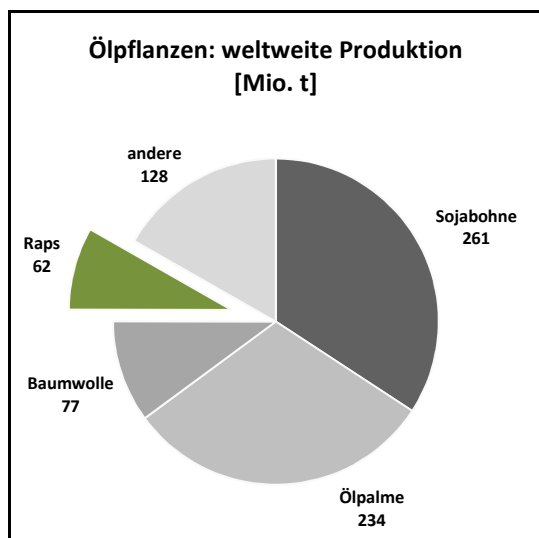


Abb. 1: Weltweite Produktionsmengen der wichtigsten Ölpflanzen (faostat.fao.org)

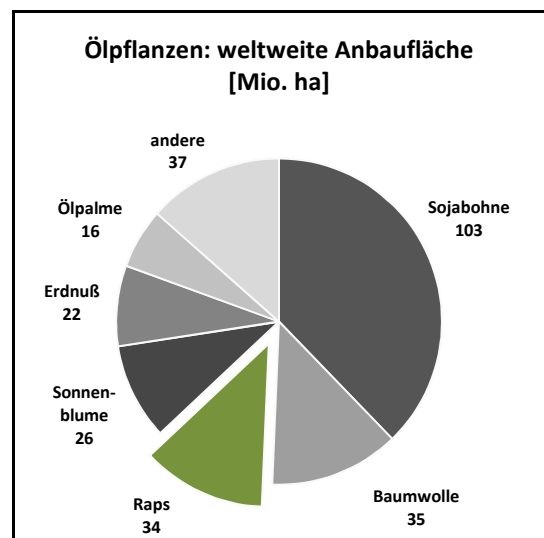


Abb. 2: Weltweite Anbauflächen der wichtigsten Ölpflanzen (faostat.fao.org)

Haupterzeugerregionen von Rapssaat waren 2011 die Europäische Union (19 Mio. t), Kanada (14 Mio. t), China (13 Mio. t) und Indien mit 8 Mio. t (vorrangig *B. juncea*). In Deutschland wurde im Jahr 2011 auf einer Anbaufläche von 1,3 Mio. ha ein Durchschnittsertrag von ca. 29 dt/ha Rapssaat erzielt (FAOSTAT 2013, faostat.fao.org). Aufgrund der ungünstigen Witterung lag dieser Wert jedoch deutlich unter dem mehrjährigen Mittel von 36,6 dt/ha (Jahre 2007 bis 2011).

Rapsöl: Zusammensetzung, Verwendung, züchterische Verbesserung

Der Ölgehalt der Rapssaat (*B. napus*) variiert zwischen 40 und 50%. Es überwiegt der Anteil ungesättigter Fettsäuren. Ölsäure (C18:1) ist darunter mit 50 bis 65% die wichtigste Monoenfettsäure. Linolsäure (C18:2) mit einem Anteil von 15 bis 30% und Linolensäure (C18:3) mit 5 bis 13% Anteil ergänzen das ernährungsphysiologisch günstige Profil des Rapsöls in Niedrigerucasäuretypen. Hier liegt der Erucasäuregehalt (C22:1) bei maximal 2%. Durch Senkung des Erucasäuregehaltes auf unter 2% im Öl entstanden in den 1970er Jahren zunächst die 0- bzw. CanOLA-Rapssorten (*Canadian Oil Low Acid*). Durch züchterische Reduzierung des Glucosinolatgehaltes um 80% auf maximal 30 µmol/g entfettete Trockenmasse wurden in den 1980er Jahren Sorten mit Doppelnullqualität (00) entwickelt, die sich für den Einsatz von Öl und Schrot in der menschlichen bzw. tierischen Ernährung eignen. Als Basis dafür diente die polnische Sommerrapssorte „Bronowski“ mit niedrigem Glucosinolatgehalt, die zur Einkreuzung dieses quantitativen Merkmals verwendet wurde. Auf Bronowski gehen sämtliche heutigen 00-Sorten zurück (Abadi und Leckband 2011). Öl aus 00-Raps wird in der Nahrungsmittelindustrie als hochwertiges Speiseöl oder zur Herstellung von Margarine verwendet. Hochölsäuretypen mit 70 bis 90% Ölsäureanteilen und geringem Linolensäuregehalt, sogenannte HolLi-Typen (*high oleic low linolenic*) sind zur Gewinnung von Frittieröl sowie für industrielle Anwendungen von Bedeutung, da dieses Öl besonders oxidationsstabil ist (Abadi und Leckband 2011). Generell eignen sich Rapsöle mit einem hohen Anteil an gesättigten Fettsäuren oder einfach ungesättigten Fettsäuren, den Monoensäuren, als Kraftstoff bzw. für die Gewinnung von Rapsmethylester (RME), sogenanntem Biodiesel. Daneben stehen für industrielle Anwendungen hoch-erucasäurehaltige Rapssorten (HEAR, *high erucic acid rape*) zur Verfügung. Diese zeichnen sich durch hohe Gehalte an Erucasäure von 40 bis 64% zu Lasten der Ölsäure mit nur 13 bis 38% Anteil aus (Schuster 1992).

Rapsschrot: Zusammensetzung, Verwendung, züchterische Verbesserung

Raps liefert neben dem Öl für die menschliche Ernährung und chemisch-technische Verwendung mit dem Rapsschrot ein hochwertiges eiweißreiches Futtermittel (Rahman und Mc Vetty 2011). Der besondere Wert des Rapsschrotes liegt insbesondere in der günstigen Aminosäurezusammensetzung der Eiweißfraktion und dem Gehalt an wertvollen Mineralstoffen (Tabelle 1).

Im Vergleich zur Sojabohne ist Rapsschrot in den Parametern Bruttoenergie und Eiweiß ungünstiger zu bewerten. Darüber hinaus ist der Fasergehalt deutlich höher. Neben dem Rohfasergehalt limitieren weitere antinutritive Inhaltsstoffe den Einsatz des Rapsschrotes in der Tierernährung: Glucosinolate haben neben einem unangenehmen Geschmack auch einen negativen Effekt auf das Wachstum und die Gesundheit monogastrischer Tiere, insbesondere der Schweine. Der Anteil an Glucosinolaten liegt bei etwa 30 µmol/g entfettete Trockenmasse, das entspricht etwa einem Fünftel dessen, was frühere Rapsorten, d.h. Erucasäure- und 0-Typen aufwiesen. Zu den sogenannten antinutritiven Inhaltsstoffen gehören weiterhin Tannine, Phytinsäuren und Sinapine, die zusammen etwa 10% des Rapsschrotes ausmachen. Braune Legehennen, die einen rezessiv vererbten Defekt einer Monooxygenase tragen, können das bei der Verdauung von Sinapin aus Rapsschrot entstehende Trimethylamin nicht oxidieren und ausscheiden. Sie akkumulieren es in ihren Eiern, was jedoch zu einem unerwünschten fischigen Aroma der Eier führt (Ward et al. 2009). Züchterisch konnte dieses Problem sowohl auf Seiten der Rapszüchtung als auch in der Legehennenzucht gelöst werden. Rapsgene wie UDP-Glucose:Sinapat-Glucosyltransferase 1 (*BnSGT1*), 1-O-Sinapoyl-β-Glucose:Cholin-Sinapoyltransferase 1 und 2 (*BnSCT1* und *BnSCT2*), deren Transkripte an der Sinapinestersynthese, insbesondere von Sinapinoylcholin (syn. Sinapin) beteiligt sind, wurden identifiziert und konnten durch Nutzung des RNA-Interferenzmechanismus (RNAi) stillgelegt werden. Dies bewirkte eine Reduzierung des Sinapingehaltes (Hüsken et al. 2005, Weier et al. 2008). Bei Hennen konnte erfolgreich gegen das homozygote Auftreten der funktionellen Punktmutation selektiert werden.

Vor allem in dunkelsamigen Rapsgenotypen inhibieren die hier reichlich vorkommenden Tannine die Eiweißhydrolyse und reduzieren somit die Proteinverdaulichkeit. Phytinsäuren akkumulieren im Samenembryo und binden Phosphor sowie einige essentielle Minerale, die dadurch einer Verwertung durch das Tier entzogen werden (Bell 1993). Kondensierte Tannine in der Samenschale führen zu einer kontinuierlichen Variation der Samenfarbe in Abhängigkeit von der Samenschalendicke und von Umwelteinflüssen wie der Temperatur (Rahman und McVetty 2011). Marles und Gruber (2004) fanden in ihrer histochemischen Untersuchung verschiedener Brassicaceen, daß Gelbsamigkeit mit einer geringeren Samenschalenmasse, einem reduzierten Proanthocyanidingehalt und einem geringen

Ligningehalt in der Samenschale assoziiert ist. Proanthocyanidine reichern sich in Vesikeln des endoplasmatischen Retikulums oder in der Vakuole der Zellen des inneren Integuments der Samenschale an, während Lignin üblicherweise in der Zellwand der Palisadenzellen akkumuliert. Die räumliche Nähe der Kompartimente kann insbesondere während der Samendehydration zur Reife hin dazu führen, daß Proanthocyanidine und Lignin vernetzen. Dies kann die quantitative Bestimmung der Proanthocyanidin-Fraktion beeinträchtigen, da deren tatsächlicher Gehalt so u. U. unterschätzt wird (Marles und Gruber 2004).

Tab.1: Inhaltsstoffe von Raps- und Sojaextraktionsschrot bezogen auf 1kg Schrot mit einem Trockensubstanzgehalt von 88% (nach Lindermayer et al. 2010)

Inhaltsstoffe	Einheit	Rapsextraktionsschrot	Sojaextraktionsschrot
ME	MJ	10.71	12.9
Rohprotein	g	332	442
Lysin	g / % i. Rp.	17.1 / 5.2	26.7 / 5.9
Methionin	g / % i. Rp.	7.0 / 2.1	5.9 / 1.3
Cystein	g	8.6	6.5
Threonin	g / % i. Rp.	15.2 / 4.6	17.1 / 3.9
Tryptophan	g / % i. Rp.	5.2 / 1.6	5.9 / 1.3
Rohfett	g	40	12
Polyensäuren	g	8.0	8.0
Rohfaser	g	128	70
Stärke	g	44	62
Zucker	g	70	95
Rohasche	g	67	59
Ca	g	7.4	2.7
P	g	10.7	5.7
Na	g	0.6	0.3
K	g	12.9	19.0
Cu	mg	3.9	17.0
Zn	mg	63	62.0
Glucosinolate	mmol	4.9	-

Gelbsamigkeit wird mit einer dünneren Samenschale und niedrigerem Tanningehalt sowie einem höheren Öl- und Proteingehalt verbunden und ist daher seit bald 40 Jahren ein Zuchtziel, vor allem in Kanada und China, aber auch in Frankreich, Polen und Deutschland (Friedt und Snowdon 2009, Rahman und McVetty 2011, Somers et al. 2001, Shirzadegan et

al. 1985, Stringam et al. 1974). In Übereinstimmung mit den genannten Befunden berichten Wittkop et al. (2012), daß die Samenfarbe auch mit dem Gehalt an „Rohfaser“ korreliert. Über den Zusammenhang und die gemeinsame genetische Determinierung von Samenfarbe und Rohfaser in Rapskörnern (Badani et al. 2006) ist heute jedoch noch immer nicht viel bekannt. Gesichert aber ist, daß eine Verringerung des Rohfasergehalts (*dietary fibre*) in der Samenschale sowie eine Reduktion der Schalendicke (Shirzadegan et al. 1985, Slominski et al. 1999) zu einer verbesserten Schrotqualität führt (Wittkop et al. 2012).

Rohfaser

Unter dem Begriff Rohfaser werden jene Substanzen zusammengefaßt, die Gerüstfunktionen in pflanzlichen Zellen haben. Dazu zählen Zellulose, Hemizellulose, Lignin, Pektine und Polyphenole, sowie im weiteren Sinne gebundene Polysaccharide und Asche. Das von Van Soest (1964) entwickelte Analyseverfahren unterscheidet Gruppen von Faserkomponenten nach ihrer Löslichkeit: Säuredetergentes Lignin (*acid detergent lignin, ADL*) beinhaltet demnach Lignin sowie Monolignole und Lignan. Im Weiteren wird in dieser Arbeit in der biochemischen Beschreibung diesbezüglich zusammenfassend von Lignin gesprochen, während der Begriff ADL im Zusammenhang mit der hier durchgeführten speziellen Phänotypisierungsmethode Anwendung findet. Säuredetergente Faserkomponenten (*acid detergent fibre, ADF*) beinhalten neben ADL auch Zellulose, während neutraldetergente Faserkomponenten (*neutral detergent fibre, NDF*) neben ADF auch Hemizellulose umfassen.

Lignin: Bedeutung und Biosynthese

Bei Lignin handelt es sich um ein komplexes Polymer, das im Wesentlichen aus den drei Monomeren *p*-Coumaryl-Alkohol, Coniferyl-Alkohol und Sinapyl-Alkohol (den sogenannten H-, G- und S-Lignin-Untereinheiten) zusammengesetzt ist. Neben Zellulose ist Lignin die meist verbreitete Kohlenstoffverbindung in der Biosphäre (Boerjan et al. 2003, Kirk und Farrell 1987). Das Polymer wird im Wesentlichen in ausdifferenzierten Zellen stützender und wasserführender Gewebe aller Gefäßpflanzen, insbesondere im Holz eingelagert (Freudenberg 1965). Lignin wird auch als Reaktion auf biotische und abiotische Streßfaktoren, wie mechanische Verletzung oder Schädlingsbefall, gebildet und ist so für die Pflanze essentiell. Daher ist es auch gegen enzymatischen und mechanischen Abbau weitgehend resistent. Dies schränkt jedoch die Verdaulichkeit stark ligninhaltiger Futtermittel ein. Auch industrielle Verwertungen, wie beispielsweise die Bioethanolgewinnung aus

verholzten Pflanzenteilen, sind nur begrenzt möglich (Chen und Dixon 2007, Moore und Jung 2001, Taherzadeh und Karimi 2008).

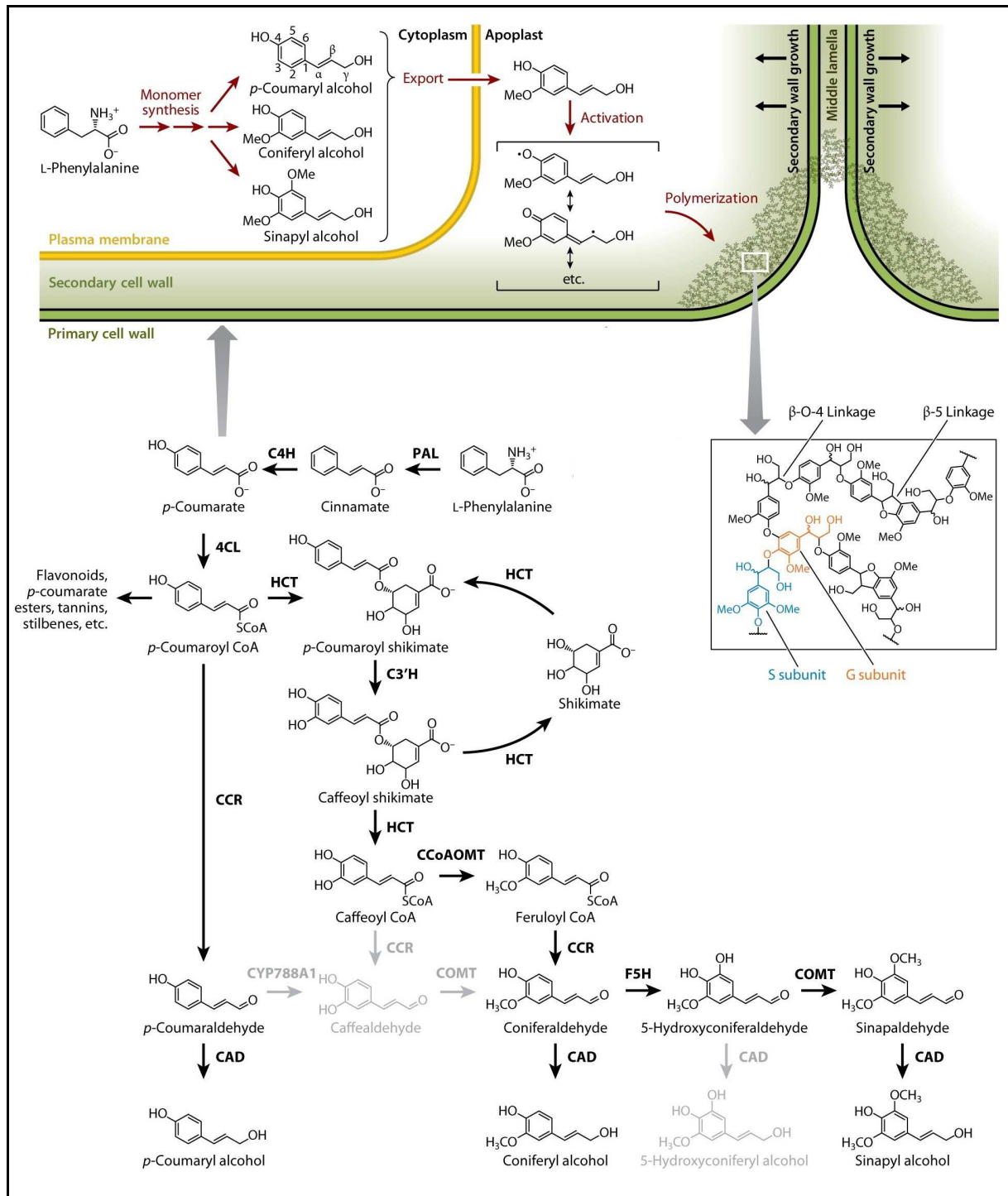


Abb. 3: Ligninbiosynthese in *Arabidopsis thaliana*, Grafik modifiziert nach Bonawitz und Chapple (2010): Deaminierung von Phenylalanin mittels Phenylalanin-Ammonium-Lyase (PAL), Coenzym A-Aktivierung durch 4-(Hydroxy-)Cinnamat-CoA-Ligase (4CL), Hydroxylierung an den Positionen 3,4 und/oder 5 im aromatischen Ring durch Cinnamat-4-Hydroxylase (C4H), *p*-Coumaroylshikimat-3'-Hydroxylase (C3'H) und Ferulat-5-Hydroxylase (F5H), Shikimatttransfer durch Hydroxycinnamoyl-CoA:Shikimat-Hydroxycinnamoyl-Transferase (HCT), sowie zwei Reduktionsschritte katalysiert durch Cinnamoyl-CoA-Reduktase (CCR) und Cinnamylalkoholdehydrogenase (CAD) liefern die Monolignole *p*-Coumaryl-Alkohol, Coniferyl-Alkohol und Sinapyl-Alkohol.

Die spezifische Zusammensetzung und Struktur des Lignins ist zwischen verschiedenen Pflanzenarten durchaus variabel. Aber auch innerhalb einer Art kann in Abhängigkeit vom Gewebe, dem Ort der Anreicherung innerhalb der Zelle, von Umwelteinflüssen und der Verfügbarkeit von Monomeren eine unterschiedliche Zusammensetzung des Lignins bezüglich der Ligninmonomere und ihrer Verknüpfungen vorliegen (Campbell und Sederoff 1996, Ralph et al. 2004, Moore und Jung 2001, Rogers und Campbell 2004). Die Ligninbiosynthese ist ein komplexer Prozeß, der insbesondere in seiner Regulation noch nicht vollständig aufgeklärt ist und für Raps bzw. Brassica-Arten bislang nur anhand von Daten für Modellorganismen wie *Arabidopsis* (*Arabidopsis thaliana* (L.) Heynh.) und Pappel (*Populus* spp. L.) abgeleitet werden kann. Generell liefert der Phenylpropanoid-Syntheseweg des Sekundärstoffwechsels in Pflanzen lösliche und unlösliche Polymere, u.a. Flavonoide, Tannine, Stilbene, Lignane und Lignin (Bonawitz und Chapple 2010, Umezawa 2010). Diese Polymere gehen jeweils auf die Vorstufe Phenylalanin zurück, eine Aminosäure aus dem Primärstoffwechsel. Eine Reihe enzymatischer Reaktionen führt letztlich zur Bildung der Monolignole *p*-Coumaryl-Alkohol, Coniferyl-Alkohol und Sinapyl-Alkohol. Die Abbildung 3 veranschaulicht den Biosyntheseweg und die räumliche Anordnung der Reaktionsschritte in der Zelle.

Die Monolignole werden auf noch nicht genau verstandene Weise in den Apoplasten transportiert, wo sie radikalisiert werden. Die Radikale gehen kovalente Bindungen ein und formen ein Polymer, das vor jeder weiteren Addition neuerlich radikalisiert werden muß. Die Struktur des Polymers wird durch die Verfügbarkeit der drei Monolignolspezies und ihrer Ringssubstitutionen bestimmt (zusammengefaßt in Bonawitz und Chapple 2010). Darüber hinaus sind auch alternative Polymerbildungen bei blockierten Syntheseschritten möglich. Hier werden anstelle der jeweiligen Produkte der Zwischenschritte die akkumulierten Substrate - also eigentlich Zwischenprodukte der Ligninsynthese - in das Ligninpolymer eingebaut (Goujon et al. 2003a, Halpin et al. 1994, Leplé et al. 2007, Piquemal et al. 1998, Ralph et al. 2001). Beispielsweise werden enzymatische Defekte von Cinnamoyl-CoA-Reduktase (CCR), Cinnamylalkoholdehydrogenase (CAD) und Coffeinsäure/5-Hydroxyferulasäure-O-Methyltransferase (COMT) teilweise durch die Inkorporation der jeweiligen Substrate kompensiert. Dieses Phänomen bezeichnen Bonawitz und Chapple (2010) als Ligninplastizität.

Samenfarbe: Bedeutung und genetische Determinierung

In vielen Mitgliedern der *Brassicaceae* ist eine gelbe Samenfarbe mit einem höheren Eiweiß- und Ölgehalt und einem reduzierten Rohfaseranteil korreliert (Simbaya et al. 1995). Dadurch sind der Energiegehalt und die Proteinverdaulichkeit gelbsamiger Genotypen höher. Dem gegenüber stehen die positiven Eigenschaften dunkelsamiger Genotypen: eine stärkere Krankheitsresistenz, längere Keimfähigkeit und bessere Lagerfestigkeit des Saatgutes. Flavonoide schützen darüber hinaus vor oxidativem Streß, Wasseraufnahme und Ausschwemmung gelöster Substanzen (Debeaujon et al. 2001, Halloin 1982). Auch die technische Bearbeitung ist für die vergleichsweise hartschaligen schwarzsamigen Rapsgenotypen optimiert. So wurden Schwierigkeiten bei der Ölpressung hellsamiger Genotypen beobachtet, da der Preßwiderstand des Schrotes und damit der Ölertrag aufgrund des reduzierten Faseranteils sank. Erste Versuche im Pilotmaßstab zeigten, daß im Rapskuchen aus gelber Saat deutlich erhöhte Mengen Öl verblieben (Mühmel et al. 2011).

Verantwortlich für die Ausprägung der Samenfarbe ist die Akkumulation von Flavonoiden, bzw. Polyphenolen (Polymere der Leucocyanidine). Flavonoide umfassen eine sehr heterogene Gruppe aromatischer Moleküle, die folgenden sieben Untergruppen zugeordnet werden: Aurone, Chalcone, Flavone, Flavonole, Flavandiole, Anthocyanine und Proanthocyanidine (kondensierte Tannine) (Winkel-Shirley 2001). Sie reichern sich in den Palisadenzellen des äußeren Integuments und den Parenchymzellen des inneren Integuments der Samenschale an (Stringam et al. 1974). Die innerste Schicht, das Endothelium, synthetisiert Proanthocyanidine, die zunächst in der Vakuole akkumulieren und durch spätere enzymatische und/oder chemische Oxidation zu der braunen Samenfarbe führen (Haughn und Chaudhury 2005). Entsprechend histologischer Studien von Beeckman et al. (2000) handelt es sich dabei jedoch um eine transiente Färbung, die sich im Verlaufe der Embryonalentwicklung wieder auflöst. Die endgültige Pigmentakkumulation findet in den Zellen des äußeren Integuments, den Palisadenzellen, statt. Diese Zellen komprimieren mit den Parenchymzellen des inneren Integuments bei der Reifung und bilden so die Pigmentschicht. Diese Schicht verbindet sich dann fest mit der äußeren Schicht des Endosperms und bildet eine Aleuronschicht (Moïse et al. 2005). Die Abbildung 4 zeigt diesen Zusammenhang am Beispiel von *A. thaliana* in einer Gegenüberstellung eines braunschaligen Wildtyps mit einer gelbschaligen Mutante (*tt12*), bei der die Akkumulation von Proanthocyanidinen in den Vakuolen der Endothelzellen gestört ist.

In gelbsamigen Rapslinien ist die Palisadenschicht um etwa zwei Drittel dünner als in schwarzsamigen Rapslinien. Dadurch ist auch der Anteil an deponierten Polyphenolen und Lignin in den Schalen gelbsamiger Linien geringer (Rahman 2001, Slominski et al. 1994,

Simbaya et al. 1995). Beeckman et al. (2000) führen die gelbe Samenfarbe auf eine schwache Pigmentierung der Kotyledonen und eine transparente Samenschale zurück. Umgekehrt ist mit der Ölräuke (*Eruca sativa*) bisher nur ein einziger Fall innerhalb der *Brassicaceae* beschrieben, bei dem eine schwarze Samenfarbe eindeutig auf die starke Färbung der Kotyledonen bei gleichzeitig transparenter Samenschale zurückzuführen ist (Marles und Gruber 2004).

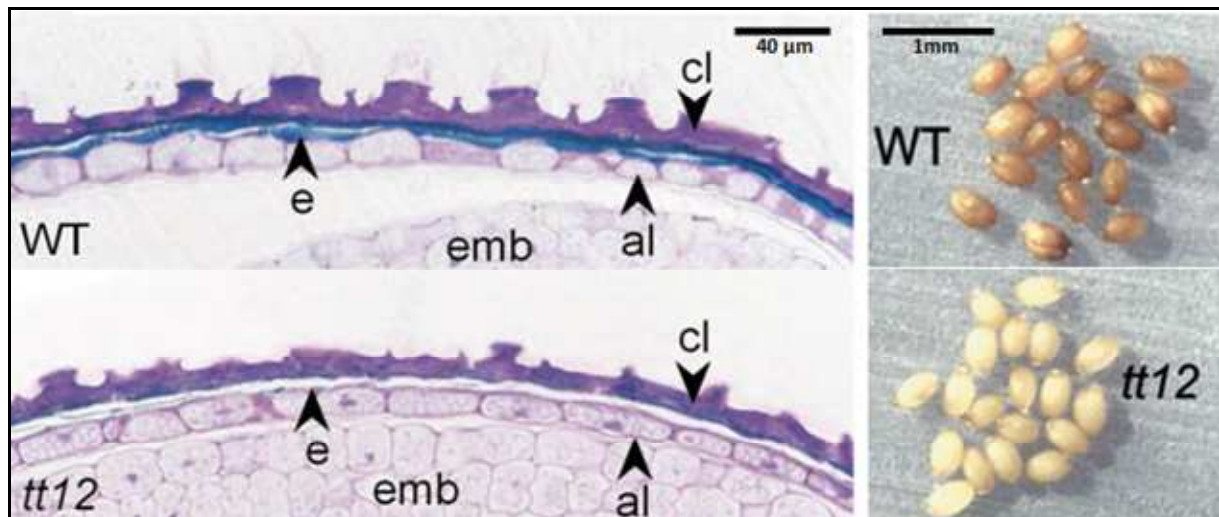


Abb. 4: Samenschalenaufbau und Phänotyp in *Arabidopsis thaliana* Wildtyp (WT) und der *tt12*-Mutante; blau: phenolische Verbindungen in der Samenschale, gefärbt durch Toluidinblau, al: Aleuronschicht, cl: vier komprimierte Schichten von Parenchymzellen, e: Endotheliumschicht, emb: Embryo; aus Debeaujon et al. 2001

Bei *A. thaliana* werden die meisten Mutationen, die zu Gelbsamigkeit führen, als *transparent testa*- oder *tt*-Mutationen bezeichnet. Mit Hilfe von *A. thaliana*-Mutanten mit veränderter Samenfarbe (*tt*, *transparent testa glabra* [*ttg*], *banyuls* [*ban*]) oder mit verändertem Proanthocyanidingehalt (*tannin deficient seed*, [*tds*]) konnten zahlreiche Gene identifiziert werden, die an der Differenzierung der Zellen des inneren Integuments beteiligt sind. Diese Gene können in zwei Gruppen unterteilt werden: Die erste umfaßt diejenigen Gene, die an der Biosynthese und Kompartimentierung der Flavonoidverbindungen beteiligt sind. Die zweite Gruppe beinhaltet Gene, die für Transkriptionsfaktoren kodieren. Zur ersten Gruppe gehören *TT3*, *TT4*, *TT5*, *TT6*, *TT7*, *BAN*, *TT12*, *TT19*, *TDS4/TT18* und *AUTOINHIBITED H⁺-ATPase ISOFORM 10 (AHA10)*. In der zweiten Gruppe sind bisher die regulatorischen Gene *TT1*, *TT2*, *TT8*, *TT16*, *TTG1* und *TTG2* bekannt (Haughn and Chaudhury 2005).

Natürliche gelbsamige Formen innerhalb der Gattung *Brassica* wurden in *B. rapa*, *B. carinata* und *B. juncea* beobachtet, nicht jedoch in *B. napus*. Das Merkmal Gelbsamigkeit muß daher durch interspezifische Kreuzung in Raps eingebracht werden. Dabei bereitet die genaue Phänotypisierung gewisse Schwierigkeiten, da gelbsamige Linien nicht immer eine homogene Färbung der Samenschale aufweisen und das Merkmal darüber hinaus stark

durch Umwelteinflüsse modifiziert werden kann. Die Vererbung der Samenfarbe scheint auf drei bislang nicht näher identifizierten Genen zu basieren, wobei offenbar in verschiedenen gelbsamigen Rapslinien teilweise unterschiedliche Gene für die Ausprägung der Samenfarbe verantwortlich sind (zusammengefaßt von Rahman und McVetty 2011).

Allopolyploidie und ihre Bedeutung für die Züchtung von Raps

Die Spezies *B. rapa* und *B. oleracea* entstanden vor etwa 4 Millionen Jahren (Rana et al. 2004). *B. rapa* besitzt das sogenannte A-Genom ($n=10$, 500Mb) und *B. oleracea* das C-Genom ($n=9$, 600Mb). Trotz dieser Unterschiede sind die Genome aber weitgehend kollinear. Diese großenteils einheitliche Struktur findet sich selbst im allopolyploiden Raps *B. napus* wieder, der vor etwa 10.000 Jahren aus der natürlichen Hybridisierung von Rübsen (*B. rapa*) und Kohl (*B. oleracea*) hervorgegangen ist (U 1935, Rana et al. 2004). Dabei addierten sich die Genome von *B. rapa* und *B. oleracea* zu dem von *B. napus* ($n=19$). Das gesamte Rapsgenom umfaßt ca. 1200Mb und ist damit zehnmal so groß wie das Genom von *A. thaliana* (120Mb). Vergleichende genetische Kartierung zwischen *B. napus* und *A. thaliana* zeigte, daß einige Regionen des Arabidopsis-Modellgenoms durch drei Kopien in jedem der diploiden Genome in *B. napus* repräsentiert sind. Dies bestätigt, daß die diploiden Brassica-Spezies offenbar aus einem gemeinsamen hexaploiden Vorfahren hervorgegangen sind (Parkin et al. 2003).

Chromosomale Umstrukturierungen

Durch die Addition der Genome der diploiden Spezies *B. rapa* und *B. oleracea* zu *B. napus* liegt die überwiegende Mehrzahl der Gene als ähnliche Kopien aus den beiden Genomen vor, d.h. sie sind zueinander homöolog. Für fast jeden Locus im A-Genom von Raps kann ein homöologer Locus im C-Genom identifiziert werden. Die meisten duplizierten Loci in jedem der diploiden Genome sind in kollinearen Blöcken angeordnet. Chromosomale Umordnungen aufgrund von Inversion, Translokation, Deletion, Insertion und Duplikation treten in der Folge von Fragmentaustauschen als Reparaturmechanismus von Doppelstrangbrüchen während der Meiose auf (*crossing over*). Sie führen zur Neukombination von Allelen, Chromosomensegmenten oder ganzen Chromosomen und fördern damit die genetische Variabilität. Durch vergleichende Kartierung lassen sich anhand der mehr oder weniger stark ausgeprägten Kollinearität einzelner Loci Aussagen über die genetische Distanz der untersuchten Spezies und damit ihre Kreuzbarkeit treffen. Bei *B. rapa*

und *B. oleracea* sind einige Umordnungen innerhalb der Genome identisch. Sie sind demnach bereits vor der Differenzierung der beiden Arten entstanden. Wenigstens 16 größere chromosomale Umordnungen dagegen unterscheiden die beiden diploiden Genome, da sie erst nach ihrer Aufspaltung von dem gemeinsamen Vorfahren auftraten (Parkin et al. 2003). Nach der Arthybridisierung und Fusion der beiden Genome zur Bildung von *B. napus* haben keine größeren Umordnungen stattgefunden, welche das A-Genom und das C-Genom von Raps signifikant von denen ihrer Vorfahren *B. rapa* und *B. oleracea* unterscheiden würden (Parkin et al. 2003). Trotz allem können aber Umordnungen zwischen dem A- und dem C-Genom im allotetraploiden Raps beobachtet werden (Osborn et al. 2003, Udall et al. 2005, Sharpe et al. 1995), insbesondere wenn es sich um neue Rapsformen, sog. „Resynthesen“ handelt (Parkin et al. 1995, Song 1995). Das wesentliche Ereignis zur Etablierung des neuen Genoms war dabei die erste Meiose nach der Addition der elterlichen Genome, wie Szadkowski et al. (2010) mit ihrer Untersuchung zur Entstehung der Genome von Rapsresynthesen aufklärten. Resynthesen sind natürliche oder künstlich herbeigeführte Hybridisierungen von *B. rapa* und *B. oleracea*. Bei den hierbei auftretenden Umordnungen handelt es sich um homöologe reziproke sowie homöologe nicht-reziproke Umstrukturierungen (Parkin et al. 1995, Sharpe et al. 1995). Unter homöologer reziproker Translokation versteht man den Austausch von Genomfragmenten zwischen homöologen Chromosomen aus dem A-Genom und dem C-Genom, also beispielsweise den Austausch eines A9-Fragmentes gegen ein C8-Fragment. Da die Chromosomen A9 und C8 homöolog sind, also genetisch weitgehend identisch sind, kann eine solche Umordnung im Idealfall ohne Auswirkung auf den Phänotyp erfolgen. Bei einer homöologen nicht-reziproken Translokation erfolgt der Transfer eines Genfragmentes von einem Chromosom auf sein Homöolog, also in dem gewählten Beispiel von A9 nach C8, ohne daß ein Austausch erfolgt. Damit verkleinert sich das Chromosom A9, während der homöologe Partner C8 größer wird. Leflon et al. (2010) beobachteten, daß im allotetraploiden Resyntheseraps die Frequenz von *crossing overs* gegenüber der Frequenz in diploiden Hybriden deutlich erhöht ist. Dies bedeutet auch, daß die genetische Neukombination beschleunigt ist und damit die Möglichkeit besteht, ungewünschte Kopplungen schneller zu brechen.

In diploiden Spezies können normalerweise keine homöologen Umordnungen auftreten, daher sind hier Methoden wie die Identifizierung von Kandidatengenen und die Entwicklung molekularer Marker zur Selektion deutlich einfacher und schneller durchzuführen als in allopolyploiden Pflanzen. Das Vorliegen von Syntänie, also der gleichen Anordnung von orthologen Loci (gleiche Genorte in verschiedenen Arten), ist dabei zwischen phylogenetisch eng verwandten Arten von zentraler Bedeutung und besonders hilfreich in der Entwicklung molekularer Marker, der Annotation von Kartenpositionen und der Identifikation von Genen. Syntänie erleichtert auch die Arbeit mit allopolyploiden Spezies

durch die Möglichkeit, auf diploide Referenzgenome zurückgreifen zu können, wie im Fall von Raps auf *B. rapa*, *B. oleracea* und *A. thaliana*. Selbstverständlich sind Referenzgenomsequenzen der gleichen Spezies das wertvollste Hilfsmittel für Kartierungsbemühungen. Das Referenzgenom für Raps (*B. napus*) steht ebenso wie das für *B. oleracea* kurz vor der Veröffentlichung (Stand April 2013). Damit können nachträglich auch die in dieser Arbeit gewonnenen Ergebnisse und Erkenntnisse validiert werden.

Molekulargenetik als Werkzeugkasten der Züchtung: Markergestützte Selektion und Next Generation Sequencing

Mit den Fortschritten in der DNA-Sequenzierung entwickelte sich die Molekularbiologie in den letzten Jahren rasant weiter. So ermöglichen Sequenzierungstechnologien der zweiten und dritten Generation (sogenannte „second“ und „third generation sequencing“ Technologien) die Sequenzierung größerer eukaryotischer Genome, die Massenentwicklung von molekularen Markern, die Identifizierung agronomisch wichtiger Gene und die Nutzung dieser Kenntnisse in Genomics-basierten Zuchtprogrammen. Referenzgenome werden für immer mehr Nutzpflanzen zugänglich und erlangen damit große Bedeutung. Mit Sequenzinformationen lassen sich agronomisch wichtige Gene schnell identifizieren (Edwards und Batley 2010). Im Falle von Raps steht derzeit das Referenzgenom von *B. rapa* (The Brassica rapa Genome Sequencing Project Consortium 2011) zur Verfügung.

Unter „second generation sequencing“ oder „next-generation sequencing“ versteht man die Gewinnung von Sequenzdaten im großen Umfang durch Hochdurchsatzverfahren. Dabei werden große Mengen, d.h. mehrere Millionen kurzer DNA-Fragmente mit Längen von 25 bis 400 Basenpaaren (bp) sequenziert. Anbieter dieser Verfahren sind Roche (Basel, Schweiz) mit dem Roche 454 GS FLX Titanium System (<http://454.com/products/gs-flx-system/>), das Fragmentlängen von 350 bis 500bp generiert und über 400Mbp mit einer Genauigkeit von >99,5% (single read accuracy) produzieren kann, sowie Life Technologies (Carlsbad, California, USA) mit dem Ion Torrent™ Semiconductor Sequencing System (Rothberg et al. 2011, <http://www.iontorrent.com/>) und Illumina (San Diego, California, USA) mit dem HiSeq2500 und MiSeq System (<http://www.illumina.com/systems.ilmn>). Die Sequenzier-technologie hat sich bis heute weiterentwickelt und grundlegend verändert. Die sogenannte „third generation sequencing“-Technologie kommt ohne Amplifizierung der DNA-Fragmente aus und vermeidet damit Fehler infolge fehlerhafter Polymerasekettenreaktion (PCR). Anbieter sind derzeit Helicos Biosciences (Cambridge, MA, USA) mit einer Einzelmolekülsequenzierungstechnologie tSMS (<http://www.helicosbio.com/Technology/tabid/64/>

Default.aspx), sowie Pacific Biosciences (Cambridge, MA, USA) mit einem Einzelmolekül-Realtime-Sequenzierungssystem PacBioRS (Carneiro et al. 2012, <http://www.pacificbiosciences.com/products/pacbio-rs-system/>).

Die Vorteile der Genomsequenzierung gegenüber der aufwendigen Entwicklung einzelner molekularer Marker zur Unterstützung der Züchtung liegen auf der Hand. Der finanzielle Aufwand der Markerentwicklung ist zwar nach wie vor hoch, jedoch ist heute der Informationsgewinn ungleich höher, da in einem Lauf große Sets an informativen Sequenzmarkern entwickelt werden können. Damit lassen sich sehr vielfältige Fragestellungen beantworten. Markertechnologien wie die aufwendigen und teuren RFLP-Marker oder Genmarker wie SCAR- und CAPS-Marker erlauben eine Selektion auf bestimmte Allele, sind aber nur bei Merkmalen sinnvoll, die von einem oder wenigen Genen vererbt werden. Selektion auf komplex vererbte Merkmale oder sogar Merkmalskombinationen kann mit diesen Methoden schlecht im Routineablauf erfolgen. SNP-Arrays bieten hier eine elegante Alternative. Auch für Anwendungen wie Kopienzahlbestimmung, die Ermittlung der genetischen Diversität innerhalb von Genpools und die Aufklärung komplex vererbter Merkmale sind umfassende Sequenzinformationen weitaus zielführender als traditionelle Methoden.

Perspektivisch läßt sich analog zur Tierzucht mithilfe umfangreicher genotypischer Daten wie SNPs der geschätzte genomische Zuchtwert (*genomic estimated breeding value, GEBV*), berechnen, so daß aufwendige und dadurch oftmals limitierende Phänotypisierungen vermieden werden können (Snowdon und Luy 2012). Auch können so Diversitätsstudien mit dem Ziel der Erweiterung der genetischen Diversität im Zuchtmaterial durchgeführt werden. Züchtung mithilfe von Sequenzierung vereinfacht durch die Nutzung von genomweiten Markern die gezielte Integration neuen Materials in den Genpool, ohne daß gleichzeitig auf die Zuchtfortschritte der Vergangenheit verzichten werden muß (Snowdon und Luy 2012). Eine große Herausforderung an Züchter ist derweil die Auswertung und Zuordnung der enormen Datenmengen, welche mit neuen Sequenzierungstechnologien erzeugt werden. Insbesondere in komplexen Genomen wie *B. napus* sind hierfür Referenzsequenzen essentiell, um zwischen Sequenzierungsfehlern, artspezifischen SNPs, aber auch paralogen, homologen und homöologen Varianzen unterscheiden zu können. Aber erst die Kombination des Wissens aus Sequenzdaten und zuverlässigen phänotypischen Daten zu agronomisch relevanten Merkmalen ermöglicht eine umfassende Auswertung und Nutzung der Ergebnisse für die Züchtung.

Kandidatengenansatz: Ligninsynthesegene *Cinnamylalkoholdehydrogenase* (CAD) und *Cinnamoyl-CoA-Reduktase* (CCR)

Enzyme des Ligninbiosyntheseweges waren bereits in Modellpflanzen Kandidaten für Untersuchungen zur Ligninzusammensetzung und zu Modifizierungsmöglichkeiten. Einige Beispiele sind O-Methyltransferase (OMT) in Tabak (Atanassova et al. 1995) und Laccase in *A. thaliana* (Liang et al. 2006), aber auch Zellulosesynthase 3 (CESA3) in *A. thaliana* (Caño-Delgado et al. 2003). Daneben liegen viele Untersuchungen zu Cinnamylalkoholdehydrogenase (CAD) vor (z. B. Halpin et al. 1994, Sibout et al. 2003 und 2005, Kim et al. 2004, Goujon et al. 2003b). Viele Studien beschäftigten sich mit Cinnamoyl-CoA-Reduktase (CCR) (z. B. Goujon et al. 2003a, Patten et al. 2005). Ralph et al. (2001) untersuchten CAD und Caffeinsäure-O-Methyltransferase (COMT) in Pappel, Tabak und Kiefer. Die Ligninsynthese ist ein komplexes Netzwerk. So stellten Chabannes et al. (2001) in Kreuzungsexperimenten mit transgenen Tabakpflanzen fest, daß CCR und CAD offenbar synergistisch den Ligningehalt beeinflussen. In früheren Studien an Tabak führte die Stilllegung von CCR zu einem niedrigeren Ligningehalt, während CAD-Mutanten eine veränderte Ligninzusammensetzung aufwiesen. Doppelmutanten wiesen einen stark reduzierten Ligninanteil mit veränderter Extrahierbarkeit auf, jedoch keine Beeinträchtigung der Pflanzenentwicklung. Solche Mutanten sind in der Züchtung sehr wertvoll, da die wünschenswerte Reduzierung antinutritiven Lignins keinen Verlust in der generellen Fitness der Pflanze nach sich zieht. Viele dieser Studien beziehen sich jedoch auf vaskuläres Lignin. Die Bedeutung jener Enzyme in der Synthese von Monolignolen und Lignin in der Samenschale ist bislang wenig untersucht worden. Diese Doktorarbeit soll einen Beitrag zur Aufklärung dieser Zusammenhänge leisten. Die besondere Bedeutung der Arbeit liegt darin, daß die Untersuchungen direkt an der Kulturpflanze Raps anstelle wie bisher in einer Modellpflanze vorgenommen wurden. Die Ergebnisse haben daher einen unmittelbar nutzbaren Wert für potentielle Züchtungsansätze mit dem Ziel der Reduzierung des Faseranteils im Rapsschrot.

Zielsetzung der Arbeit

Übergeordnetes Ziel der Arbeit war es, Kandidatengene mit Einfluß auf die Qualitätsparameter Samenfarbe und Samenfasergehalt in Raps zu identifizieren. Die molekulargenetischen Arbeiten wurden an doppelhaploiden (DH) Populationen durchgeführt. Markerdaten und phänotypische Daten standen teilweise zur Verfügung und wurden im Laufe der Arbeit um neue Jahrgangsdatensätze ergänzt. Die Populationen gehen auf Eltern zurück, die sich in den Parametern Samenfarbe und Samenfasergehalt unterscheiden. Sie kombinieren die Eigenschaften Gelb- bzw. Schwarzsamigkeit mit hohem, mittlerem bzw. niedrigem Fasergehalt.

Informationen über Kandidatengene für diese Zielmerkmale können in der Züchtung neuer Rapslinien genutzt werden, um zusätzlich zu sonstigen Saatgutqualitätsmerkmalen auch auf einen verbesserten Futterwert des Schrotes zu selektieren. Dies würde eine verstärkte Nutzung des Rapsschrotes in der Tierernährung ermöglichen und damit einen beträchtlichen Mehrwert aus dem Nebenprodukt der Ölgewinnung erzielen. Zielstellung ist also letztlich die Reduzierung des unverdaulichen Fasergehaltes, der mit rund 10% im Schrot eine bedeutende Fraktion darstellt und darüber hinaus negativ mit dem Anteil des wertvollen Proteins korreliert ist.

Die vorliegende Arbeit behandelt dieses Thema einerseits durch genetische Analyse spaltender Populationen und die Kartierung der quantitativ vererbten Merkmale durch die Berechnung von *quantitative trait loci* (QTL). Auf dieser Basis sollten anschließend Kandidatengene identifiziert werden und durch Sequenzierung in den Kartierungseltern charakterisiert werden. Sequenzunterschiede in den kodierenden Bereichen der Gene sollten dann in Hinblick auf allelische Diversität und Kopienzahl analysiert werden, um spezifische Genmarker für die Rapszüchtung zu entwickeln.

Publikationen

1. A knockout mutation in the lignin biosynthesis gene *CCR1* explains a major QTL for acid detergent lignin content in *Brassica napus* seeds

Liu, L. Z., A. Stein, B. Wittkop, P. Sarvari, J. Li, X. Yan, F. Dreyer, M. Frauen, W. Friedt, R.J. Snowdon, Theoretical and Applied Genetics 2012, 124:1573-1586

2. Dissection of a major QTL for seed colour and fibre content in *Brassica napus* reveals colocalisation with candidate genes for phenylpropanoid biosynthesis and flavonoid deposition

A. Stein, B. Wittkop, L.-Z. Liu, C. Obermeier, W. Friedt, R. J. Snowdon, Plant Breeding 2013, accepted and in production

A knockout mutation in the lignin biosynthesis gene *CCR1* explains a major QTL for acid detergent lignin content in *Brassica napus* seeds

Liezhao Liu · Anna Stein · Benjamin Wittkop ·
Pouya Sarvari · Jiana Li · Xingying Yan · Felix Dreyer ·
Martin Frauen · Wolfgang Friedt · Rod J. Snowdon

Received: 14 August 2011 / Accepted: 31 January 2012
© Springer-Verlag 2012

Abstract Seed coat phenolic compounds represent important antinutritive fibre components that cause a considerable reduction in value of seed meals from oilseed rape (*Brassica napus*). The nutritionally most important fibre compound is acid detergent lignin (ADL), to which a significant contribution is made by phenylpropanoid-derived lignin precursors. In this study, we used bulked-segregant analysis in a population of recombinant inbred lines (RILs) from a cross of the Chinese oilseed rape lines GH06 (yellow seed, low ADL) and P174 (black seed, high ADL) to identify markers with tight linkage to a major quantitative trait locus (QTL) for seed ADL content. Fine mapping of the QTL was performed in a backcross population comprising 872 BC₁F₂ plants from a cross of an F₇ RIL from the above-mentioned population, which was

heterozygous for this major QTL and P174. A 3:1 phenotypic segregation for seed ADL content indicated that a single, dominant, major locus causes a substantial reduction in ADL. This locus was successively narrowed to 0.75 cM using *in silico* markers derived from a homologous *Brassica rapa* sequence contig spanning the QTL. Subsequently, we located a *B. rapa* orthologue of the key lignin biosynthesis gene *CINNAMOYL CO-A REDUCTASE 1* (*CCR1*) only 600 kbp (0.75 cM) upstream of the nearest linked marker. Sequencing of PCR amplicons, covering the full-length coding sequences of *Bna.CCR1* homologues, revealed a locus in P174 whose sequence corresponds to the *Brassica oleracea* wild-type allele from chromosome C8. In GH06, however, this allele is replaced by a homologue derived from chromosome A9 that contains a loss-of-function frameshift mutation in exon 1. Genetic and physical map data infer that this loss-of-function allele has replaced a functional *Bna.CCR1* locus on chromosome C8 in GH06 by homoeologous non-reciprocal translocation.

Communicated by R. Visser.

L. Liu and A. Stein contributed equally to the work.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-012-1811-0) contains supplementary material, which is available to authorized users.

L. Liu · A. Stein · B. Wittkop · P. Sarvari · W. Friedt ·
R. J. Snowdon (✉)
Department of Plant Breeding, IFZ Research Centre for
Biosystems, Land Use and Nutrition, Justus Liebig University,
Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany
e-mail: rod.snowdon@agr.uni-giessen.de

L. Liu · J. Li · X. Yan
College of Agronomy and Biotechnology, Southwest University,
Beibei, Chongqing 400716, China

F. Dreyer · M. Frauen
Norddeutsche Pflanzenzucht Hans-Georg Lembke KG,
24363 Hohenlieth, Germany

Introduction

Oilseed rape/canola (*Brassica napus* L., genome AACC) is widely grown throughout the world for vegetable oil and biodiesel production, and after oil extraction it also provides a high-quality meal used primarily for livestock feeding (Downey and Rakow 1987). In contrast to other *Brassica* species, which show a range of seed colour from bright yellow to black, *B. napus* seeds normally have a dark pigmentation. Nevertheless, breeding for yellow seed colour is considered advantageous for the meal quality because yellow-seeded genotypes have a thinner seed coat associated with higher protein content (Wittkop et al. 2009)

and reduced quantities of non-energetic fibre and antinutritional polyphenolics (Simbaya et al. 1995). Hence, substantial efforts have been made to introgress genes imparting yellow seed colour into *B. napus* via interspecific hybridisation (e.g. Shirzadegan 1986; Van Deynze and Pauls 1994; Rashid et al. 1994; Tang et al. 1997; Li et al. 1998, 2011).

Seed colour is difficult to use in breeding programmes as a morphological marker for improved meal quality, however, because seed coat pigment biosynthesis is complex and highly sensitive to temperature, light intensity and other abiotic factors (Boesewinkel and Bouman 1995). Genetic studies of seed colour in different genetic backgrounds often revealed different major-effect quantitative trait loci (QTL; Tang et al. 1997; Li et al. 1998; Badani et al. 2006; Liu et al. 2005; Xiao et al. 2007; Fu et al. 2007), along with allelic variation among black-seeded crossing partners at major seed colour loci and other contributing QTL (Badani et al. 2006; Fu et al. 2007; Snowdon et al. 2010). This complexity of inheritance and the environmental instability of seed pigmentation complicate introgression of the yellow-seed trait into elite breeding lines. Early reports described the dominance of the black seed trait over yellow seed colour in *B. napus*, and suggested control by three independent gene loci (Shirzadegan 1986; Van Deynze et al. 1993). However, further analysis of other yellow-seed sources revealed that different major loci can be involved and that semi-dominant or dominant inheritance of yellow seed colour over black seed colour is also possible (Tang et al. 1997; Li et al. 1998).

Biosynthesis of lignin and its phenylpropanoid precursors in xylem and stem tissues has been studied extensively in numerous model and crop plants (see Bonawitz and Chapple 2010 for a recent review). In contrast, seed coat phenylpropanoid biosynthesis is less well understood, although the seeds from many crops play a huge role both in livestock feeding and in human nutrition. Whereas insufficient dietary fibre is commonly associated with increased obesity in human populations, high quantities of indigestible, non-energetic fibre are considered undesirable in livestock feeds. On the other hand, seed coat compounds have obvious roles in protection of the embryo against damage by biotic and abiotic stress factors (reviewed by Boesewinkel and Bouman 1995).

Marles and Gruber (2004) found that most of the lignin-related compounds in seeds of the closely related *Brassica carinata* are found in the testa. In accordance with this result, experiments with *B. napus* seeds showed that de-hulling achieves considerably reduced crude fibre and condensed tannin contents in the extraction meal from different oilseed rape varieties (Matthäus 1998). We have observed that near-infrared reflectance spectroscopy (NIRS) measurements of acid detergent lignin (ADL) in

B. napus seeds correlate strongly to concentrations of phenylpropanoid-derived lignin precursors measured by mass spectrometry analysis of seed coat preparations (A. Frolov and B. Wittkop, unpublished results). Hence, our NIRS estimates of ADL from whole, intact *B. napus* seeds can be regarded as quantitative measurements of lignin-related metabolites in the seed coat. By NIRS analysis of genetically diverse *B. napus* accessions, we found considerable phenotypic variation for ADL content, and a considerably greater heritability of ADL content than of seed colour (Snowdon et al. 2010). Most importantly, genotypes with very similar seed colour can show large variation in seed ADL, implying that seed colour alone is not always suitable as a selection marker for meal digestibility.

In different studies, the two homoeologous *B. napus* chromosomes A9 and C8 have been reported to contain a major QTL with a large effect on seed colour in different genetic backgrounds (Li et al. 1998; Badani et al. 2006; Fu et al. 2007). Recently, the major QTL on chromosome A9 was found to coincide with a corresponding major QTL for seed ADL content (Snowdon et al. 2010). Previous studies of seed colour and antinutritive phenolics in *Brassica* oilseeds focused on *Arabidopsis transparent testa (tt)* or *tannin-deficient seed (tds)* mutations as potential candidate gene targets for improved meal quality (e.g. Marles et al. 2003; Badani et al. 2006; Burdzinski and Wendell 2007; Wei et al. 2007; Xu et al. 2007; Akhov et al. 2009; Auger et al. 2009; Chai et al. 2009). Most *TT* and *TDS* genes are involved in flavonoid biosynthesis. In contrast, little attention has been paid to the phenylpropanoid biosynthesis pathway. In this study, we exploited genomic DNA sequence data from the *Brassica* A-genome of *Brassica rapa* for sequence-based fine mapping of a major QTL for seed ADL content in *B. napus*. By *in silico* chromosome walking across the QTL, we identified a key phenylpropanoid biosynthesis pathway gene within the QTL. DNA sequence analysis uncovered functional mutations that appear to explain the low-ADL phenotype.

Materials and methods

Mapping populations

A population of 232 F₉ recombinant inbred lines (RILs) was derived by single seed descent (SSD) from F₂ offspring of a cross between the Chinese semi-winter oilseed rape parental lines, GH06 (yellow seeds, low ADL) and P174 (black seeds, high ADL). The pedigree of GH06 includes a yellow-seeded genotype derived from interspecific hybridisation between *B. napus* and *Brassica juncea*, and another yellow-seeded *B. napus* genotype developed by radiation mutagenesis of a dark-seeded breeding line.

QTL mapping of seed colour and fibre traits was performed using data collected from field trials of the RIL population in four different environments over 2 years as follows: 2008 in Giessen (central Germany), Hohenlieth (northern Germany) and Chongqing (southwest China), and 2009 again in Giessen. For the trials in Germany in 2008, the population was sown in the greenhouse in winter and transplanted to the field in early spring, after vernalisation, when plants reached the four-leaf stage. In Chongqing in 2008, the seeds were sown in nursery beds in autumn and transplanted to the field 1 month later. In Giessen in 2009, the population was sown directly in the field in spring. All trials were performed with a plot size of 4.5 m² (1.5 m × 3 m) with five or six rows per plot depending on the location.

Based on the results of genetic mapping and QTL analysis, a backcross population for subsequent fine mapping was derived from a cross of an F₇ plant of the yellow-seeded, low-ADL inbred line RIL138, derived from the cross GH06*P174, with the black-seeded, high-ADL parent P174. Phenotypic analysis of F₂–F₃ progenies from the selected RIL138 plant revealed that it was heterozygous for the major QTL controlling ADL content and seed colour (see results), although all markers surrounding this QTL in RIL138 were derived from the black-seeded parent P174. Hence, we expected the resulting backcross to have a high resolution for fine mapping. A heterozygous, high-ADL BC₁ plant from RIL138 × P174 was selected for selfing to produce BC₁F₂ seeds. A total of 872 BC₁F₂ plants were sown in the greenhouse and transplanted to the field in Giessen in 2010. We collected open-pollinated seed samples from all individual BC₁F₂ plants to avoid micro-environmental temperature and humidity effects from self-pollination bags. NIRS analysis of ADL content was performed in two technical replicates.

DNA marker screening

Genomic DNA from the RILs and parental lines was extracted from young leaf tissue using the method described by Doyle and Doyle (1990). Genomic DNA samples from the BC₁F₂ population were extracted using a BioSprint 96 magnetic bead extraction system (Qiagen, Hilden, Germany) with corresponding BioSprint 96 DNA Plant Kits, in accordance to the manufacturer's instructions. For bulked-segregant analysis, equimolar DNA samples from ten low-ADL and ten high-ADL RILs, respectively, were pooled into low- and high-ADL bulks. The two bulked samples and the two parents were used to generate amplified fragment length polymorphism (AFLP; Vos et al. 1995) and simple sequence repeat (SSR) fingerprints to identify markers tightly linked to QTL for ADL. AFLP analysis was carried out using *EcoRI/MseI* AFLP kits from Invitrogen (Carlsbad, CA, USA), in

accordance with the manufacturer's instructions. A total of 1,024 selective AFLP primer combinations were applied with fluorescently labelled selective primers in the two bulks and parents, with dual IRD 700/800 fluorescent fragment detection on a LI-COR 4200 DNA Analyzer. Detection of SSR marker fragments was performed with a fluorescently labelled universal M13 primer, as described by Berg and Olaisen (1994). *Brassica* SSR primer pairs were obtained from publicly available collections listed at <http://www.brassica.info/> and by *in silico* SSR detection from *B. rapa* sequences available at the *Brassica* Genome Gateway (<http://brassica.bbsrc.ac.uk/>) and the BRAD *Brassica* Database (<http://brassicadb.org/>), respectively. Primer sequences for the informative SSR markers are provided as electronic supplementary material (ESM 1).

Bulked-segregant analysis and fine mapping

Two AFLP markers tightly linked with seed ADL content were selected for cloning and sequencing after re-amplification with unlabelled AFLP primers. The target fragments were extracted from silver-stained polyacrylamide gels by boiling in 50 µL of sterile double-distilled water. The corresponding AFLP selective primers and 5 µL of the supernatant as template were used to re-amplify the recovered DNA fragment. Cloned PCR products containing inserts of the expected size were sequenced, and sequences were BLASTed to the *B. rapa* sequences available at BRGB, South Korea (<http://www.brassica-rapa.org/BRGP/index.jsp>) and BRAD, China (<http://brassicadb.org/brad/>) to identify corresponding *Brassica* genomic sequences. These were subsequently used to mine SSR markers in close proximity to the sequences of the trait-linked AFLP markers, using the *in silico* SSR finder software WebSat (Martins et al. 2009).

Codominant *Brassica* SSR markers associated with ADL content in the RIL population were re-mapped in the BC₁F₂ population, and sequences of all markers flanking the major QTL for ADL content were identified in corresponding sequence scaffolds from *B. rapa* chromosome A9 for comparison of the genetic and physical maps across the QTL region. An *in silico* chromosome walking approach was used to narrow the QTL confidence interval, using the physical sequence of *B. rapa* chromosome A9 to successively select *in silico* SSR markers with reduced recombination frequencies to the causal locus. For fine mapping, backcross progenies were screened for recombinations of marker alleles from the low-ADL parent in recessive, high-ADL plants.

Trait analysis

Measurements for seed colour and fibre components were obtained using an NIR System 6500 with WinISI II

software (FOSS GmbH, Rellingen, Germany). Phenotype values for seed colour (visual light absorbance), ADL (% seed dry weight), acid detergent fibre content (ADF, % seed dry weight) and neutral detergent fibre content (NDF, % seed dry weight) were extrapolated from NIRS spectra using calibrations developed specifically for the measurement of these traits in *B. napus* (Wittkop et al. 2009). NIRS-derived estimates for each trait and genotype were averaged over two technical repetitions. In the RIL population, mean trait values from up to three self-pollinated plants from each genotype over all four environments were used as the basis for the subsequent QTL analysis of seed colour and fibre components, while single-plant analyses were performed in the BC₁F₂ population. Cellulose concentrations were calculated as the difference between NDF (predominantly phenylpropanoid compounds, cellulose and hemicellulose) and ADF (predominantly phenylpropanoid compounds and hemicellulose), and hemicellulose concentration as the difference between ADF and ADL (phenylpropanoids). Statistical analysis of the phenotype data was performed using the software package SPSS 13.0.

Linkage analysis and QTL mapping

Genetic linkage analysis was performed using JoinMap 3.0 (Van Ooijen and Voorrips 2001), with the mapping function of Kosambi (1994) and a minimum logarithm of odds (LOD) score of 3.0. Detection of QTL for seed colour, cellulose, hemicellulose and ADL was performed in the RIL population by composite interval mapping using WinQTLcart2.5 (Wang et al. 2006; see <http://statgen.ncsu.edu/qtlcart/WQTLcart.htm>). The LOD threshold for the detection of significant QTL was set by permutation analysis with 300 permutations.

454 amplicon sequencing and structural analysis of *Bna.CCR1* homologues

A 454 amplicon sequencing approach was employed to survey DNA sequence variation over homologous *B. napus* copies of the phenylpropanoid biosynthesis candidate gene *CCR1* in P174 and GH06, respectively. Using genomic *CCR1* sequences from *B. rapa* (*BraA.CCR1*) and *Brassica oleracea* (*BolC.CCR1*), which each consist of five exons, we designed barcoded fusion PCR primers to generate amplicons of ~450–500 bp in length, from both A and C genome homologues, covering exons 1 and 2 in a single amplicon and exons 3, 4 and 5 with one amplicon each. Details of the tailed fusion primer sequences are provided in the electronic supplementary material (ESM 2).

After quantification, normalisation and equimolar pooling, the barcoded amplicons were sequenced in both directions by SeqIT (Kaiserslautern, Germany), as part of a larger

pooled sequencing reaction on a 1/8th sequencing plate of a 454-FLX genome sequencer (454 Life Sciences, Branford, CT, USA). The resulting sequence reads were clipped to remove emulsion PCR primer sequences, 4 bp key sequences and (after sorting into genotype-specific pools) MID barcodes. The sequences were then processed with the Sequencher[®] version 4.10 sequence analysis software (Gene Codes Corporation, Ann Arbor, MI, USA) as follows:

For each genotype, the sequences were individually grouped into amplicon-specific pools, according to their corresponding forward/reverse primer sequences and trimmed to remove M13/T7 overhang sequences. Subsequently, the raw reads were assembled into locus-specific contigs using a preferential 3' gap alignment with a stringent 95% minimum match and a minimum overlap of 20 bp. The chromosomal origin of each locus-specific contig was identified by aligning to *B. rapa* and *B. oleracea* chromosomes via the online BLAST available at <http://brassicadb.org/brad/>. For structural analysis, the sequences were trimmed to exons corresponding to the coding DNA sequence for *BraA.CCR1* and subsequently translated to the protein sequence. Positions of conserved functional domains were derived by protein BLAST analysis with translated amino acid sequences from full-length plant *CCR1* coding sequences available in the NCBI database.

Results

Inheritance of seed coat ADL content

Considerable variation was found in the expression of seed coat pigmentation phenotypes of individual RILs between German and Chinese environments, with genotypes that produced bright yellow seeds in China having considerably darker seeds at the two locations in Germany. In contrast, the fibre traits showed much less genotype-by-environment (G×E) interaction, reflecting the higher heritability of seed fibre components compared to seed pigmentation.

Correlations among seed traits in the RIL and BC₁F₂ populations are shown in Table 1. Although the seed colour showed high variation between the Chinese and German environments, the seed ADL content was still significantly positively correlated with seed colour in both the RIL population ($R^2 = 0.67$, $P < 0.01$) and the BC₁F₂ population ($R^2 = 0.82$, $P < 0.01$). The correlations of ADL to cellulose and hemicellulose content were also significant ($P < 0.01$), though considerably lower in both populations.

Segregation patterns for seed colour and fibre traits in the RIL and BC₁F₂ populations are shown in Fig. 1. Cellulose and hemicellulose, which are primarily embryonic traits and thus independent of seed colour, showed continuous normal distributions. For both of these traits,

Table 1 Correlations between seed components and seed colour in a RIL population (above diagonal) and a BC₁F₂ population (below diagonal) derived from the cross GH06 (yellow seeds, low seed ADL) × P174 (black seeds, high seed ADL)

Seed fractions	Seed ADL (% DW)	Seed colour	Oil (% DW)	Protein (% DW)	Cellulose (% DW)	Hemicellulose (% DW)
Seed ADL (% DW)	–	0.67**	0.06	–0.30**	0.14*	–0.23**
Seed coat colour	0.82**	–	0.20**	–0.29**	0.17**	–0.03
Oil (% DW)	–0.04	–0.18**	–	–0.50**	–0.19**	–0.82**
Protein (% DW)	–0.09	–0.05	–0.58**	–	–0.34**	0.42**
Cellulose (% DW)	0.25**	0.11**	0.05	–0.31**	–	–0.22**
Hemicellulose (% DW)	–0.25**	–0.13**	–0.63**	0.56**	–0.60**	–

* Significant correlation at $P < 0.05$ (two-tailed)** Significant correlation at $P < 0.01$ (two-tailed)

very little difference was observed between the parental lines and strong transgressive segregation was evident. The phenotypic distributions for seed colour also suggested a polygenic inheritance with environmental influence, although in both populations the distribution deviated significantly from normality. In contrast, the clear bimodal segregation for ADL content in the BC₁F₂ corresponds to a simple Mendelian segregation with dominant inheritance of a single major locus conferring reduced ADL content ($\chi^2_{(3:1)} = 0.523 < \chi^2_{0.05}$). The low-ADL phenotype of heterozygous F₁ seeds from the cross between P174 and GH06 also corresponds to a dominant inheritance of low ADL in this cross. On the other hand, however, full dominance is not exhibited over all high-ADL *B. napus* parents (data not shown). A strong maternal effect on seed ADL, as expected for a trait primarily expressed in the testa, was not observed in reciprocal crosses between GH06 and P174.

Bulked-segregant analysis

Of the 1,024 AFLP primer combinations tested in the bulked-segregant analysis, 202 exhibited polymorphic amplification products between bulked DNA samples from ten high-ADL RILs and ten low-ADL RILs, respectively. Testing of these 202 selected primer combinations in ten selected low-ADL and ten high-ADL RILs identified only two polymorphisms associated to the seed ADL phenotype. The AFLP primer combination E46/M35 generated a 362 bp dominant marker specific to lines with high-ADL content, while primer combination E35/M43 amplified a codominant marker with a 395 bp allele in high-ADL lines and a 399 bp allele in low-ADL lines, respectively. Genotyping of these two AFLP markers in the entire RIL population revealed a genetic distance of 1.84 cM between them and confirmed their correlation to seed ADL content.

Identification and mapping of ADL-associated markers

BLAST alignments of the cloned AFLP sequences to the Chinese *B. rapa* sequence database BRAD (accessed in

January 2011) revealed 100% matches of both ADL-associated markers to *B. rapa* chromosome A9 scaffold 000074. The physical distance between the markers was a mere 173 kb, despite the relatively large genetic distance of 2.6 cM between them in the RIL population. Based on the sequence information from *B. rapa* scaffold 000074, we designed two SCAR markers corresponding to the respective AFLP markers, and screened 18 *in silico* SSR primer pairs from the neighbouring chromosome region. Five of the SSRs showed polymorphisms between the mapping parents and the ADL bulks, and were henceforth genotyped in the entire RIL population. In addition, 101 *Brassica* SSR primer pairs with previously mapped loci on *B. napus* chromosome A9 were also used for bulk analysis. Seven of these SSR markers, which closely matched the segregation pattern for high- and low-ADL content in the bulk lines, were subsequently used to genotype the RIL population.

Linkage mapping of the ADL-associated markers from chromosome A9 revealed that the identified markers segregated into two separate linkage groups (Fig. 2). The first of these linkage groups (designated RIL_A9-1) contained ten markers covering 19.7 cM, whereas the second linkage group (RIL_A9-2) had a length of 9.3 cM and contained the two AFLP-derived SCAR markers. Interestingly, RIL_A9-1 contains an 8.5 cM gap that corresponds closely to the length of RIL_A9-2, and alignment of the two groups to the *B. rapa* physical map of chromosome A9 (between 2.5 and 3.5 Mbp) revealed that the sequence of RIL_A9-2 indeed corresponds to the *B. rapa* sequence between the markers CN79_145 and KBrH095F22.3_230, which flank the gap on RIL_A9-1 (Fig. 2).

QTL analysis and fine mapping

QTL analysis by composite interval mapping in the linkage maps for RIL_A9-1 and RIL_A9-2 revealed highly significant QTL for ADL content and seed colour (LOD = 25.3 and 20.0, respectively). These QTL colocalise at the same position on RIL_A9-1 (Fig. 2),

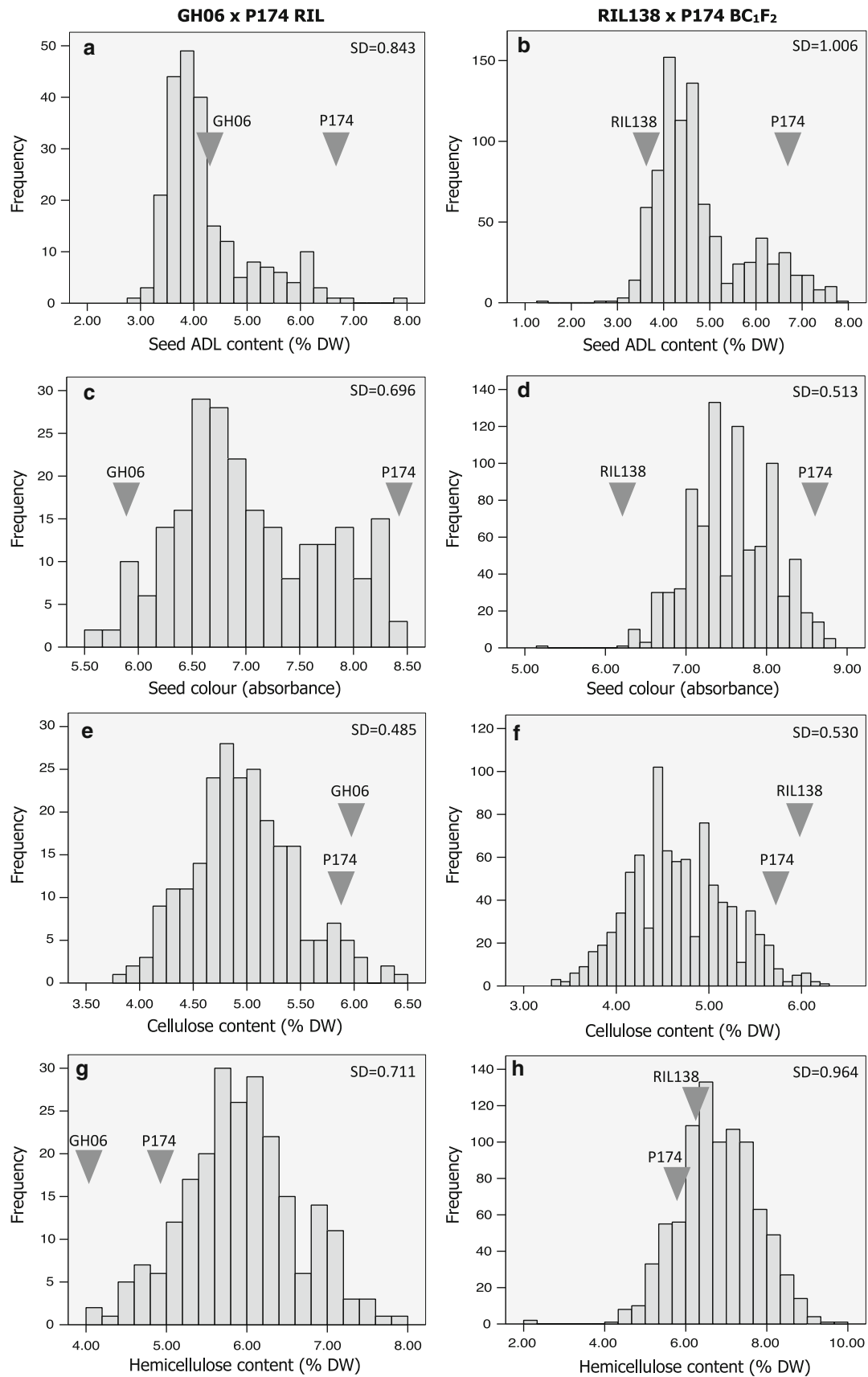


Fig. 1 Histograms showing phenotype distributions for **a, b** ADL content, **c, d** seed colour, **e, f** cellulose content and **g, h** hemicellulose content in seeds from **a** 232 RILs from the cross GH06 × P174, and **b** 872 BC₁F₂ plants from the backcross of RIL138 × P174. Arrows show the mean values for the parental lines GH06, P174 and RIL138, respectively. *SD* standard deviation

explaining 39.3 and 31.0%, respectively, of the phenotypic variation for ADL content and seed colour in the RIL population. Negative additive effects of -0.48 and -0.52 for seed colour and ADL content, respectively, indicate that the allele from parent P174 increases both seed coat pigmentation and ADL. The QTL peak spans the two SCAR markers developed from the bulked-segregant AFLP analysis.

For fine mapping, 200 BC₁F₂ plants with high seed ADL content were analysed with all polymorphic codominant markers spanning the QTL region on RIL_A9-2. The segregation pattern in the BC₁F₂ population revealed a single locus (hereinafter referred to as *ADL*) with dominance of the low-ADL allele. Consequently, codominant markers involved in recombination events with the *ADL* locus must be heterozygous in high-ADL BC₁F₂ individuals, whereas these in turn must be homozygous-recessive at *ADL*. Conversely, markers cosegregating with *ADL* must be homozygous for marker alleles from the high-ADL parent in all high-ADL BC₁F₂ plants. Pairwise genetic distances in relationship to the *ADL* locus could, thus, be calculated based on frequencies of such recombinations. This allowed reconstruction of an 8 cM long linkage group, designated BC₁F₂_A9-2, corresponding to RIL_A9-2 but

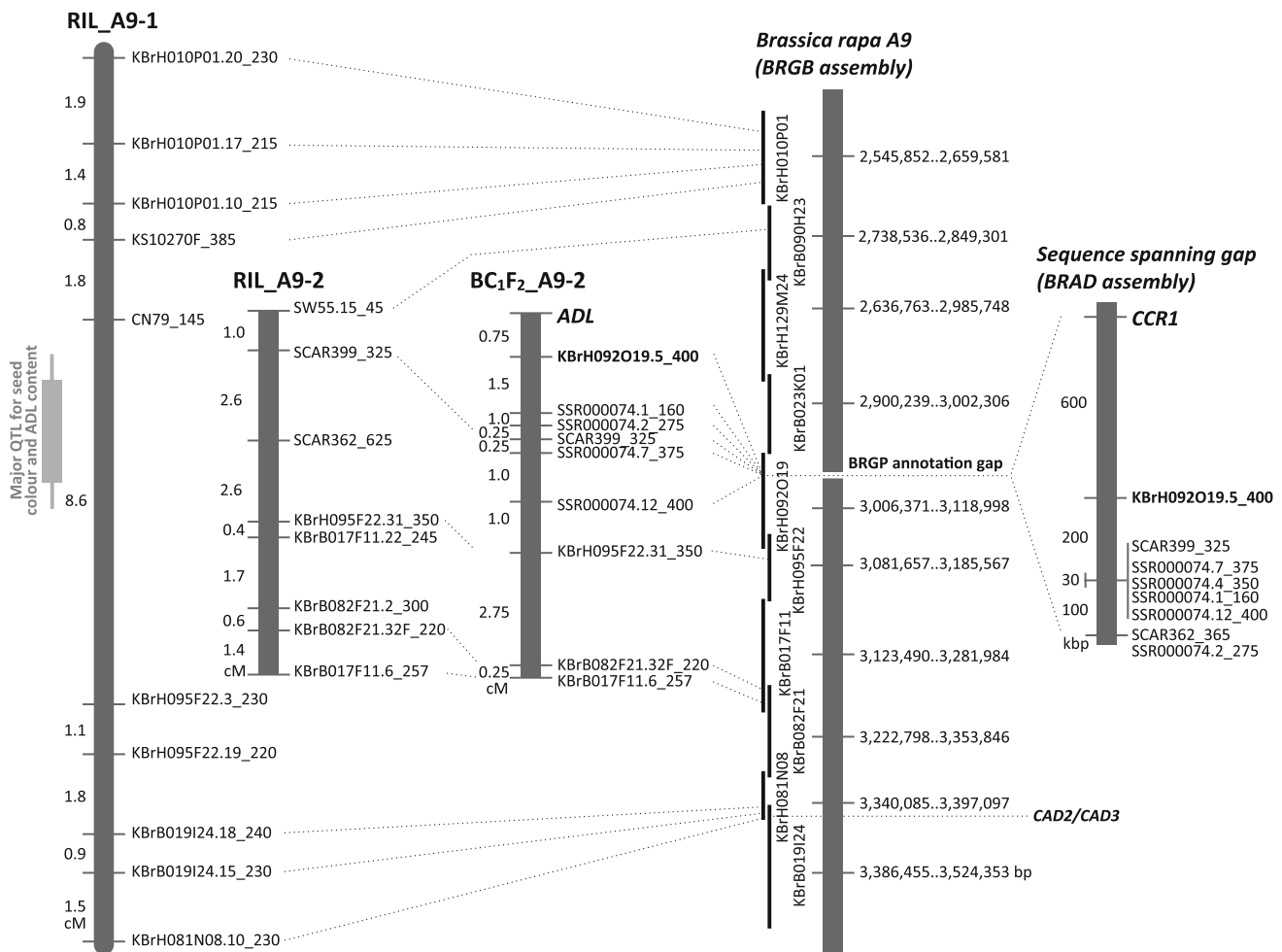


Fig. 2 Comparison of the genetic and physical maps for the region of *Brassica napus* chromosome A9 spanning the major QTL for seed ADL content and seed colour in the RIL population from GH06 × P174, and the corresponding *ADL* locus in a BC₁F₂ population derived from the heterozygous RIL138 backcrossed with P174. Alignment of marker sequences to the *B. rapa* chromosome A9 sequence assembly revealed that linkage group RIL_A9-2, which was

remapped with additional markers as BC₁F₂_A9-2 in the BC₁F₂ population, corresponds to the gap in linkage group RIL_A9_1. The lack of linkage between markers on these two groups is explained by a putative homoeologous non-reciprocal translocation of the chromosome block represented by RIL_A9-2 to chromosome C8 in the parental line GH06

including *ADL* as a Mendelian locus along with additional markers in closer proximity to *ADL*. The nearest SSR marker we were able to locate in BC₁F₂-A9-2 was KBrH090019.5, which mapped 0.75 cM below the *ADL* locus. Thirty-three further *in silico* SSR markers were tested from the 1 Mbp physical sequence interval between KBrH090019.5 and CN79_145, the most proximal marker on the other side of the putative translocation breakpoint on RIL_A9-1. Unexpectedly, all but 12 of these putative SSRs failed to amplify and the 12 that did amplify were all monomorphic between the parental lines. This meant that the *ADL* locus mapped at the proximal end of BC₁F₂-A9-2, a few centimorgan upstream from the position of the QTL for seed ADL content on RIL_A9-2. This discrepancy may result from less accurate mapping of RIL_A9-1 due to the lower population size and the presence of both dominant and codominant markers, whereas only codominant markers were used to map *ADL* on BC₁F₂-A9-2.

CINNAMOYL CO-A REDUCTASE 1 is a positional and functional candidate gene for the ADL locus

Based on the fine-mapping results, we expected the causative gene or genes for the *ADL* locus to be located in the vicinity upstream of the marker KBrH090019.5 on *B. napus* chromosome A9. Due to the high homology of the A-genome chromosomes between *B. rapa* and *B. napus*, and the strong conservation of marker order between our *B. napus* genetic map and the available *B. rapa* genome sequence, we queried the BRGP and BRAD databases to locate nearby candidate genes. No known *TT* genes were found in the vicinity of KBrH090019.5. On the other hand, only 600 kbp upstream of KBrH090019.5, we located a *B. rapa* orthologue of the major lignin biosynthesis gene, *CINNAMOYL CO-A REDUCTASE 1* (*BraA.CCR1*), in the BRAD gene annotation. This gene was, therefore, revealed for the first time as an important positional candidate for a

key role in the regulation of seed coat fibre content in *B. napus*. Unexpectedly, the sequence containing *BraA.CCR1* was not found in the BRGP assembly of the same region of *B. rapa* chromosome A9, which instead exhibited a gap of around 260 kbp at the corresponding position in comparison to the matching BRAD sequence scaffold (accessed in May 2011). In the BRAD annotation, the region in question also contained, besides *BraA.CCR1*, five codominant markers that mapped to the putative translocation represented by linkage group BC₁F₂-A9-2 and are at the same time associated with the major QTL for seed ADL content on chromosome A9.

Identification of functional mutations in *Bna.CCR1* homologues from GH06

Amplicon sequencing of the complete lengths of all five *Bna.CCR1* exons in GH06 and P174 revealed sequence contigs corresponding to *Brassica* homologues from chromosomes A9 of *B. rapa* and C8 of *B. oleracea*. Figure 3 compares the translated protein sequences derived from the coding regions of these homologues from GH06 and P174 with the corresponding *B. rapa* and *B. oleracea* reference sequences. In the black-seeded, high-ADL *B. napus* parent P174, we identified homologues corresponding to both the A9 locus from *B. rapa* (*BraA.CCR1.A9*) and the C8 locus from *B. oleracea* (*BolC.CCR1.C8*), whereas the latter locus was not found in GH06. Instead, GH06 contains a knockout allele that corresponds to the homologous *BraA.CCR1.A9* sequence but carries a loss-of-function frameshift (fs) mutation in exon 1. We propose that this *bnaA.ccr1.A9-fs* knockout is the causal dominant allele for the *ADL* locus in this cross.

The translated protein sequence of the chromosome C8 allele in P174 (*BnaC.CCR1.C8*) is identical to that from *B. oleracea* chromosome C8 (*BolC.CCR1.C8*). This differs to the predicted protein from *BraA.CCR1.A9* by a P > A

	Position	8	13		123	136		183		273
<i>BolC.CCR1.C8</i> (Bol009845)	A	V			V	S		L		I
<i>BnaC.CCR1.C8</i> (P174)	A	V			V	S		L		I
<i>BraA.CCR1.A9</i> (Bra026373)	P	V			V	N		L		L
<i>bnaA.ccr1.A9</i> (P174)	P	V			A	N		M		L
<i>bnaA.ccr1.A9</i> (GH06)	P	V			A	N		M		L
<i>bnaA.ccr1.A9-fs</i> (GH06)	P	frameshift/STOP								
Active site										
NAD(P) binding site										
Substrate binding site										

Fig. 3 Positions of amino acid alterations in the derived protein sequences of *BnaCCR1* homologues from chromosomes C8 and A9 of the dark-seeded, high-ADL *B. napus* line P174 and the yellow-seeded low-ADL mutant GH06. The protein sequences are aligned to the *BolC.CCR1* and *BraA.CCR1* reference sequences from *B.*

oleracea chromosome C8 (Bol009845) and *B. rapa* chromosome A9 (Bra026373), respectively. Boxes show putative functional mutations adjacent to enzyme active sites, NAD(P) binding sites and/or substrate binding sites, respectively

substitution at amino acid position 8, an N > S substitution at position 136 and an L > I substitution at position 273. While this last substitution occurs in a non-conserved terminal protein domain, particularly the proline–alanine substitution at position 8 is expected to significantly alter the binding properties of the protein near the first predicted NAD(P) binding site of the enzyme. The strong conservation of the protein sequences from the chromosome C8 allele from both *B. napus* and *B. oleracea* wild types suggests that *BnaC.CCR1.C8* encodes a functional CCR1 enzyme in P174.

The *B. napus* A-genome allele corresponding to *BraA.CCR1.A9* in both P174 and GH06 additionally carries two putatively functional amino acid substitutions: V > A at position 123 and L > M at position 183. Both of these changes are likely to further influence the adjacent substrate and NAD(P) binding sites, while the former mutation is also adjacent to the enzyme active site. We, therefore, suggest that the native A-genome allele found on chromosome A9 in both P174 and GH06 (hereinafter designated *bnaA.ccr1.A9*) is likely to encode a protein with impaired function and possibly an altered substrate specificity compared to the putative wild-type enzyme encoded by *BolC.CCR1.C8* and *BnaC.CCR1.C8*. Besides the *bnaA.ccr1.A9*-fs knockout GH06 possesses only this functionally impaired locus.

Discussion

Oilseed rape and canola extraction meals after oil extraction contain around 40% high-quality protein (dry weight after oil extraction) with a favourable composition of amino acids, including comparatively high contents of the essential sulphuric amino acids, methionine and cysteine. Modern oilseed rape varieties with “double-low” (00 or canola-quality) seeds, having no erucic acid in the seed oil and low concentrations of glucosinolates in the extraction meal, therefore represent a valuable feedstuff for livestock nutrition. On the other hand, the extraction meal contains a relatively large proportion of non-energetic crude fibre, particularly ADL, along with condensed tannins that reduce the protein digestibility. This strongly limits the quantities of oilseed rape meal that can be used in feed rations for monogastric livestock. Reductions in the seed ADL content would significantly increase the proportion that could be included in poultry and swine feed mixtures, considerably improving the overall economical value of the crop. Interestingly, low-ADL lines identified in this study tended to have an increased protein content, with no corresponding effect on oil content despite the strong negative correlation between oil and protein content. This coincides with the findings of Wittkop et al. (2009) in other *B. napus*

germplasm and contrasts the general belief that a reduction of seed coat thickness in light-seeded *B. napus* lines should result in a pleiotropic increase of both oil and protein content, due to the proportionally greater contribution of the embryo to the seed. In fact, the observed tendency for higher protein in low-ADL lines, with no corresponding increase in oil content, suggests a direct biochemical relationship between the seed coat and the biosynthesis of storage proteins in the embryo. Biosynthesis of the *B. napus* seed storage protein *napin* competes with the phenylpropanoid biosynthesis pathway for methionine. An accumulation of free methionine in the embryos of low-ADL genotypes (unpublished observations) might feasibly cause a feedback into increased storage proteins. The negative correlation between ADL and protein content has a dual positive effect on the nutritional composition of the seed meal for animal nutrition, by simultaneously increasing the energy content and the digestibility of the meal. Identification of genes directly influencing ADL content, independently of the flavonoid pathway, will enable considerably more targeted approaches for identification and use of genetic resources for this trait in practical breeding.

Inheritance of seed colour and ADL content

Strong correlations were observed between seed colour and ADL content, but not between ADL and cellulose or hemicellulose contents. These findings confer to our expectation that seed phenylpropanoids contributing to antinutritive ADL are largely confined to the testa, whereas overall seed contents of cellulose and hemicellulose should also contain a significant contribution from embryo cell walls. Oil content in the embryos was highly negatively correlated with hemicellulose content in the two populations. This is presumably due to competition for carbon partitioning between the two pathways, confirming that a substantial proportion of the hemicellulose is embryo derived. Interestingly, the lack of a clear maternal effect on the *ADL* locus suggests that the regulation of phenylpropanoid biosynthesis in the testa may be influenced by genes expressed in the seed embryo or endosperm.

A non-reciprocal translocation involving the *bnaA.ccr1.A9*-fs mutation putatively disrupts seed coat phenylpropanoid biosynthesis

We identified the important lignin biosynthesis gene *CCR1* as a positional candidate with very close physical vicinity to markers that are tightly linked to the major dominant locus for low seed coat ADL content in *B. napus*. Furthermore, we discovered that a functional *BnaC.CCR1* locus on chromosome C8 appears to have been replaced in

GH06 by a loss-of-function frameshift mutant allele. These findings implicate the *bnA.ccr1.A9_{fs}* knockout as the effector of the *ADL* locus.

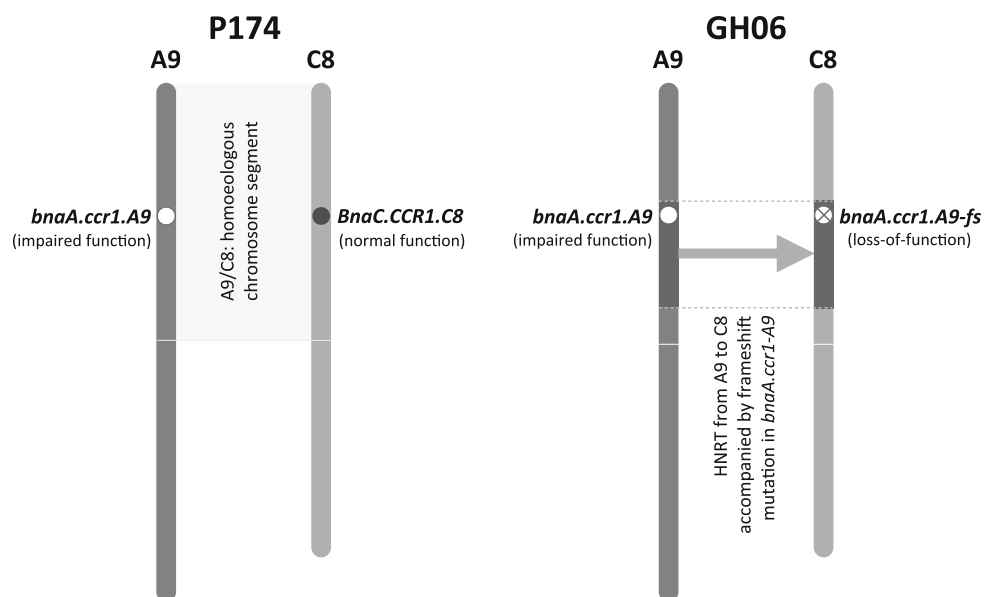
Unexpected discrepancies between our detailed genetic maps and two alternative physical maps, corresponding to the chromosome region containing the *ADL* locus in *B. rapa*, suggest that translocation as the likely mechanism for the replacement of *BnaC.CCR1.C8*. The linkage groups RIL_A9-1 and RIL_A9-2 showed no genetic linkage, although they represent physically adjacent sequences on chromosome A9 (Fig. 2). These observations are consistent with a homoeologous non-reciprocal translocation (HNRT; Udall et al. 2005) of a section of chromosome A9 to its homoeologue C8 in the yellow-seeded parent GH06. Figure 4 presents a model explaining how an HNRT and accompanying knockout mutation could give rise to the DNA sequence, mapping and phenotype anomalies in GH06 and its low-ADL progenies. The pedigree of GH06 includes a yellow-seeded line derived from resynthesised *B. napus* and another ancestor originating from radiation mutagenesis. HNRT events are common during resynthesis of *B. napus* (Udall et al. 2005; Szadkowski et al. 2010), while gene knockouts caused by deletions are a common consequence of radiation mutation. According to Nicolas et al. (2009), the chromosome arms at the top of the partially homoeologous *B. napus* chromosomes A9 and C8 are particularly prone to homoeologous recombination.

CCR1, also known in Arabidopsis as *IRREGULAR XYLEM 4* (*IRX4*; At1g15950), is the first specific committed step in the biosynthesis of monolignols that in xylem are later assembled into cell wall lignin (Lacombe et al. 1997; Piquemal et al. 1998). Xylem cell walls of Arabidopsis *irx4* (*ccr1*) mutants were found to contain 50%

less lignin than wild-type plants, although their cellulose and hemicellulose content remained unchanged (Jones et al. 2001). This phenotype is consistent with the seed phenotypes seen in low seed ADL progenies from the yellow-seeded oilseed rape genotype GH06.

Few studies have specifically investigated the role of phenylpropanoid biosynthesis genes in seed testa development. In one rare example, in which no distinction was made between *CCR* homologues, Bezold et al. (2005) confirmed that hull-less mutants of pumpkin (*Cucurbita pepo* L.) show considerably lower expression of *CCR* and other genes involved in secondary cell wall biosynthesis. In Arabidopsis, the laccase gene *LAC15* (*TRANSPARENT TESTA 10*) was found by Liang et al. (2006) to play a role in monolignol polymerisation in seeds. However, *lac15/tt10* mutants also exhibit an accumulation of soluble proanthocyanidins (Pourcel et al. 2005), an association that we have not observed in *B. napus* genotypes with low seed ADL content (Snowdon et al. 2010). The role of *CCR1* in xylem lignification has been investigated in considerable detail, however, microarray data show that the same gene is also expressed at moderate to strong levels in Arabidopsis seeds during the mid to late stages of seed development (data from the Arabidopsis eFP Browser, <http://bbc.botany.utoronto.ca/>; Winter et al. 2007). In other *B. napus* genotypes, we confirmed a strong expression of *CCR1* during seed development by quantitative real-time PCR (W. Bekele, C. Obermeier and R. Snowdon, unpublished data). In Arabidopsis, the seed expression is particularly pronounced in the peripheral endosperm and the testa (data from <https://www.genevestigator.com/>; Hruz et al. 2008), supporting its participation in seed coat phenylpropanoid biosynthesis. Our low-ADL phenotype is not associated

Fig. 4 Explanation for the replacement of the functional *BnaC.CCR1.C8* locus on *B. napus* chromosome C8 with a functionally impaired homologue from chromosome A9 in GH06 by a homoeologous non-reciprocal translocation (HNRT). The native *bnA.ccr1.A9* alleles in both P174 and GH06 share functional mutations affecting amino acids near the enzyme active sites that are expected to impair the enzyme activity. Furthermore, the translocated *bnA.ccr1.A9_{fs}* allele carries a frameshift mutation (fs) that causes complete loss-of-function due to a stop codon in exon 1 (cf. Fig. 3)



with any developmental defects like those seen in *Arabidopsis irx4* mutants, hence it is likely that the *bnA.ccr1.A9-fs* mutant knocks out a homologue with seed-specific expression and that independent homologues are responsible in *B. napus* for lignin biosynthesis in xylem tissues.

CCR1 encodes the enzymatic reduction of hydroxycinnamoyl CoA esters to their corresponding cinnamoyl aldehydes during monolignol biosynthesis. In *Medicago truncatula*, Zhou et al. (2010) found that the enzyme encoded by *CCR1* prefers feruloyl CoA as a substrate, generating S and G lignin subunits in the presence of coniferaldehyde 5-hydroxylase (F5H), 5-hydroxyconiferaldehyde 5-*O*-methyltransferase (COMT) and cinnamyl alcohol dehydrogenase (CAD). The substrate preference was confirmed by Lauvergeat et al. (2001), who found that *CCR1* is five times more efficient than *CCR2* in *Arabidopsis* with feruloyl CoA and sinapoyl CoA as a substrate. This suggests a primary role of *CCR1* in constitutive phenylpropanoid biosynthesis, while *CCR2* is more active in the synthesis of phenolic compounds in response to biotic or abiotic stress (Lauvergeat et al. 2001).

BnCAD2/CAD3 may also be involved in seed-specific phenylpropanoid biosynthesis

The multiple members of the *CAD* gene family encode NADPH-dependent oxidoreductases catalysing reduction of various phenylpropenyl aldehyde derivatives (Mansell et al. 1974). This multigene family exhibits high sequence diversity, but relatively broad functional redundancy (Kim et al. 2004). In *Arabidopsis*, the xylem-expressed homologues *AtCAD4* (At3g19450) and *AtCAD5* (At4g34230), which are the most closely related family members to proven *CAD* genes from other species, were found by Kim et al. (2004) to show the highest enzymatic activity on phenylpropenyl aldehyde substrates. In contrast, the tandemly duplicated homologues *AtCAD2* (At2g21730) and *AtCAD3* (At2g21890) were found to have far lower catalytic activity. The physiological function of these two *CAD* family members has yet to be established; however, *AtCAD2* was found to use caffeoyl aldehyde most effectively as a substrate, whereas *AtCAD3* had a slight preference for *p*-coumaroyl and 5-hydroxyconiferyl aldehydes (Kim et al. 2004).

Somewhat surprisingly, both *AtCAD2* and *AtCAD3* show strong expression levels in the peripheral endosperm of developing embryos (data from <https://www.genevestigator.com/>). The spatial and temporal proximity of *AtCCR1*, *AtCAD2* and *AtCAD3* gene expression at the periphery of developing embryos are consistent with a possible joint role in seed coat phenylpropanoid biosynthesis. Most interestingly, we located a *Brassica* orthologue of the tandem

CAD2/CAD3 duplication ~660 kbp downstream of KBrH090019.5 in *B. rapa* (BRAD annotation). This raises the question of whether *BnCAD2/CAD3* may not be a similarly suitable positional candidate for the *ADL* locus in *B. napus*. However, by mapping of markers closely flanking *CAD2/CAD3* in the *B. rapa* genome sequence, we established that the *Bna.CAD2/CAD3* locus does not correspond with the *ADL* locus, instead mapping around 13 cM away on chromosome A9 (Fig. 2). Furthermore, disruption of *CAD* function on its own does not necessarily have a negative impact on overall lignin content (e.g. Chabannes et al. 2001; Kim et al. 2004), although the lignin composition may be altered (e.g. Halpin et al. 1998; Patten et al. 2005). Hence, disruption of *Bna.CAD2/CAD3* alone might not be expected to cause the dominant, low seed coat ADL phenotype of GH06 and its progenies.

Interaction of *Bna.CCR1* and *Bna.CAD2/CAD3* in seed coat phenylpropanoid biosynthesis

Nevertheless, results from other studies strongly suggest that disrupted activity of both *Bna.CCR1* and *Bna.CAD2/CAD3* in seeds may be required for the drastic, dominantly inherited reduction of seed coat ADL that we observed in GH06 and its low-ADL progenies. In particular, simultaneous reduction of CCR and CAD enzymatic activity can reduce lignin in xylem by 50% compared to wild-type levels (Thévenin et al. 2011). Chabannes et al. (2001) reported that transgenic silencing of *CCR1* by itself induced a strong reduction in lignin content in tobacco, however, only an intermediate effect was seen in hemizygous plants. This corresponds to a codominant inheritance of wild type and silenced alleles due to a gene dosage effect. On the other hand, the low-lignin phenotype was fully dominant in hemizygous *CCR1/ccr1* plants that were also hemizygous for a silenced copy of *CAD*, whereas the *cad* mutation on its own (either in homozygous or hemizygous form) had no negative effect on lignin content. Chabannes et al. (2001) concluded that the ectopic expression of *ccr1* mutants depends on the genetic background in which the mutation occurs, and that a double mutant with non-functional *ccr1* + *cad* can show drastic (dominant) reduction in lignin content even when both mutations are present in hemizygous form.

These observations fit perfectly into the context of our own data, if we assume that *Bna.CCR1* loss-of-function underlies the *ADL* locus in GH06. Figure 5 gives a hypothetical inheritance model for the *ADL* locus based on the knowledge of the interaction between *ccr1* and *cad* mutations in xylem. According to the results of Chabannes et al. (2001), the mutant *bnA.ccr1.A9-fs* from GH06 would be expected to be completely dominant over a functional *BnaC.CCR1.C8* from P174 only when the nearby

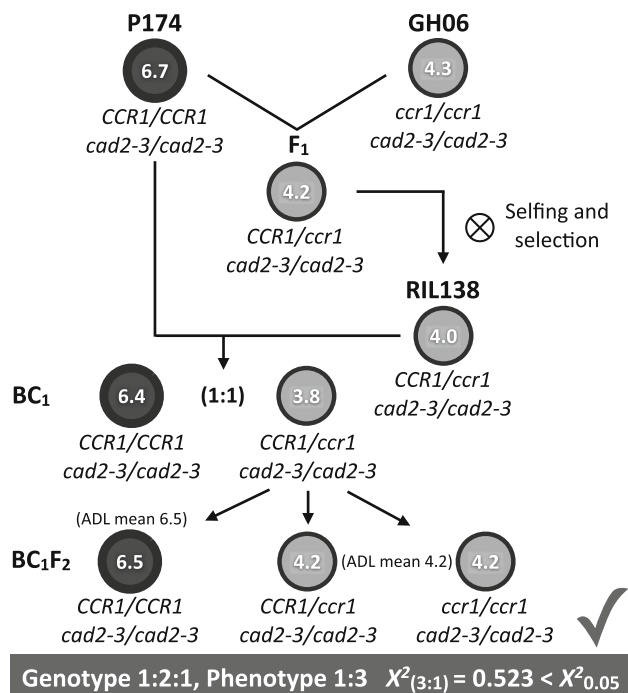


Fig. 5 Hypothesis explaining the observed BC₁F₂ segregation with the *bnaA.ccr1.A9-fs* mutation in GH06 accompanied by a putative *bna.cad2-3* mutation in both parents. Observed seed ADL contents (% DW) in each generation are shown within the circles. The *bnaA.ccr1.A9-fs* knockout mutation is only expected to be dominant in the presence of a non-functional *bna.cad2/cad3* (cf. Chabannes et al. 2001). Disruption of CAD alone does not disrupt phenylpropanoid biosynthesis, hence P174 has a normal, high-ADL phenotype. All *cad2-3* mutants containing the *bnaA.ccr1.A9-fs* knockout in either homozygous or hemizygous form should have ~50% total seed ADL, but show a normal developmental phenotype

Bna.CAD2/CAD3 locus is also silenced or non-functional in both parents. A *bna.cad2/cad3* mutation on its own should have no phenotypic effect on seed coat ADL in P174 due to the presence of a functional copy of *BnaC.CCR1.C8*. Furthermore, in crosses of high-ADL *B. napus* lines with normal *Bna.CAD2/CAD3*, only partial dominance or codominance of *bnaA.ccr1* should be observed. Crosses of GH06 to genetically diverse black-seeded lines often show an incomplete dominance consistent with this theory.

Use of *bna.ccr1* and *bna.cad2/cad3* mutants for breeding

Diagnostic markers for the *bnaA.ccr1.A9-fs* knockout mutation are expected to be particularly useful for marker-assisted selection of oilseed rape with reduced seed fibre content. Whereas NIRS is commonly used in selection for other important seed quality traits, many breeders do not have access to effective NIRS calibrations for important seed fibre compounds. Gene-based markers, therefore, offer a promising alternative for uncovering allelic

variation in breeding materials and introgressing promising alleles for low seed fibre content into elite breeding lines. Whereas previous efforts to reduce fibre content in oilseed rape relied heavily on selection based on the environmentally unstable yellow-seed trait, marker-assisted selection for functional seed coat phenylpropanoid mutants may provide a means for effective improvement of meal quality independently of seed colour. Combination of markers for *bna.ccr1* and *bna.cad2/cad3* mutations may also facilitate breeding of hybrid cultivars that take advantage of the dominant inheritance of the low seed ADL trait derived from GH06.

Acknowledgments This work was funded with a grant to LL from the “Forschungsfonds Raps” and additional support from Norddeutsche Pflanzenzucht H.G. Lembke KG. The authors thank Wubishet Bekele and Christian Obermeier for unpublished data on *BnCCR1* gene expression in *B. napus* seeds.

References

- Akhov L, Ashe P, Tan YF, Datla R, Selvaraj G (2009) Proanthocyanidin biosynthesis in the seed coat of yellow-seeded, canola quality *Brassica napus* YN01-429 is constrained at the committed step catalyzed by dihydroflavonol 4-reductase. *Botany* 87:616–625
- Auger B, Baron C, Lucas MO, Vautrin S, Bergès H, Chalhoub B, Fautrel A, Renard M, Nesi N (2009) *Brassica* orthologs from *BANYULS* belong to a small multigene family, which is involved in procyanidin accumulation in the seed. *Planta* 230:1167–1183
- Badani AG, Snowdon RJ, Wittkop B, Lipsa FD, Baetzel R, Horn R, De Haro A, Font R, Lühs W, Friedt W (2006) Colocalization of a partially dominant gene for yellow seed color with a major QTL influencing acid detergent fibre (ADF) content in different crosses of oilseed rape (*Brassica napus*). *Genome* 49:1499–1509
- Berg ES, Olaisen B (1994) Hybrid PCR sequencing—sequencing of PCR products using a universal primer. *Biotechniques* 17:896–901
- Bezold TN, Mathews D, Loy JB, Minocha SC (2005) Molecular analysis of the hull-less seed trait in pumpkin: expression profiles of genes related to seed coat development. *Seed Sci Res* 15:205–217
- Boesewinkel FD, Bouman F (1995) The seed: structure and function. In: Kigel J, Galili G (eds) *Seed development and germination*. Marcel Dekker, New York, pp 1–24
- Bonawitz ND, Chapple C (2010) The genetics of lignin biosynthesis: connecting genotype to phenotype. *Annu Rev Genet* 44:337–363
- Burdzinski C, Wendell DL (2007) Mapping the anthocyaninless (*anl*) locus in rapid-cycling *Brassica rapa* (RBr) to linkage group R9. *BMC Genet* 8:64–70
- Chabannes M, Barakate A, Lapiere C, Marita JM, Ralph J, Pean M, Danoun S, Halpin C, Grima-Pettenati J, Boudet AM (2001) Strong decrease in lignin content without significant alteration of plant development is induced by simultaneous down-regulation of *cinnamoyl CoA reductase (CCR)* and *cinnamyl alcohol dehydrogenase (CAD)* in tobacco plants. *Plant J* 28:257–270
- Chai YR, Lei B, Huang HL, Li JN, Yin JM, Tang ZL, Wang R, Chen L (2009) *TRANSPARENT TESTA 12* genes from *Brassica napus* and parental species: cloning, evolution, and differential involvement in yellow seed trait. *MGG* 81:109–123
- Downey RK, Rakow G (1987) Rapeseed and mustard. In: Fehr W (ed) *Handbook of Plant Breeding*, vol II. Macmillan Publishing Co., New York, pp 437–486

- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissues. *Focus* 12:13–15
- Fu FY, Liu LZ, Chai YR, Chen L, Yang T, Jin MY, Ma AF, Yan XY, Zhang ZS, Li NJ (2007) Localization of QTLs for seed color using recombinant inbred lines of *Brassica napus* in different environments. *Genome* 50:840–854
- Halpin C, Holt K, Chojecki J, Oliver D, Chabbert B, Monties B, Edwards K, Barakate A, Foxon GA (1998) *Brown-midrib* maize (*bml*)—a mutation affecting the *cinnamyl alcohol dehydrogenase* gene. *Plant J* 14:545–553
- Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, Oertle L, Widmayer P, Gruissem W, Zimmermann P (2008) Genevestigator V3: a reference expression database for the meta-analysis of transcriptomes. *Adv Bioinforma* 2008:420747
- Jones L, Ennos AR, Turner SR (2001) Cloning and characterization of *IRREGULAR XYLEM 4 (IRX4)*: a severely lignin-deficient mutant of *Arabidopsis*. *Plant J* 26:205–216
- Kim SJ, Kim MR, Bedgar DL, Moinuddin SGA, Cardenas CL, Davin LB, Kang CH, Lewis NG (2004) Functional reclassification of the putative cinnamyl alcohol dehydrogenase multigene family in *Arabidopsis*. *Proc Natl Acad Sci USA* 10:1455–1460
- Kosambi DD (1994) The estimation of map distances from recombination values. *Ann Eugen* 12:172–175
- Lacombe E, Hawkins S, Van Dorsselaere J, Piquemal J, Goffner D, Poeydomenge O, Boudet AM, Grima-Pettenati J (1997) Cinnamoyl CoA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: cloning, expression and phylogenetic relationships. *Plant J* 11:429–441
- Lauvergeat V, Lacomme C, Lacombe E, Lasserre E, Roby D, Grima-Pettenati J (2001) Two cinnamoyl-CoA reductase (CCR) genes from *Arabidopsis thaliana* are differentially expressed during development and in response to infection with pathogenic bacteria. *Phytochemistry* 57:1187–1195
- Li JN, Zhang XK, Chen L, Chi C, Wang R (1998) Initial study on the genetics of seed color of yellow-seeded lines (*Brassica napus* L.) from different genetic sources. *Chin J Oil Crop Sci* 20:16–19
- Li AM, Jiang JJ, Zhang YT, Snowdon RJ, Liang GH, Wang YP (2011) Molecular and cytological characterization of introgression lines in yellow seed derived from somatic hybrids between *Brassica napus* and *Sinapis alba*. *Mol Breed*. doi:10.1007/s11032-010-9540-z
- Liang M, Davis E, Gardner D, Cai X, Wu Y (2006) Involvement of *AtLAC15* in lignin synthesis in seeds and in root elongation of *Arabidopsis*. *Planta* 224:1185–1196
- Liu Z, Fu TD, Tu JX, Chen BY (2005) Inheritance of seed colour and identification of RAPD and AFLP markers linked to the seed colour gene in rapeseed (*Brassica napus* L.). *Theor Appl Genet* 110:303–310
- Mansell R, Gross GG, Stöckigt J, Franke H, Zenk MH (1974) Purification and properties of cinnamyl alcohol dehydrogenase from higher plants involved in lignin biosynthesis. *Phytochemistry* 13:2427–2435
- Marles MAS, Gruber MY (2004) Histochemical characterisation of unextractable seed coat pigments and quantification of extractable lignin in the Brassicaceae. *J Sci Food Agric* 84:251–262
- Marles MAS, Gruber MY, Scoles GJ, Muir AD (2003) Pigmentation in the developing seed coat and seedling leaves of *Brassica carinata* is controlled at the dihydroflavonol reductase locus. *Phytochemistry* 62:663–672
- Martins WS, Lucas DC, Neves KF, Bertoli DJ (2009) WebSat—a web software for microsatellite marker development. *Bioinformatics* 3:282–283
- Matthäus B (1998) Effect of dehulling on the composition of antinutritive compounds in various cultivars of rapeseed. *Fett Lipid* 100:295–301
- Nicolas SD, Leflon M, Monod H, Coriton O, Huteau V, Chèvre A-M, Jenczewski E (2009) Genetic regulation of meiotic cross-overs between related genomes in *Brassica napus* haploids and hybrids. *Plant Cell* 21:373–385
- Patten AM, Cardenas CL, Cochran FC, Laskar DD, Bedgar DL, Davin LB, Lewis NG (2005) Reassessment of effects on lignification and vascular development in the *irx4* *Arabidopsis* mutant. *Phytochemistry* 66:2092–2107
- Piquemal J, Lapierre C, Myton K, O'Connell A, Schuch W, Grima-Pettenati J, Boudet AM (1998) Down-regulation of cinnamoyl-CoA reductase induces significant changes of lignin profiles in transgenic tobacco plants. *Plant J* 13:71–83
- Pourcel L, Routaboul J-M, Kerhoas L, Caboche M, Lepiniec L, Debeaujon I (2005) *TRANSPARENT TESTA 10* encodes a laccase-like enzyme involved in oxidative polymerization of flavonoids in *Arabidopsis* seed coat. *Plant Cell* 17:2966–2980
- Rashid A, Rakow G, Downey RK (1994) Development of yellow seeded *Brassica napus* through interspecific crosses. *Plant Breed* 112:127–134
- Shirzadegan M (1986) Inheritance of seed color in *Brassica napus* L. *Zeitung für Pflanzenzüchtung* 96:140–146
- Simbaya J, Slominski BA, Rakow G, Campbell LD, Downey RK, Bell JM (1995) Quality characteristics of yellow-seeded *Brassica* seed meals: protein, carbohydrates, and dietary fiber components. *J Agric Food Chem* 43:2062–2066
- Snowdon RJ, Wittkop B, Rezaidad A, Hasan M, Lipsa F, Stein A, Friedt W (2010) Regional association analysis delineates a sequenced chromosome region influencing antinutritive seed meal compounds in oilseed rape. *Genome* 53:917–928
- Szadkowski E, Eber F, Huteau V, Lodé M, Huneau C, Belcram H, Coriton O, Manzanares-Dauleux MJ, Delourme R, King GJ, Chalhou B, Jenczewski E, Chèvre AM (2010) The first meiosis of resynthesized *Brassica napus*, a genome blender. *New Phytol* 186:102–112
- Tang ZL, Li JN, Zhang XK, Chen L, Wang R (1997) Genetic variation of yellow-seeded rapeseed lines (*Brassica napus* L.) from different genetic sources. *Plant Breed* 116:471–474
- Thévenin J, Pollet B, Letarnec B, Saulnier L, Gissot L, Maia-Grondard A, Lapierre C, Jouanin L (2011) The simultaneous repression of CCR and CAD, two enzymes of the lignin biosynthetic pathway, results in sterility and dwarfism in *Arabidopsis thaliana*. *Mol Plant* 4:70–82
- Udall JA, Quijada PA, Osborn TC (2005) Detection of chromosomal rearrangements derived from homologous recombination in four mapping populations of *Brassica napus* L. *Genetics* 169:967–979
- Van Deynze A, Pauls KP (1994) The inheritance of seed color and vernalization requirement in *Brassica napus* and using doubled haploid populations. *Euphytica* 74:77–83
- Van Deynze AE, Beversdorf WD, Pauls KP (1993) Temperature effects on seed colour in black-and yellow-seeded rapeseed. *Can J Plant Sci* 73:383–387
- Van Ooijen JW, Voorrips RE (2001) JoinMap® 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen
- Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Freijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* 23:4407–4414
- Wang S, Basten CJ, Zeng ZB (2006) Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC, USA
- Wei YL, Li JN, Lu J, Tang ZL, Pu DC, Chai YR (2007) Molecular cloning of *Brassica napus* *TRANSPARENT TESTA 2* gene family encoding potential MYB regulatory proteins of proanthocyanidin biosynthesis. *Mol Biol Rep* 34:105–120

- Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart N (2007) An “Electronic Fluorescent Pictograph” browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* 2:e718
- Wittkop B, Snowdon RJ, Friedt W (2009) Status and perspectives of breeding for enhanced yield and quality of oilseed crops for Europe. *Euphytica* 170:131–140
- Xiao S, Xu J, Li Y, Zhang L, Shi S, Shi S, Wu J, Liu K (2007) Generation and mapping of SCAR and CAPS markers linked to the seed coat color gene in *Brassica napus* using a genome-walking technique. *Genome* 50:611–618
- Xu BB, Li JN, Zhang XK, Wang R, Xie LL, Chai YR (2007) Cloning and molecular characterization of a functional flavonoid 3'-hydroxylase gene from *Brassica napus*. *J Plant Physiol* 164:350–363
- Zhou R, Jackson L, Shadle G, Nakashima J, Temple S, Chen F, Dixon R (2010) Distinct cinnamoyl CoA reductases involved in parallel routes to lignin in *Medicago truncatula*. *Proc Natl Acad Sci USA* 107:17803–17808

Dissection of a major QTL for seed colour and fibre content in *Brassica napus* reveals colocalization with candidate genes for phenylpropanoid biosynthesis and flavonoid deposition

ANNA STEIN, BENJAMIN WITTKOP, LIEZHAO LIU, CHRISTIAN OBERMEIER, WOLFGANG FRIEDT and ROD J. SNOWDON*

Department of Plant Breeding, Justus Liebig University, Heinrich-Buff-Ring 26-32, 35392, Giessen, Germany; *Corresponding author, E-mail: rod.snowdon@agrar.uni-giessen.de

With 3 figures and 4 tables

Received October 31, 2012/Accepted March 21, 2013

Communicated by A.-M. Chevre

Abstract

A major quantitative trait locus (QTL) influencing seed fibre and colour in *Brassica napus* was dissected by marker saturation in a doubled haploid (DH) population from the black-seeded oilseed rape line 'Express 617' crossed with a yellow-seeded *B. napus* line, '1012-98'. The marker at the peak of a sub-QTL with a strong effect on both seed colour and acid detergent lignin content lay only 4 kb away from a *Brassica* (H⁺)-ATPase gene orthologous to the transparent testa gene *AHA10*. Near the peak of a second sub-QTL, we mapped a copy of the key phenylpropanoid biosynthesis gene cinnamyl alcohol dehydrogenase, while another key phenylpropanoid biosynthesis gene, cinnamoyl co-a reductase 1, was found nearby. In a cross between 'Express 617' and another dark-seeded parent, 'V8', *Bna.CCR1* was localized *in silico* near the peak of a corresponding seed fibre QTL, whereas in this case *Bna.CAD2/CAD3* lay nearby. Re-sequencing of the two phenylpropanoid genes via next-generation amplicon sequencing revealed intragenic rearrangements and functionally relevant allelic variation in the three parents.

Key words: oilseed rape — *CCR1* — *CAD* — *AHA10* — seed meal quality

Oilseed rape or canola (*Brassica napus* L.) varieties with reduced levels of antinutritive seed phenolic acid and fibre compounds would considerably improve the value of the seed meal after oil extraction (Kracht et al. 2004). Recently, strong associations were reported between seed colour and the content of acid detergent lignin (ADL), the most important antinutritive fibre compound in oilseed rape and canola seed meal (Wittkop et al. 2009, 2012 and Liu et al. 2012). A major quantitative trait locus (QTL) with very large effects on seed colour, condensed tannin content and seed acid detergent fibre (ADF) was previously mapped to chromosome A9 in the black-seeded x yellow-seeded doubled haploid population YE2-DH (Badani et al. 2006, Snowdon et al. 2010, Lipsa et al. 2012). In crosses between yellow-seeded and dark-seeded Chinese oilseed rape, the phenylpropanoid biosynthesis gene *CINNAMOYL CO-A REDUCTASE 1* (*CCR1*) was revealed as a positional and functional candidate gene for a major QTL on chromosome A9 that explained a considerable proportion of the variation in ADL content (Liu et al. 2012). That study suggested that this *Bna.CCR1* homologue might potentially interact with neighbouring homologues of a *CINNAMYL ALCOHOL DEHYDROGENASE* (*CAD*) gene family member to regulate seed coat phenylpropanoid biosynthesis in

B. napus (Liu et al. 2012). Functional mutations in either of these genes can significantly influence the biosynthesis of monolignols, lignans and other antinutritive phenolic acid derivatives, the effect being greatest in double mutants (Chabannes et al. 2001). Hence, both genes are potentially interesting targets for molecular breeding of oilseed rape and canola with improved nutritional quality of the seed meal.

On the other hand, *TRANSPARENT TESTA* (*TT*) genes are also considered as important candidates for the explanation of seed colour differences in *Brassica* species (Fu et al. 2007, Wei et al. 2007, Xu et al. 2007, Akhrov et al. 2009, Auger et al. 2009). A number of *TT* genes have been found on chromosome A9 in *B. rapa* (Fu et al. 2007), but efforts to associate important QTL with *TT* loci in *B. napus* were unsuccessful so far.

Materials and Methods

Plant materials and field trials: Two half-sib doubled haploid (DH) winter oilseed rape populations were used in this study. The first population (YE2-DH) comprised 165 DH lines from a cross between the black-seeded inbred line 'Express 617' and the yellow-seeded line '1012-98', a resynthesized *B. napus* that originated from a spontaneous hybridization between *B. oleracea* and *B. rapa* (Gland 1982, see Badani et al. 2006 for further details). The second population (ExV8-DH) comprised 94 DH lines from the cross of 'Express 617' to the black-seeded, semisynthetic *B. napus* line 'V8'. The 'V8' parent was derived by backcrossing of a resynthesized rapeseed genotype derived from an interspecific cross between the Indian *B. rapa* ssp. *trilocularis* (Yellow Sarson) accession 'YSPb-24' and the cauliflower (*B. oleracea* L. convar. *botrytis*) accession 'Super Regama' (BK 2096) with a high erucic acid line (Lühs and Friedt 1995, Basunanda et al. 2007). To account for genotype-by-environment interactions, the phenotyping for the QTL analysis was conducted on seed samples produced at a total of six different locations in Germany, over multiple years from 2003 until 2009 (Table 1). Inflorescences were bagged at the onset of flowering to ensure self-pollination, and ripe seeds were collected from two to five plants per genotype.

Quantification of seed colour and fibre compounds: All collected seed samples were screened by near-infrared reflectance spectrometry (NIRS) for non-destructive quantification of seed pigmentation and fibre content. Trait values were measured on seeds from two to five self-pollinated plants per genotype per environment and averaged over two technical repetitions. Phenotype values for seed colour (visual light absorbance), along with contents of ADL, ADF and neutral detergent fibre (NDF), were extrapolated from NIRS calibrations developed specifically for the

Table 1: Locations and harvesting seasons used for quantitative trait locus (QTL) analysis of seed colour and fibre traits in the winter oilseed rape doubled haploid mapping populations YE2-DH ('Express 617' × '1012-98') and ExV8-DH ('Express 617' × 'V8')

Population	No. of DH lines	Location	Years
YE2-DH	165	Rauischholzhausen	2003, 2009
		Gross Gerau	2006
		Seligenstadt	2006
		Einbeck	2007
		Total environments	5
ExV8-DH	94	Einbeck	2006
		Reinshof	2006
		Grundschwalheim	2006, 2007
		Rauischholzhausen	2006, 2007, 2009
		Total environments	7

measurement of these traits as a proportion of seed dry weight in *B. napus* (Wittkop *et al.* 2012). The ADL fraction represents lignin and precursors including lignans, monolignols and related phenolic complexes. The ADF fraction additionally contains cellulose, while NDF contains both cellulose and hemicellulose in addition to the ADL fraction.

Genetic mapping and QTL analysis: Genomic DNA was extracted from leaf samples of all parental and DH lines using standard procedures (Doyle and Doyle 1990). For ExV8-DH, a new genetic map was generated using a large set of simple sequence repeat (SSR) markers, single-nucleotide polymorphism (SNP) markers and restriction site-associated DNA (RAD) markers. A previously available genetic map for YE2-DH was used as a basis for marker saturation in the region of the major QTL for seed colour and ADF content that was previously identified by Badani *et al.* (2006). For comparative mapping, the region surrounding this QTL was first saturated using *in silico* SSR and gene-derived markers within the corresponding region of chromosome A9 from the *Brassica rapa* genome sequence (*Brassica rapa* Genome Sequencing Project Consortium. 2011). Primer sequences for the new markers are given in Table 2. The primers were tested for polymorphism in the parents of YE2-DH and ExV8-DH and used to saturate and align the genetic maps at the QTL position. Additional SSR markers published by Liu *et al.* (2012) were used to map the corresponding QTL region in our material.

New dense maps of A9 for YE2-DH and ExV8-DH were generated with the software Joinmap 4.0 (Van Ooijen 2006) using the mapping function of Kosambi (1944). The complex polyploid structure of the *B. napus* genome, along with chromosome rearrangements in this genome region in synthetic *B. napus*, can lead to potential map distortion (Liu *et al.* 2012). Therefore, in the linkage analysis, we only included markers whose segregation did not show excessive distortion from the expected Mendelian ratio of 1 : 1, represented by a goodness-of-fit χ^2 -value <4.0. Linkage was calculated with a maximum recombination

frequency of 0.4 and a logarithm of odds (LOD) score of >1.0. QTL were localized by composite interval mapping (CIM) using the software WinQTL Cartographer 2.5 (Wang *et al.* 2006) and PLABQTL 1.2 (Utz and Melchinger 1996). LOD thresholds for QTL detection were set to $P = 0.05$ and were determined using permutation tests with 1000 permutations (Churchill and Doerge 1994). The QTL analysis was performed for each trait separately and over individual environments for both populations.

Where possible, the sequences of SSR, SNP and RAD markers near the peaks of the detected QTL were physically located in the chromosome A9 genome sequence of *Brassica rapa* v1.1, accessed in January 2013 via the *B. rapa* genome browser available at <http://www.phytozome.net/> (Goodstein *et al.* 2012) and the BLAST tool available at <http://brassicadb.org/> (Cheng *et al.* 2011). This enabled the determination of estimated physical distances from QTL-linked markers to phenylpropanoid or *TT* genes potentially associated with the respective QTL.

Identification of the candidate gene *Bna.CAD2/CAD3* in 'Express 617': High-density filters of our BRASSIKEY *B. napus* BAC library for the common parent 'Express 617', with 8x genome coverage, were hybridized with probes developed from the sequence-characterized amplified region (SCAR) markers At3g17310 and KBrB025M01.1, respectively. These two markers closely flanked the peak of the major main-effect QTL for ADL and seed colour on chromosome A9 in a previous low-density version of the YE2-DH genetic map. A total of 63 clones found to be positive for one or both of these markers were subsequently isolated. DNA was extracted with the Large Construct Kit from Qiagen (Hilden, Germany) from single-cell cultures of positive BAC clones. Fingerprint patterns were established for all BAC clones using PCR markers spanning the major QTL region on chromosome A9. Based on these patterns, three BAC clones were selected that best represented this region at putative homoeologous loci. The clones were subjected to shotgun sequencing by Seq-IT (Kaiserslautern, Germany) on a Roche 454-FLX Titanium next-generation sequencer (454 Life Sciences, Branford, CT, USA) at an average depth of coverage of 30x for comparative sequence analysis. The *B. rapa* BAC KBrH108D07 was used as a reference sequence for mapping of the reads. Contig assembly and analysis of DNA sequence and structural variation among the three homoeologous BACs was performed using the software MAUVE 2.3.1 (Darling *et al.* 2004). Gene content of the three 'Express 617' BACs was derived from the BAC annotation tool supplied by the *Brassica* Genome Gateway (<http://brassica.nbi.ac.uk/>) and revealed a homologue of a mannitol dehydrogenase gene with extremely high similarity (95% identity) to both *A. thaliana* *CAD2* (At2g21730) and *A. thaliana* *CAD3* (At2g21890). Sequence identity to a *B. rapa* cinnamyl alcohol dehydrogenase (Bra031216) was even higher (100%).

Curiously, both *B. rapa* and *B. napus* 'Express 617' exhibit an identical chimeric structure within the gene sequence of *CAD*, possibly originating from a fusion of *AtCAD2* and *AtCAD3* through chromosomal rearrangements during the ancestral polyploidization of the *Brassica* diploids. We thus designated the chimeric *B. napus* gene as *Bna.CAD2/CAD3*. Using the *Bna.CAD2/CAD3* sequence from 'Express 617', we

Table 2: Primer sequences for newly developed cleaved amplified polymorphic sequence (CAPS), sequence-characterized amplified region (SCAR) and simple sequence repeat (SSR) markers spanning the quantitative trait locus (QTL) region on *Brassica napus* chromosome A9

Markers	Forward primer sequence	Reverse primer sequence	Type
BnCADb2VspI	CCCGGGTATCCACACAATA	CCAACGATTTCATGCCTAGC	CAPS (VspI)
BnCCR1HphI	AATCATCTATTTCAAAACATGTAACG	TTTCACACCAAATCTTAATTGGTAAA	CAPS (HphI)
At3g17310	CCTATGTTTCCCAAGGGTT	TGTCAGAAAAATGCAACGG	SCAR
F7D8	CTACCCTCTCTCAACTCTCC	GGTCTTGGTCTCCTTCTGT	SSR
KBrB023K01.9	CAAAATCCCTCCTCCTCTC	TGGCAAAAACAGAAAAGTCG	SSR
KBrB025M01.1	ACAAGACCCTGGAACGG	CCTAAGGAAAGACGAAAGAAT	SCAR
KBrB036I13.1	CTGAACAATTTAGCGTGTGAGC	GATCGCTCTTGAGAATCGTCTT	SSR
KBrB082F21-2F	CCCCAAAAACAGTGCCAAAG	GGCCCTAATCCCCATCACA	SSR
KBrH092O19.SP6	GCCAGATACGTTGCAGGTTT	TTGACATACCTGCATACTCAAAGA	SSR
KBrH108D07	GCGTCCCCTGATGTTGAG	CGAATCCGAATCCCCTTG	SCAR

subsequently developed the CAPS marker BnCADb2VspI to map this gene in the YE2-DH population.

Identification of the candidate gene *Bna.CCR1*: Comparison of the genetic linkage maps for YE2-DH and P174*GH06 (Liu et al. 2012) revealed shared SSR loci within the respective QTL from these two populations. In a fine-mapping study, Liu et al. (2012) identified *CCR1* as a positional and functional candidate gene in very close proximity to a major locus for ADL on chromosome A9. This raised the possibility that *CCR1* might simultaneously be a potential candidate gene for the QTL for ADL on chromosome A9 in both the German and the Chinese rapeseed material. Additional SSR markers from the linkage maps reported by Liu et al. (2012) were therefore mapped in YE2-DH to further saturate the QTL region on chromosome A9. Unfortunately, none of the markers linked to the QTL of Liu et al. (2012) were polymorphic in ExV8-DH. Instead, we added *B. napus* SNP and RAD markers (Obermeier, unpublished data) to chromosome A9 of ExV8-DH, in order to accurately align the genetic map to the *B. rapa* genome sequence for indirect alignment with the YE2-DH map.

Genomic re-sequencing of *Bna.CCR1* and *Bna.CAD2/CAD3* homoeologues: The universal-tailed 454-FLX amplicon sequencing procedure described by Gholami et al. (2012) was used to recover and re-sequence *B. napus* homoeologues of the phenylpropanoid biosynthesis genes *Bna.CCR1* (At1g15950; Bra026737) and *Bna.CAD2/CAD3* (At2g21730/At2g21890; Bra031216). Recovery of *Bna.AHA10* (Bra039228) sequences was not feasible due to the very high number of closely related gene family members, all of which have multiple homoeologues in *B. napus*.

Pairs of primers were designed to generate genomic PCR amplicons that covered the full lengths of the coding sequences (exons) of the selected genes. Oligonucleotide sequences for the tailed, gene-specific primers are listed in Table 3. All amplicons were sequenced together by Seq-IT (Kaiserslautern, Germany) as part of a larger sequencing run on a 1/4 plate of a Roche 454-FLX Titanium next-generation DNA sequencer (454 Life Sciences, Branford, CT). Raw reads were sorted according to the amplicon-specific primers and trimmed to exon borders. The reads were aligned for each genotype and amplicon individually and grouped into contigs to identify homoeologous loci using the Dirty Data Alignment Algorithm with preferred 3' gap placement, a minimum match percentage of 85–95 and a minimum overlap of 20. Consensus sequences were generated by plurality for each contig separately. After the removal of low-frequency contigs, which are likely to represent PCR errors, the grouped contigs were aligned to the *Bra.A.CAD2/CAD3* (Bra031216) and *Bra.A.CCR1* (Bra026737) coding sequences obtained from The Brassica Database (BRAD, Cheng et al. 2011). Sequences of the *B. oleracea* orthologues of these genes (Bol044768 and Bol009845, respectively)

were obtained from the same database. Comparison to these references enabled assignment of all amplicons to their putative chromosomes in the *B. napus* genome. The trimmed consensus nucleotide sequences of all detected *Bna.A.CAD2/CAD3* and *Bna.A.CCR1* exons are given in Electronic Supporting Information (Data S1 and Data S2), respectively. Additionally, the amplicon sequence data were used to develop *CCR1* gene markers for YE2-DH. Unfortunately, these markers were not polymorphic in ExV8-DH, thus preventing mapping of the gene in this population.

Results

Key phenylpropanoid biosynthesis genes are associated with an overlapping QTL for ADF and ADL in yellow-seeded and dark-seeded genetic backgrounds

Figure 1 illustrates the recalculated genetic maps for linkage group A9 in YE2-DH and ExV8-DH, along with QTL positions for seed fibre compounds and seed colour. Additional information on the QTL is given in Table 4. Alignment of the two genetic maps was achieved by projection of markers onto the physical genome sequence of *B. rapa*.

A total of 8 new SSR markers derived from the *B. rapa* genome sequence were mapped to chromosome A9 in YE2-DH. Unfortunately, either most of the SSR markers in the QTL region on chromosome A9 in YE2-DH were not polymorphic in ExV8-DH, or only homoeologous loci that mapped to chromosome C8 were polymorphic (data not shown). Hence, we generated a completely new map for chromosome A9 in ExV8-DH, which is based on SNP and RAD marker data that could be indirectly aligned to the YE2-DH map via the *B. rapa* genome sequence, based on the physical marker positions.

Quantitative trait loci for seed colour, NDF, ADF and ADL were re-calculated after mapping in both populations. Due to the high recombination frequency in the chromosome region spanning the QTL (Snowdon et al. 2010, Liu et al. 2012), the increase in marker density enabled a higher-resolution mapping of the QTL for seed colour and fibre. This enabled the QTL to be split into two peaks with varying effects on seed colour and/or fibre components, respectively.

In YE2-DH, a sub-QTL with phenotypic effects of 5.9%, 4.1%, 7.7% and 7.3% on seed colour, NDF, ADF and ADL, respectively, was flanked by the SSR markers CN79 and KBrB082F21. These markers also flanked the major QTL for seed ADL content reported by Liu et al. (2012) and fine-mapped

Table 3: Universal-tailed primer sequences for multiplex 454 amplicon sequencing of full-length coding sequences from *Bna.CAD2/CAD3* and *Bna.CCR1* homoeologues. Details of the tailed PCR and sequencing procedure can be found in Gholami et al. (2012)

Genes	Amplicon primer name	Primer sequence
<i>Bna.CAD2/CAD3</i>	BnaCAD3.Exon1+2 FWD	TTTCCCAGTCACGACGTTTGACAGTCTGACAGAGAGACGA
	BnaCAD3.Exon1+2 REV	TAATACGACTCACTATAGGGCGAGAGAGTATTTGTGTGGATACC
	BnaCAD3.Exon3a FWD	TTTCCCAGTCACGACGTTTTTCGAATGTATTTCTATTTTGTCT
	BnaCAD3.Exon3a REV	TAATACGACTCACTATAGGGAAGGCTTTGCCAATCTTCAC
	BnaCAD3.Exon3b FWD	TTTCCCAGTCACGACGTTTCCAATGAAATATTATGGTATGACTGA
	BnaCAD3.Exon3b REV	TAATACGACTCACTATAGGGGAAGAGTCATCTGTTTCTTTAC
	BnaCAD3.Exon4 FWD	TTTCCCAGTCACGACGTTTGTTTACGCTTTTGTGCAAG
	BnaCAD3.Exon4 REV	TAATACGACTCACTATAGGGTCGGAATCTTTTGTCTTTATCA
	BnaCCR1.Exon1+2 FWD	TTTCCCAGTCACGACGTTATGCCAGTCGACGAATCTAC
	BnaCCR1.Exon1+2 REV	TAATACGACTCACTATAGGGTCATGTAGGTCCCTAACTTTATCAA
<i>Bna.CCR1</i>	BnaCCR1.Exon3 FWD	TTTCCCAGTCACGACGTTAATCATCTATTTCAAAACATGTAACG
	BnaCCR1.Exon3 REV	TAATACGACTCACTATAGGGTTTCACACCAATCTTAATTGGTAAA
	BnaCCR1.Exon4 FWD	TTTCCCAGTCACGACGTTTGCTTTTGCACTTAACAAACATGA
	BnaCCR1.Exon4 REV	TAATACGACTCACTATAGGGAAGTGGCATTGGTTTCACTTA
	BnaCCR1.Exon5 FWD	TTTCCCAGTCACGACGTTTCTTTGGTCTATGCCCACTTT
	BnaCCR1.Exon5 REV	TAATACGACTCACTATAGGGCCGATCCTGATACCGTTTGT

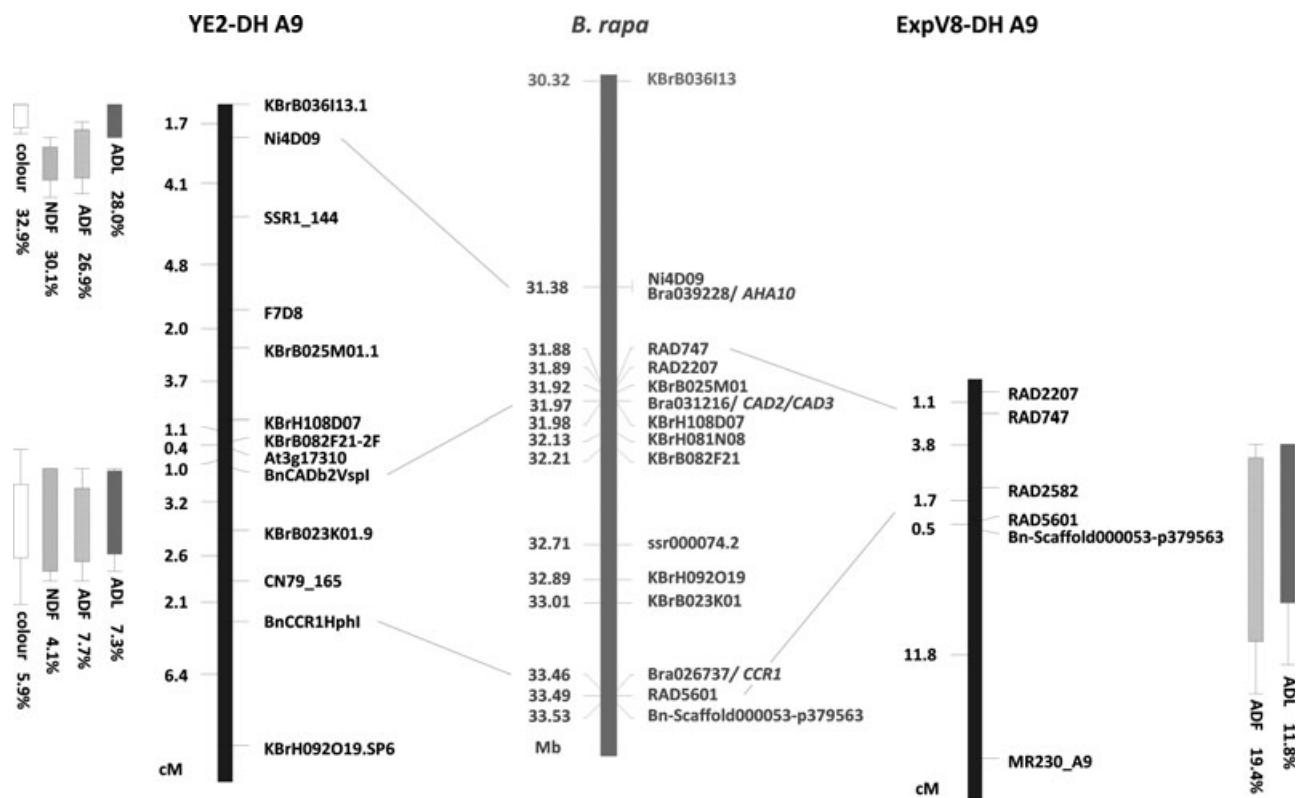


Fig. 1: Comparison of physical marker positions to the reference sequence of *B. rapa* revealed the colocalization of quantitative trait loci (QTL) for acid detergent fibre (ADF) and acid detergent lignin (ADL) on *Brassica napus* chromosome A9 in doubled haploid (DH) mapping populations from the crosses ‘Express 617’ × ‘1012–98’ (YE2-DH) and ‘Express 617’ × ‘V8’ (ExV8-DH), respectively. The linkage groups show incomplete chromosome sections with high marker densities, spanning the respective QTL on chromosome A9. The markers BnCADb2Vspl and BnCCR1HphI in the YE2-DH map are gene-derived markers for *Bna.CAD2/CAD3* (Bra031216) and *Bna.CCR1* (Bra026737), respectively

Table 4: List of quantitative trait loci (QTL) for seed colour and fibre compounds detected on chromosome A9 by composite interval mapping (CIM) in the populations YE2-DH and ExV8-DH. 1-LOD and 2-LOD confidence interval (IV) state the 1- or 2-LOD decrease from the QTL peak, respectively, as interval lengths in cM

Population	Trait	Position [cM]	Nearest marker to peak	LOD	Additive effect	Partial R ² [%]	1-LOD IV [cM]	2-LOD IV [cM]
YE2-DH	Seed colour	1.0	Ni4D09	16.0	0.66	32.9	1.2	1.5
	Seed colour	20.8	KBrB023K01.9	3.6	0.27	5.9	3.8	8.0
	NDF	2.7	Ni4D09	11.9	0.63	30.1	2.6	2.2
	NDF	20.8	KBrB023K01.9	2.1	0.22	4.1	5.3	5.8
	ADF	2.7	Ni4D09	12.0	0.57	26.9	3.3	2.9
	ADF	20.8	KBrB023K01.9	4.0	0.29	7.7	3.8	5.8
	ADL	1.0	KBrB036113.1	14.0	0.50	28.0	1.7	1.7
	ADL	20.8	KBrB023K01.9	4.3	0.25	7.3	4.3	5.3
ExV8-DH	ADF	88.8	RAD5601	5.2	0.35	19.4	9.5	13.9
	ADL	89.8	Bn-Scaffold000053-p379563	3.4	0.29	11.8	11.4	8.2

NDF, neutral detergent fibre; ADF, acid detergent fibre; ADL, acid detergent lignin.

in that study to the position of the phenylpropanoid gene *Bna.CCR1*. Unexpectedly, however, the gene-derived marker BnCCR1HphI in YE2-DH was mapped 2.1 cM downstream of CN79, outside of the 2-LOD QTL confidence interval.

On the other hand, the gene-derived marker BnCADb2Vspl for *Bna.CAD2/CAD3* mapped directly within the 1-LOD confidence interval for NDF, ADF and ADL in YE2-DH. This suggests that *Bna.CAD2/CAD3* may have more influence on seed coat phenolic fibre compounds than *Bna.CCR1* in this particular cross.

In ExV8-DH, QTL peaks for ADF and ADL were detected that explained 19.4% and 11.8%, respectively, of the phenotypic

variation for these traits. The QTL peaks are flanked by the markers RAD747 (*B. rapa* position A9:31884700 bp), RAD5601 (*B. rapa* position A9:33492962 bp) and SNP marker BnScaffold000053-p379563 (*B. rapa* position A9:33529680 bp). In the *B. rapa* genome, *CCR1* (Bra026737) is located between these QTL-flanking markers at position A9:33454096–33456663. A nearby *CAD2/CAD3* orthologue (Bra031216) is annotated at position A9:31975046–31973109, very close (approximately 90 kb) to the physical position of the QTL-flanking marker RAD747. Synteny of this region between *B. rapa* and *B. napus* has been confirmed by a stepwise BLAST analysis to sequence contigs from ‘Express 617’ (B. Samans, University of Giessen,

personal communication), suggesting that *Bna.CAD2/CAD3* is also close to the QTL peak in ExV8-DH. Unfortunately, we failed to genetically map *Bna.CCR1* or *Bna.CAD2/CAD3* in ExV8-DH due to a lack of polymorphism in this cross for the markers we developed for YE2-DH. Nevertheless, the sequence projections of the flanking markers predict that *Bna.CCR1* is present in the major QTL for seed fibre in the dark-seeded genetic background of ExV8-DH, as it was suggested to also be in the yellow-seeded Chinese *B. napus* investigated by Liu et al. (2012).

A *TRANSPARENT TESTA* gene maps in silico to a major QTL for seed colour

Addition of new markers to the genetic map of chromosome A9 in YE2-DH resulted in rearrangement of marker orders, achieving a better correspondence to the physical order in the *B. rapa* genome. At the same time, the major QTL for seed colour, which also has significant (presumably pleiotropic) effects on seed coat thickness and fibre content, was dissected into the minor sub-QTL described above along with a major sub-QTL that had a very strong influence particularly on seed colour (partial $R^2 = 32.9\%$).

The improved marker order presumably comes about because the greater marker density more accurately resolves mapping errors due to putative homoeologous translocations, which are particularly common between this part of A9 and a highly homoeologous region of chromosome C8 (Nicolas et al. 2009). Despite these inaccuracies, in all versions of the genetic map, the SSR marker Ni4D09 was consistently located very close to the peak of the major QTL for seed colour and seed fibre traits. Given the very high recombination frequencies in this chromosome region (Snowdon et al. 2010, Liu et al. 2012), this suggests a close physical proximity of this marker to the responsible gene(s). Ni4D09 is an EST-derived SSR whose sequence is derived from the Arabidopsis unigene At2g24100. A BLAST search for this gene in the *Brassica* A genome, using the *B. rapa* genome browser available on www.phytozome.net (Goodstein et al. 2012), revealed high homology (e-41) to the *Brassica* unigene Bra039229, located as expected on chromosome A9. The very next gene, located only 4 kbp away in the *B. rapa* genome,

is the plasma membrane proton-ATPase Bra039228, which is an orthologue of the Arabidopsis *TRANSPARENT TESTA* gene *AHA10*. *AHA10* is a seed-expressed H⁺-ATPase (Harper et al. 1994) required for proanthocyanidin formation in the seed coat endothelium (Baxter et al. 2005). Thus, the ATPase Bra039228 was revealed for the first time as a highly promising positional and functional candidate gene that may have a major impact on seed colour variation in the *B. napus* population investigated here.

Deep re-sequencing of *Bna.CAD2/CAD3* and *Bna.CCR1* reveals considerable genomic rearrangements and amino acid substitutions

Amino acid positions identifying the derivation of non-terminated *Bna.CAD2/CAD3* exons and *Bna.CCR1* exons from the diploid A or C genome, respectively, are shown in Figs 2 and 3, respectively. Complex rearrangements within A and C genome homoeologues along with a number of putatively functional amino acid modifications were identified in both *Bna.CAD2/CAD3* and *Bna.CCR1*.

For *Bna.CAD2/CAD3*, 'Express 617' exhibited three allelic variants for exons 1 and 2, two of which correspond to the amino acid sequence from the A genome of *B. rapa* and one of which corresponds to the sequence from the C genome of *B. oleracea*. In all three *B. napus* genotypes 'Express 617', '1012-98' and 'V8', the longest exon of *Bna.CAD2/CAD3*, exon 3, was found to represent a chimeric sequence comprising partial gene sequences from both *B. rapa* and *B. oleracea*. In particular, the region of exon 3, which was uniquely represented by the *B. oleracea* homologue, contains the alcohol dehydrogenase (ADH) zinc-binding domain. Additionally, one of the A genome alleles of *Bna.CAD2/CAD3* was missing in 'V8', which has an intermediate-ADL content, and neither of the A genome alleles was found in the low-ADL line '1012-98'. Instead, '1012-98' contains an additional different C genome copy of exons 1 and 2. Compared to the wild-type allele from *B. rapa*, however, the short (90-bp) exon 1 encodes four consecutive amino acids (positions 12, 13, 14 and 15) with considerably differing properties. Furthermore, at amino acid position 35, asparagine (N) is replaced by aspartate (D), a modification with a potential

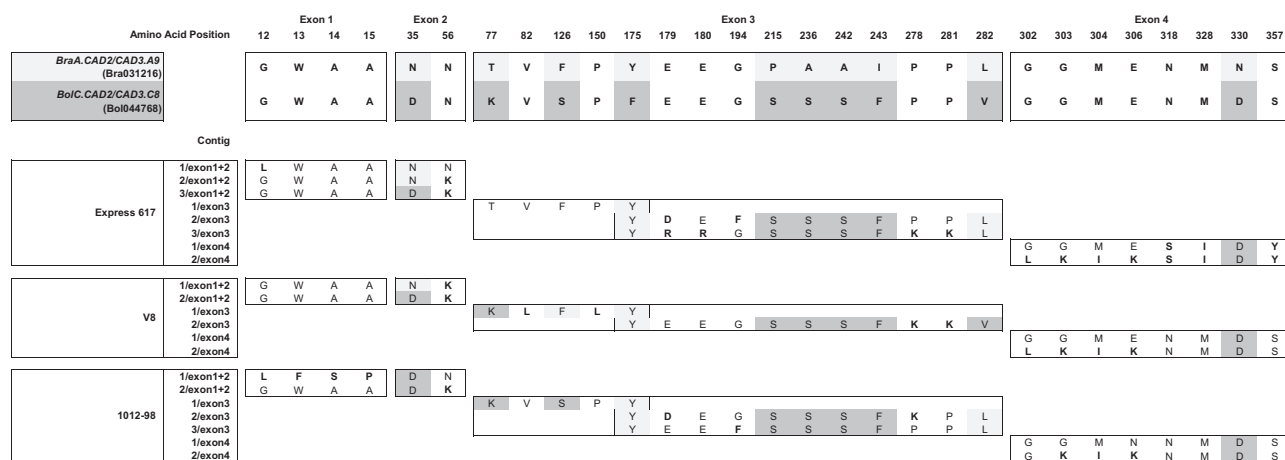


Fig. 2: Amino acid variation in non-truncated copies of the four exons of *Bna.CAD2/CAD3* between the genotypes 'Express 617', 'V8' and '1012-98' compared to the A and C genome reference sequences Bra031216 (light shading) and Bol044768 (dark shading). Only those positions are shown which allow discrimination between A genome and C genome or diverge from both (bold type). Contigs are numbered consecutively and do not correspond to locus numbers. The vertical lines represent the borders of the two different PCR amplicons that were used to sequence the full length of exon 3

Amino Acid Position		Exon 1		Exon 2					Exon 3					Exon 4			Exon 5		
		8	23	45	47	77	93	94	106	107	123	136	151	183	253	273	321	322	326
BraA.CCR1.A9 (Bra026737)		P	A	D	P	A	T	D	G	A	V	N	L	L	T	L	H	L	P
BolC.CCR1.C8 (Bol009845)		A	A	D	P	S	T	D	G	A	V	S	L	L	S	I	H	L	S
Contig																			
Express 617	1/exon1+2	P	P	D	P	A	T	D											
	2/exon1+2	P	A	D	P	A	T	E											
	3/exon1+2	A	A	D	P	S	T	D											
	1/exon3								G	A	V	N	F						
	2/exon3								G	A	V	S	F						
	1/exon4													L	S	L			
V8	1/exon1+2	P	A	D	P	A	T	E											
	2/exon1+2	A	A	D	P	S	I	E											
	1/exon3								G	A	A	N	F						
	2/exon3								G	A	V	S	F						
	1/exon4													M	S	L			
	2/exon4													L	S	I			
1012-98	1/exon1+2	P	P	A	D	A	T	D											
	1/exon3								P	T	V	S	F						
	2/exon3								G	G	V	S	F						
	1/exon4													L	S	I			
	1/exon5																H	L	S
	1/exon5																		

Fig. 3: Amino acid variation in the five exons of *Bna.CCR1* between the genotypes 'Express 617', 'V8' and '1012-98' compared to the A and C genome reference sequences Bra026737 (light shading) and Bol009845 (dark shading). Only those amino acids are shown which allow discrimination between A genome and C genome or diverge from both (bold type). For exon 5, only the first diverging amino acid is shown, which allows discrimination between the A and C genome locus. Bra026737 and Bol009845 differ in the last 17 amino acids. Contigs are numbered consecutively and do not correspond to locus numbers

functional influence on the GroES-like zinc-binding domains encoded by exon 2. In addition, all detected copies of exon 3 in 'V8' and '1012-98' share a basic lysine (K) at amino acid position 77, close to the substrate-binding site, instead of the polar neutral threonine (T) present in all copies of this exon in the high-ADL line 'Express 617'. Finally, in all three genotypes, non-functional paralogues of *Bna.CAD2/CAD3* were also amplified, containing frameshift mutations resulting in premature stop codons in exon 3. Further modification of between 4 and 7 amino acids in the GroES-like domain encoded by exon 4 was seen in *B. napus* compared to both diploid progenitors.

Comparison of the amplicon sequences of *Bna.CCR1* homoeologues in 'Express 617', '1012-98' and 'V8' with the reference coding sequences from the diploid *Brassica* species revealed a minor translocation in exon 4 of *B. napus*, however without obvious impact on functionality. For '1012-98', however, rearranged gene sequences were detected that corresponded neither to the complete A genome homologue from *B. rapa* chromosome A9, nor to the C genome homologue from *B. oleracea* chromosome C8. Instead, the sequence data suggest that a homoeologous translocation has created a novel chimeric allele, because parts of the coding sequence correspond to the sequence of the A genome homologue from *B. rapa* and parts match the C genome homologue from *B. oleracea*. However, long-range PCR successfully amplified the whole gene length of *Bna.CCR1* in '1012-98', thus suggesting that a chimeric but intact novel copy of the gene was formed (data not shown). Frameshift mutations were observed which led to a knockout in one allele. Another frameshift mutation in exon 5 was identified consistently in one allele of all three genotypes '1012-98', 'Express 617' and 'V8'. For the intermediate-ADL line 'V8', a putatively functional amino acid substitution at position 123 from valine (V) to alanine (A) and another at position 183 from leucine (L) to methionine (M) was observed in the A genome homologue

from *B. rapa*. The CCR protein harbours active sites, NAD(P)H-binding and substrate-binding sites in close vicinity to these amino acids (Liu *et al.* 2012).

Discussion

The genetic map of YE2-DH was enriched in the major QTL region for seed colour on chromosome A9, using markers developed from *B. rapa* sequence data. The map saturation in YE2-DH revealed skewed marker frequencies and mapping ambiguities consistent with an assumption of a C8 to A9 translocation in the parental line '1012-98', as described by Liu *et al.* (2012) for another synthetic, yellow-seeded *B. napus* genotype. This may explain why linkage group A9 in the YE2-DH map was previously identified by Badani *et al.* (2006) as C8 (N18), using only terminal consensus markers. The addition of further markers indicates that the middle of the linkage group in YE2-DH, where the QTL are located, in fact corresponds to chromosome A9. Translocation events are frequently observed in allopolyploid *B. napus*, particularly in resynthesized genotypes (Parkin *et al.* 1995, Sharpe *et al.* 1995). The terminal chromosome arms of the partially homoeologous *B. napus* chromosomes A9 and C8 are known to be especially prone to homoeologous recombination (Nicolas *et al.* 2009).

The genetic and/or physical positions of the SSR markers KBrH092O19 and KBrB025M01, which flank the common QTL for fibre compounds in ExV8-DH and YE2-DH, correspond exactly with a major QTL for seed colour and ADL content found previously in a completely unrelated Chinese oilseed rape cross segregating for seed colour (Liu *et al.* 2012). In that study, Liu *et al.* (2012) fine-mapped this locus to the position of a *Bna.CCR1* candidate gene. The apparent colocalization with markers from the seed fibre QTL in the present study suggests that this gene may influence seed fibre not only in yellow-seeded

backgrounds, but also in dark-seeded oilseed rape. The effect putatively involves a restriction of monolignol accumulation due to an enzymatic block involving CCR1 and/or CAD2/CAD3. In both genes, we observed a high degree of allelic diversity with amino acid substitutions, some within putative active sites of the enzymes. The low-ADL mutant '1012-98' did not carry an original A or C genome copy of *Bna.CCR1*. Instead, '1012-98' carried a novel chimeric allele of *Bna.CCR1*. Furthermore, '1012-98' also contained no sequences corresponding to an A genome copy of *Bna.CAD2/CAD3*. This constitution of a double mutant at the A9 locus may contribute to the low-ADL phenotype observed for '1012-98', whereby the mapping of gene-derived markers suggests that the major causal gene in YE2-DH is *Bna.CAD2/CAD3*.

In 'V8', which exhibits an intermediate-ADL phenotype, we detected only a functionally impaired *Bna.CCR1.A9* sequence along with one completely non-functional C genome sequence that contained a frameshift knockout mutation in exon 3 (*bna.ccr1.C8-fs*). The colocalization of *Bna.CCR1* with the QTL for seed fibre in ExV8-DH suggests that one or both of these two mutations may be responsible for the intermediate-ADL phenotype of 'V8'. In contrast, the high-ADL line 'Express 617' possessed a fully functional *Bna.CCR1.A9* locus accompanied by the *bna.ccr1.C8-fs* knockout locus and functional A9 and C8 copies of *Bna.CAD2/CAD3*. Further work is needed to confirm whether *Bna.CAD2/CAD3* also lies within the confidence interval of the QTL in ExV8-DH, as it does in YE2-DH.

In YE2-DH, the marker nearest to the peak of a major QTL for seed colour was found to be extremely close (4 kbp) to the deduced position of a proton-ATPase gene highly similar to the *TT* gene *AHA10*. Baxter et al. (2005) reported that a knockout mutation of *A. thaliana AHA10*, which is involved in transporting flavonoid intermediates for proanthocyanidin production, resulted in a *transparent testa* phenotype, which is analogous to the yellow-seeded phenotype in *Brassica* species. A blocked transport prevents the accumulation of proanthocyanidins in the seed coat and their oxidation into brown pigment. Further work is required to confirm the presence and identity of an endothelium-specific (H⁺)-ATPase gene at this major QTL, a task that is complicated by the large number of highly similar orthologues distributed across the genome. Seven suspected gene family members alone are annotated to chromosome A9 in *B. rapa*. Nevertheless, the close physical proximity to the peak of this major QTL suggests a potential role of this gene. Confirmation of this suspicion would significantly advance our understanding of the genetics of pigmentation and fibre accumulation in *B. napus* seed coats and represent a major step towards marker-assisted breeding for improved oilseed rape meal quality.

Conclusions

Circumvention of blocked enzymatic steps in the phenylpropanoid biosynthesis pathway can lead to the integration of accumulating intermediates rather than the normal monolignols into the polymeric end-products. Alteration in the monolignol composition can subsequently cause changes in the final phenylpropanoid composition. In xylem, where polymerized lignin is the most abundant end-product of phenylpropanoid biosynthesis, this compensatory phenomenon for re-channelling of phenolic intermediates is known as lignin plasticity (Bonawitz and Chapple 2010). The presence of polymerized lignin has not been unequivocally demonstrated in seed coats of *Arabidopsis* and *Brassica* species; however, poorly digestible phenolic polymers

and their precursors contribute significantly to ADL, the most nutritionally relevant extractable fibre fraction in animal feeds derived from crushed oilseeds (Wittkop et al. 2012). Improvement of meal digestibility in oilseed rape lines expressing functionally modified or defective CCR1, CAD and/or AHA10 enzymes in their seed coats is therefore a promising prospect for elevation of the total seed value. We have demonstrated that a reduction in antinutritive fibre is also possible in high-performing, dark-seeded oilseed rape lines that do not seem to suffer from yield penalties associated with physiological or phytopathological effects of yellow-seeded colour. The DNA sequences and markers identified in the present study therefore represent an important step towards identifying additional useful mutants and implementing them in elite breeding programmes. However, further physiological evidence is needed to prove the hypothesized functional impact of defective gene copies.

Acknowledgements

The authors acknowledge Theophilus Tettey and Mahmood Gholami for assistance with the 454 amplicon sequencing and Alexander Grüner for assistance with mapping *Bna.CCR1*. We also thank Isobel Parkin for providing Illumina SNP data and Birgit Samans for assistance with sequence analysis. Funding was provided by the German Federal Ministry of Education and Research (BMBF) grant 0313703A and DFG grant SN14/12-1.

References

- Akhov, L., P. Ashe, Y. F. Tan, R. Datla, and G. Selvaraj, 2009: Proanthocyanidin biosynthesis in the seed coat of yellow-seeded, canola quality *Brassica napus* YN01-429 is constrained at the committed step catalyzed by dihydroflavonol 4-reductase. *Botany* **87**, 616–625.
- Auger, B., C. Baron, M. O. Lucas, S. Vautrin, H. Bergès, B. Chalhoub, A. Fautrel, M. Renard, and N. Nesi, 2009: Brassica orthologs from BANYULS belong to a small multigene family, which is involved in procyanidin accumulation in the seed. *Planta* **230**, 1167–1183.
- Badani, A. G., R. J. Snowdon, B. Wittkop, F. D. Lipsa, R. Baetzel, P. Basunanda, T. H. Spiller, M. Hasan, A. Gehringer, J. Schondelmaier, W. Lühs, and W. Friedt, 2006: Colocalization of a partially dominant gene for yellow seed colour with a major QTL influencing acid detergent fibre (ADF) content in different crosses of oilseed rape (*Brassica napus*). *Genome* **49**, 1499–1509.
- Basunanda, P., T. H. Spiller, M. Hasan, A. Gehringer, J. Schondelmaier, W. Lühs, W. Friedt, and R. J. Snowdon, 2007: Marker-assisted increase of genetic diversity in a double-low seed quality winter oilseed rape genetic background. *Plant Breed.* **126**, 581–587.
- Baxter, I. R., J. C. Young, G. Armstrong, N. Foster, N. Bogenschütz, T. Cordova, W. A. Peer, S. P. Hazen, A. S. Murphy, and J. F. Harper, 2005: A plasma membrane H⁺-ATPase is required for the formation of proanthocyanidins in the seed coat endothelium of *Arabidopsis thaliana*. *PNAS* **102**, 2649–2654.
- Bonawitz, N. D., and C. Chapple, 2010: The genetics of lignin biosynthesis: connecting genotype to phenotype. *Annu. Rev. Genet.* **44**, 337–363.
- Brassica rapa Genome Sequencing Project Consortium: Wang, X., H. Wang, J. Wang, R. Sun, J. Wu, S. Liu, Y. Bai, J. H. Mun, I. Bancroft, F. Cheng, S. Huang, X. Li, W. Hua, J. Wang, X. Wang, M. Freeling, J. C. Pires, A. H. Paterson, B. Chalhoub, B. Wang, A. Hayward, A. G. Sharpe, B. S. Park, B. Weisshaar, B. Liu, B. Li, B. Liu, C. Tong, C. Song, C. Duran, C. Peng, C. Geng, C. Koh, C. Lin, D. Edwards, D. Mu, D. Shen, E. Soumpourou, F. Li, F. Fraser, G. Conant, G. Lassalle, G. J. King, G. Bonnema, H. Tang, H. Wang, H. Belcram, H. Zhou, H. Hirakawa, H. Abe, H. Guo, H. Wang, H. Jin, I. A. Parkin, J. Batley, J. S. Kim, J. Just, J. Li, J. Xu, J. Deng, J. A. Kim, J. Li, J. Yu, J. Meng, J. Wang, J. Min, J. Poulain, J. Wang, K. Hatakeyama, K. Wu, L. Wang, L. Fang, M. Trick, M. G. Links, M. Zhao, M. Jin,

- N. Ramchiary, N. Drou, P. J. Berkman, Q. Cai, Q. Huang, R. Li, S. Tabata, S. Cheng, S. Zhang, S. Zhang, S. Huang, S. Sato, S. Sun, S. J. Kwon, S. R. Choi, T. H. Lee, W. Fan, X. Zhao, X. Tan, X. Xu, Y. Wang, Y. Qiu, Y. Yin, Y. Li, Y. Du, Y. Liao, Y. Lim, Y. Narusaka, Y. Wang, Z. Wang, Z. Li, Z. Wang, Z. Xiong, and Z. Zhang, 2011: The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* **43**, 1035–1039.
- Chabannes, M., A. Barakate, C. Lapierre, J. M. Marita, J. Ralph, M. Pean, S. Danoun, C. Halpin, J. Grima-Pettenati, and A. M. Boudet, 2001: Strong decrease in lignin content without significant alteration of plant development is induced by simultaneous down-regulation of *Cinnamoyl CoA Reductase (CCR)* and *Cinnamyl Alcohol Dehydrogenase (CAD)* in tobacco plants. *Plant J.* **28**, 257–270.
- Cheng, F., S. Liu, J. Wu, L. Fang, S. Sun, B. Liu, P. Li, W. Hua, and X. Wang, 2011: BRAD, the genetics and genomics database for *Brassica* plants. *BMC Plant Biol.* **11**, 136.
- Churchill, G., and R. Doerge, 1994: Empirical threshold values for quantitative trait mapping. *Genetics* **138**, 963–971.
- Darling, A. C. E., B. Mau, F. R. Blattner, and N. T. Perna, 2004: Multiple alignment of conserved genomic sequences with rearrangements. *Genome Res.* **14**, 1394–1403.
- Doyle, J. J., and J. L. Doyle, 1990: A rapid total DNA preparation procedure for fresh plant tissue. *Focus* **12**, 13–15.
- Fu, F. Y., L. Z. Liu, Y. R. Chai, L. Chen, T. Yang, M. Y. Jin, A. F. Ma, X. Y. Yan, Z. S. Zhang, and N. J. Li, 2007: Localization of QTLs for seed color using recombinant inbred lines of *Brassica napus* in different environments. *Genome* **50**, 840–854.
- Gholami, M., W. A. Bekele, J. Schondelmaier, and R. J. Snowdon, 2012: A tailed PCR procedure for cost-effective, two-order multiplex sequencing of candidate genes in polyploid plants. *Plant Biotechnol. J.* **10**, 635–645.
- Gland, A., 1982: Gehalt und Muster der Glucosinolate in den Samen von resynthetisierten Rapsformen. *Z. Pflanzenzüchtung* **88**, 242–254.
- Goodstein, D. M., S. Shu, R. Howson, R. Neupane, R. D. Hayes, J. Fazo, T. Mitros, W. Dirks, U. Hellsten, N. Putnam, and D. S. Rokhsar, 2012: Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* **40**, D1178–D1186.
- Harper, J. F., L. Manney, and M. R. Sussman, 1994: The plasma membrane H⁺-ATPase gene family in *Arabidopsis*: genomic sequence of *AHA10* which is expressed primarily in developing seeds. *Mol. Genet. Genet.* **244**, 572–587.
- Kosambi, D. D., 1944: The estimation of map distance from recombination values. *Ann. Eugen.* **12**, 172–175.
- Kracht, W., S. Dänicke, H. Kluge, K. Keller, W. Matzke, U. Hennig, and W. Schumann, 2004: Effect of dehulling of rapeseed on feed value and nutrient digestibility of rape products in pigs. *Arch. Anim. Nutr.* **58**, 389–404.
- Lipsa, F. D., R. J. Snowdon, and W. Friedt, 2012: Quantitative genetic analysis of condensed tannins in oilseed rape meal. *Euphytica* **184**, 195–205.
- Liu, L. Z., A. Stein, B. Wittkop, P. Sarvari, J. Li, X. Yan, F. Dreyer, M. Frauen, W. Friedt, and R. J. Snowdon, 2012: A knockout mutation in the lignin biosynthesis gene *CCR1* explains a major QTL for acid detergent lignin content in *Brassica napus* seeds. *Theor. Appl. Genet.* **124**, 1573–1586.
- Lühs, W., and W. Friedt, 1995: Natural fatty acid variation in the genus *Brassica* and its exploitation through resynthesis. *Cruciferae Newsl.* **17**, 14–15.
- Nicolas, S. D., M. Leflon, H. Monod, O. Coriton, V. Huteau, A.-M. Chèvre, and E. Jenczewski, 2009: Genetic regulation of meiotic cross-overs between related genomes in *Brassica napus* haploids and hybrids. *Plant Cell* **21**, 373–385.
- Parkin, I. A. P., A. G. Sharpe, D. J. Keith, and D. J. Lydiat, 1995: Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape). *Genome* **38**, 1122–1131.
- Sharpe, A. G., I. A. P. Parkin, D. J. Keith, and D. J. Lydiat, 1995: Frequent nonreciprocal translocation in the amphidiploid genome of oilseed rape (*Brassica napus*). *Genome* **38**, 1112–1121.
- Snowdon, R. J., B. Wittkop, A. Rezaei, M. Hasan, F. Lipsa, A. Stein, and W. Friedt, 2010: Regional association analysis delineates a sequenced chromosome region influencing antinutritive seed meal compounds in oilseed rape. *Genome* **53**, 917–928.
- Utz, H. F., and A. E. Melchinger, 1996: PLABQTL: a program for composite interval mapping of QTL. *J. Quant. Trait Loci* **2**, 1–5.
- Van Ooijen, J. W., 2006: JoinMap® 4: Software for the Calculation of Genetic Linkage Maps in Experimental Populations. Kyazma B.V., Wageningen, Netherlands.
- Wang, S., C. J. Basten, and Z. B. Zeng, 2006: Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC, USA.
- Wei, Y. L., J. N. Li, J. Lu, Z. L. Tang, D. C. Pu, and Y. R. Chai, 2007: Molecular cloning of *Brassica napus* *TRANSPARENT TESTA 2* gene family encoding potential MYB regulatory proteins of proanthocyanidin biosynthesis. *Mol. Biol. Rep.* **34**, 105–120.
- Wittkop, B., R. J. Snowdon, and W. Friedt, 2009: Status and perspectives of breeding for enhanced yield and quality of oilseed crops for Europe. *Euphytica* **170**, 131–140.
- Wittkop, B., R. J. Snowdon, and W. Friedt, 2012: New NIRS calibrations for fibre fractions reveal broad genetic in *Brassica napus* seed quality. *J. Agric. Food Chem.* **60**, 2248–2256.
- Xu, B. B., J. N. Li, X. K. Zhang, R. Wang, L. L. Xie, and Y. R. Chai, 2007: Cloning and molecular characterization of a functional flavonoid 3-hydroxylase gene from *Brassica napus*. *J. Plant Physiol.* **164**, 350–363.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Data S1. Trimmed consensus nucleotide sequences for exon-derived *Cinnamyl Alcohol Dehydrogenase (Bna.CAD2/CAD3)* amplicons in the *Brassica napus* genotypes ‘Express 617’, ‘V8’ and ‘1012-98’, respectively.

Data S2. Trimmed consensus nucleotide sequences for exon-derived *Cinnamoyl Co-A Reductase 1 (Bna.CCR1)* amplicons in the *Brassica napus* genotypes ‘Express 617’, ‘V8’ and ‘1012-98’, respectively.

Diskussion

Die beiden in dieser Arbeit zusammengefaßten Studien „A knockout mutation in the lignin biosynthesis gene *CCR1* explains a major QTL for acid detergent lignin content in *Brassica napus* seeds“ und „Dissection of a major QTL for seed colour and fibre content in *Brassica napus* reveals colocalisation with candidate genes for phenylpropanoid biosynthesis and flavonoid deposition“ beschreiben zum ersten Mal die potentielle Beteiligung von funktionellen Mutationen in Genen der Ligninbiosynthese an der Bildung von ernährungsrelevanten Samenfaserkomponenten in Rapssamen. Die Studien untersuchten insgesamt vier Rapspopulationen aus zwei grundsätzlich verschiedenen Genpools:

- 1) die chinesische DH-Population P174*GH06,
- 2) eine neue Population aus Rückkreuzung (*backcross*, *BC*) von P174 mit einer rekombinanten Inzuchtlinie (*recombinant inbred line*, *RIL*) aus der vorgenannten Kreuzung,
- 3) die DH-Population Express 617*1012-98 (YE2-DH),
- 4) die DH-Population Express 617*V8 (ExV8-DH).

Alle Populationen haben gemeinsam, daß Haupt-QTL für Samenfarbe und für Samenfaserkomponenten auf Kopplungsgruppe A9 kartieren. Ein Haupt-QTL für diese Merkmale ist in der Vergangenheit auch in anderem Material nachgewiesen worden, etwa von Zhang et al. (2011), ohne daß die genetische Grundlage geklärt werden konnte.

Mögliche funktionelle Bedeutung der QTL

Die Samenfarbe, bzw. die Intensität der Pigmentierung der Samenschale geht auf zwei wesentliche Faktoren zurück. Zum einen akkumulieren hier Proanthocyanidine, eine Gruppe von Flavonoiden, die sich durch Oxidation braun färben und damit die Samenfarbe bestimmen. Die Oxidation wird im Wesentlichen durch Catechol-Oxidasen und Laccasen katalysiert (Pourcel et al. 2006). Zum anderen spielt die Anatomie der Samenschale eine Rolle. Proanthocyanidine reichern sich im inneren Integument der Samenschale an (Marles und Gruber 2004). Die räumliche Akkumulation von Proanthocyanidinen kann daher unmittelbar durch Veränderungen in der Zellstruktur beeinträchtigt werden. Zu solchen Veränderungen zählen die tatsächliche Größe der Zellen des inneren Integuments, aber auch die innere Struktur und die Kompartimentierung dieser Zellen.

Das Merkmal Samenfarbe ist für die Verwertung des Rapsschrotes und daher für die Züchtung aus unterschiedlichen, auch konträren Gründen interessant: Einerseits erfüllen

Flavonoide in der Pflanze vielfältige Funktionen. Sie stärken die Widerstandsfähigkeit der Pflanze gegen UV-Strahlung, Pathogene und Herbivoren. Positiv zu bewerten ist auch ihre Fähigkeit, freie Radikale wie reaktive Sauerstoffspezies (*reactive oxygen species, ROS*) abzufangen. Damit stärkt und schützt eine dunkle Samenfarbe das Korn. Andererseits führt die Eigenschaft der Flavonoide, Proteine zu komplexieren, zu einer reduzierten Verdaulichkeit des Rapsschrotes in der Tierernährung (Pourcel et al. 2006). Züchtung, die darauf abzielt, den Gehalt an Flavonoiden zu senken, muß daher beide Aspekte im Blick behalten, um einen verbesserten Futterwert des Schrotes bei gleichzeitig unbeeinträchtigter Vitalität der Pflanze bzw. des Saatgutes und des Keimlings zu erzielen. Inwieweit sich hier Spielraum ergibt, wird auch aus der Arbeit von Wittkop et al. (2012) deutlich. Sie dokumentiert anhand von Nahinfrarotspektrometriedaten (NIRS) eine beträchtliche Variation der nominellen Farbwerte selbst in uniform schwarzsamig erscheinenden Rapslinien. Dieser Beobachtung entsprechend konnte auch in der vorliegenden Arbeit für die schwarzsamige Population ExV8-DH ein QTL für Samenfarbe ermittelt werden. Darüber hinaus korreliert das Merkmal Samenfarbe signifikant mit ADL ($r=0,834$ in YE2-DH und $r=0,621$ in ExV8-DH, Wittkop et al. 2012). Für die Selektion mittels NIRS bedeutet dies, daß bei einer Auslese auf niedrige Samenfarbewerte auch Individuen erfaßt werden, die ebenfalls niedrige ADL-Werte aufweisen. Dies trifft vor allem auf gelbsamige oder braunsamige Genotypen mit besonders niedrigen Samenfarbwerten zu. Umgekehrt aber können Rapskörner mit mittleren ADL-Werten mittlere bis hohe Farbwerte aufweisen. Die hohe Korrelation der Merkmale entspringt der Tatsache, daß sowohl ADL als auch Samenfarbe Konzentrationsparameter sind, die unter anderem von der Struktur der Samenschale, genauer gesagt von der Stärke der Palisadenschicht abhängen. Je dünner die Palisadenschicht ausgeprägt ist, desto geringer ist zwangsläufig auch die Akkumulation von Polyphenolen und Ligninen in der Samenschale (Anjou et al. 1977, Theander et al. 1977, Slominski et al. 1994, Stringam et al. 1974). Dicke Samenschalen mit großen Zellen und dicken Zellwänden werden ausnahmslos bei schwarzsamigen Genotypen gefunden, die außerdem hohe Fasergehalte aufweisen, während dünne Samenschalen mit Gelbsamigkeit und niedrigem Fasergehalt assoziiert sind (Baetzel 2004, Stringam et al. 1974). Neben der Samen-anatomie bietet die Biosynthese der Proanthocyanidine, Polyphenole und Phenylpropanoide weitere Variationsquellen für die qualitätsrelevanten Parameter Samenfarbe und Samenfasergehalt, etwa aufgrund genetischer Effekte oder Umwelteinflüsse. Es gibt eine Vielzahl an Beispielen von Arabidopsis-Mutanten, bei denen eine Störung der Flavonoidbiosynthese vorliegt, welche die Pigmentierung erheblich beeinflusst. Namentlich Mutationen sogenannter *transparent testa (tt)* und *tannin-deficient seed (tds)* Gene sind hier zu nennen (Shirley et al. 1995, Abrahams et al. 2002).

In der hier untersuchten schwarzsamigen Population ExV8-DH konnte jeweils ein QTL für ADL und Samenfarbe ermittelt werden. Diesen QTL liegen vermutlich Mutationen in essentiellen Strukturgenen zugrunde. Auch in der Kreuzung von Express 617 mit dem gelbsamigen Elter 1012-98 sind diese beiden QTL zu finden. Hier kolokalisiert jedoch der QTL für Samenfarbe mit einem zweiten QTL für ADL. Ein Grund dafür könnte darin liegen, daß sich in dieser Region Gene befinden, die die Struktur der Samenschale bestimmen. Ein definitiver Beweis hierfür konnte jedoch bisher nicht erbracht werden, da bislang keine Kandidatengene identifiziert wurden, die tatsächlich die Dicke der Zellwand und die Zellgröße beeinflussen. Immerhin legen die für *B. napus* abgeleiteten physischen Positionen des flankierenden SSR-Markers Ni4D09 nahe, daß sich in unmittelbarer Nähe (4kb) ein Gen für eine Protonen-ATPase befindet. Baxter et al. (2005) beschreiben für eine Mutation einer eng verwandten ATPase (*autoinhibited H⁺ATPase*, AHA10) in *Arabidopsis* einen gelbsamigen Phänotyp durch strukturelle Veränderung der zentralen Vakuole. Es handelt sich demnach um ein hochinteressantes Kandidatengen, welches auf seine Beteiligung an der Ausprägung der Samenfarbe hin weiter untersucht werden sollte.

Bewertung der genetischen Kartierung und QTL-Bestimmung: Möglichkeiten und Grenzen der Methoden

Für die markergestützte Selektion ist die Identifizierung relevanter Genomregionen, oder günstigenfalls von Kandidatengenen für die Ausprägung eines wirtschaftlich interessanten Merkmals essentiell. Dazu ist es wichtig, eng mit dem oder den QTL gekoppelte Marker zu identifizieren. Die gewonnenen Markerdaten werden dazu mit Phänotypdaten auf Basis der genetischen Karte miteinander verrechnet. Dafür sollten möglichst alle Kopplungsgruppen der Karte gleichmäßig durch Marker abgedeckt sein, wobei jedoch zu enge Markerabfolgen zu vermeiden sind. Um diese Anforderungen zu erfüllen, wurden in dieser Arbeit die Karten manuell auf eine Dichte reduziert, die bei minimal 2cM zwischen benachbarten Markern liegt. Um zu vermeiden, daß falsch kartierte Marker die Lokalisation der QTL beeinträchtigen, wurden verschiedene Markersets aus der genetischen Karte gebildet und die QTL-Berechnung mit den beiden QTL-Programmen WinQTLCart 2.5 (Wang et al. 2012) und PlabQTL 1.2 (Utz und Melchinger 1996) durchgeführt. QTL-Positionen, die sich anhand dieses Ansatzes als konsistent erwiesen, wurden als reale QTL betrachtet.

Zur Erstellung der genetischen Karten wurde das Programm Joinmap 4.0 (Stam 1993, Van Oijen 1996) verwendet. Dieses Programm berechnet die paarweisen Rekombinationshäufigkeiten der Marker, fügt die Marker sukzessive in eine lineare Anordnung ein

und erstellt so eine genetische Karte. Dieser Vorgang wird mehrfach mit variierender Reihenfolge der hinzugefügten Marker wiederholt, so daß letztlich die bestangepaßte Karte durch Versuch und Irrtum gefunden wird (Stam 1993). Die Auswahl der Marker durch das Programm ist dabei abhängig von den verfügbaren Markerinformationen. So stören abweichende Spaltungsverhältnisse ($\chi^2 \gg 0$) und vor allem fehlende Genotypwerte die Platzierung des Markers in der Karte. Bei der Erstellung einer genetischen Karte kann man jedoch selten auf einen perfekten Datensatz zurückgreifen, so daß für jeden Marker einzeln entschieden werden muß, ob seine Kartierung informativ ist, oder ob eher ein starker Störungseffekt zu erwarten ist. Im Fall der Karte für YE2-DH standen zahlreiche Marker für die Kopplungsgruppe A9 zur Verfügung. Dies ermöglichte eine stringente Auslese der zu verwendenden Marker zu einem informativen Datensatz. Da es sich bei dem Elterngenotyp 1012-98 um den Nachkommen einer resynthetisierten *B. napus*-Akzession handelt, war damit zu rechnen, daß eventuell Chromosomenmutationen (siehe folgender Abschnitt) die Chromosomenpaarung in der Meiose und damit die Rekombination beeinflussen könnten. Dies könnte dazu führen, daß in der Nachkommenschaft der Kreuzung mit Express 617 Haplotypen mit einer stark abweichenden Reihenfolge der Marker entstehen. Solche Haplotypen stören in der Berechnung einer genetischen Karte. In der Folge ist die Identifizierung einzelner relevanter Marker oder Genomregionen für die phänotypische Variation eines Merkmals sehr kompliziert und unter Umständen fehlerbehaftet. Die Ergebnisse verschiedener Kartierungsansätze in der Population YE2-DH mit unterschiedlichen Markersets deuten an, daß die doppelhaploide Nachkommenschaft bezüglich der Markerreihenfolge auf A9 tatsächlich sehr divers ist. Der genotypische Datensatz wurde daher stringent auf Marker reduziert, die einen χ^2 -Wert von maximal 2 aufwiesen. Für die Kartierung der Population ExV8-DH wurden Marker mit einem χ^2 -Wert von bis 4 toleriert, um eine ausreichende Markeranzahl berücksichtigen zu können. Weiterhin wurden solche Marker von der Kartenberechnung ausgeschlossen, deren relative Position innerhalb von A9 sich während wiederholter Kartierungen als inkonsistent erwies. Dies traf beispielsweise auf den Marker ENA21 zu. Andere Marker fielen dadurch auf, daß sie unerwartet auch auf C8 kartierten. Diese Marker wurden ebenfalls nicht in der Kartenberechnung berücksichtigt. Zwar können Marker, die aus dem A-Genom von *B. rapa* entwickelt wurden und unerwartet auf der homologen Kopplungsgruppe C8 im C-Genom kartieren, tatsächlich wegen der hohen Sequenzähnlichkeit (Syntänie) zwischen A- und C-Genom auf homologen Loci basieren. Es kann sich aber auch um A-Genom-Loci handeln, die durch Translokation zwischen A9 und C8 ausgetauscht sind. Ob es sich um eine potentielle Deletion-Insertion handelt, oder ob mit einer Insertion auf C8 eine Duplizierung des Locus verbunden ist, läßt sich aufgrund der Ergebnisse nicht definitiv sagen. Hierfür sollten zur Berechnung der genetischen Karte diejenigen Loci identifiziert und

unberücksichtigt bleiben, die fälschlich als ein Locus ausgewertet wurden, aber tatsächlich zwei oder mehr Loci repräsentieren. Solche „Phantomloci“ können gravierende Folgen für die Berechnung genetischer Karten haben und sogar Kopplung ganzer Markergruppen vortäuschen. Sharpe et al. (1995) beschreiben derartige Pseudokopplung in *B. napus*. Auch in dem Material der vorliegenden Arbeit konnte in der Vergangenheit Pseudokopplung beobachtet werden. Segmente der Kopplungsgruppe A9 der Population YE2-DH wurden früher fälschlicherweise als N18 (=C8) annotiert (Badani et al. 2006). Hier wurde auch der Haupt-QTL für Samenfarbe kartiert. Eine höhere Markerzahl und -dichte sowie eine stringendere Auswahl der kartierten Marker erlaubt nun eine sichere Unterscheidung der Loci in die zwei Kopplungsgruppen A9 (N9) und C8 (N18).

In Ermangelung einer genetischen wie physischen Referenzkarte für *B. napus* diene die veröffentlichte Genomsequenz von *B. rapa* als Referenz (<http://www.brassica-rapa.org>), um die Lokalisation der entwickelten genetischen Marker aus *B. rapa*-BAC-Sequenzen in den hier erstellten genetischen Karten für *B. napus* auf Plausibilität zu überprüfen.

Chromosomale Umstrukturierungen von Kandidatengen-Regionen und ihre phänotypischen Auswirkungen

Chromosomenmutationen können einen erheblichen phänotypischen Effekt und damit auch Einfluß auf die Berechnung genetischer Karten haben. So beschrieben Schranz und Osborn (2000, 2004) in Raps-Resynthesen umfangreiche *de novo* Variationen infolge der Allopolyploidisierung. Polyploidisierung ist ein anerkannt bedeutender Mechanismus in der Evolution und wahrscheinlich verantwortlich für die Entstehung von 2-4% neu entstandener Spezies (Wendel 2000). Rekombination findet während der Meiose primär zwischen homologen Chromosomen statt. In Allopolyploiden können aber auch Chromosomenpaarungen zwischen homöologen Chromosomen auftreten und ermöglichen so einen homöologen Gentransfer (Gaeta und Pires 2010). Darunter versteht man den Genaustausch zwischen homöologen Chromosomen der verschiedenen Genome einer allopolyploiden Pflanze bei hoher Syntänie, also einer stark konservierte Abfolgen genetischer Loci. Im Rapsgenom weisen beispielsweise die homöologen Kopplungsgruppen A1/C1, A2/C2 und A3/C3 eine ausgeprägte Syntänie auf, sowie in komplexerer Weise A4/C4/A5/C5/A6/C7/A7 und A8/C8/A10/C9/A9/C8 (Parkin et al. 2003). Udall et al. (2005) beschrieben umfangreiche chromosomale Umstrukturierungen in *B. napus*, u. a. der Kopplungsgruppen A9/C8, und unterschieden dabei verschiedene Mechanismen: homöologe reziproke Transpositionen (HRTs), homöologe nicht-reziproke Transpositionen (HNRTs) *de novo* und bereits

bestehende HNRTs. In natürlichen und resynthetisierten *B. napus*-Allopolyploiden ist homöologe Rekombination vielfach die Ursache für genetische Veränderungen und damit für die Entstehung neuer Phänotypen (Parkin et al. 1995, Sharpe et al. 1995; Udall et al. 2005, Mun et al. 2009). Dies dürfte auch auf 1012-98, einen Elter der YE2-DH-Population, zutreffen. Dafür sprechen mehrere Argumente: In der Population YE2-DH kartieren einige Markerallele entgegen der Erwartung auf C8, was in der Halbgeschwisterpopulation ExV8-DH nicht der Fall ist. Diese Kartierungsbesonderheiten in YE2-DH deuten auf homöologe Loci hin, die aufgrund der vorliegenden Syntänie so ähnlich sind, daß sie mit den gleichen Primern detektiert wurden. Die Primer wären dann nicht wie beabsichtigt lokusspezifisch. Es ist aber auch möglich, daß tatsächlich A-Loci aufgrund genetischer Umordnungen nun physisch auf C8 lokalisiert sind und folgerichtig dort auch kartieren. Beispiele aus der Literatur belegen anhand so unterschiedlicher Merkmale wie Ertrag (Osborn et al. 2003), Blühzeitpunkt (Schrantz und Osborn 2000, Pires et al. 2004, Gaeta et al. 2007), Entwicklungsparametern und morphologischen Merkmalen (Schrantz und Osborn 2004, Gaeta et al. 2007), daß phänotypische Diversität aufgrund von homöologen Transpositionen in Raps-Resynthesen generiert wird.

Ist auch die phänotypische Variabilität der Samenfarbe und vor allem der Samenfasergehalte der in dieser Arbeit untersuchten Populationen auf chromosomale Umordnungen zurückzuführen? Die Sequenzdaten der beiden aus QTL-Analysen identifizierten Kandidatengene für ADL erlauben diesbezüglich einige Hypothesen, die im Folgenden hergeleitet werden.

Das Gen *Cinnamylalkoholdehydrogenase* (*BnCAD2/CAD3*)

Die 454-Amplicon-Sequenzdaten für das Kandidatengen *Cinnamylalkoholdehydrogenase* (CAD) in den drei Eltern lassen auf mindestens zwei sehr unterschiedliche Mutationsereignisse schließen. So handelt es sich bei dem Genlocus offenbar um das Ergebnis einer Fusion zweier Homologe, die in *A. thaliana* als *AtCAD2* (At2g21730) und *AtCAD3* (At2g21890) vorliegen. In *B. rapa* dagegen kann in dem entsprechenden konservierten Block nur ein einziges CAD-Gen gefunden werden, welches jedoch offenbar chimärisch ist, d.h. eine sehr hohe Ähnlichkeit sowohl mit *AtCAD2* als auch mit *AtCAD3* aufweist. In *B. napus* liegt ebenfalls dieses fusionierte chimäre Gen mit sehr hoher Sequenzidentität zu *B. rapa* vor. Offenbar ist in der Evolution der *Brassica*-Spezies durch Rekombination ein neuartiges CAD-Gen entstanden. Dieses Gen wird in dieser Arbeit daher mit *CAD2/CAD3* bezeichnet. In Arabidopsis machen duplizierte, nebeneinander liegende Gene, sogenannte Tandem-Duplikationen, etwa 17% der rund 25.000 Gene aus (The Arabidopsis Genome Initiative, 2000). Ein solcher Tandembereich mit hoher

Sequenzübereinstimmung ist jedoch aufgrund seiner höheren Wahrscheinlichkeit der (nicht-reziproken) Rekombination anfällig für eine sich manifestierende Gendeletion, da ein letaler Phänotyp mit deutlich geringerer Wahrscheinlichkeit verursacht wird, wenn ein redundantes Gen verloren geht (Mun et al. 2009). Bei Deletionen von Kopien zuvor duplizierter Gene handelt es sich um einen Evolutionsmechanismus, der nach Polyploidisierung ggf. zu einer Stabilisierung des neuen Genoms führt und die Genomgröße wieder reduziert. Dieses Phänomen wurde von Rana et al. (2004) für *B. napus* und seine diploiden Vorfahren nachgewiesen.

Eine weitere Umordnung betrifft das Exon 3. Im Vergleich zu dem *B. rapa* *CAD2/CAD3*-Gen (Bra031216) konnte in der vorliegenden Arbeit eine intragenische Chimärbildung gezeigt werden, die wahrscheinlich auf homöologe Rekombination der jeweiligen Genkopien aus dem A-Genom und dem C-Genom zurückgeht. Die physiologische Funktion des Gens bzw. Gentandems in Arabidopsis ist bislang noch nicht geklärt und eine Beteiligung an der Ligninbiosynthese kann bislang nur vermutet werden. *AtCAD2* und *AtCAD3* weisen ähnliche, wenn auch nicht identische Substratspektren auf: *AtCAD2* setzt bevorzugt Caffeylaldehyd um, während *AtCAD3* effektiver *p*-Coumarylaldehyd und 5-Hydroxyconiferylaldehyd nutzt (Kim et al. 2004). Es ist anzunehmen, daß die Substratspezifität der neuartigen chimären Form derjenigen der beiden ursprünglichen CAD-Gene ähnlich ist, sofern sie überhaupt funktional ist. Dies bleibt jedoch zu überprüfen

Der Verlust eines Gens aus dem *CAD2-CAD3*-Tandem, wie es in Arabidopsis vorliegt, ist offenbar für die *Brassica*-Spezies nicht von Nachteil, da keinerlei Beeinträchtigung in der Bereitstellung von Monolignolen in den Brassicaceen im Vergleich zu Arabidopsis beobachtet werden konnte. Bonawitz und Chapple (2010) führen in ihrer Zusammenfassung zur Ligninbiosynthese aus, daß Lignin als ein plastisches Polymer gelten kann, da verschiedene Monolignole einander substituieren können und sich somit alternative Synthesewege für Lignin darstellen. Dies ist für die terrestrische Pflanze durchaus sinnvoll, da Lignin essentiell für aufrechtes Wachstum, Wassertransport und Streßreaktionen ist. Bildet sich nun ein neuartiges fusioniertes Gen, welches ein funktionales Enzym kodiert, kann dies für die Pflanze zunächst bedeuten, daß der Phänotyp aufrechterhalten bleibt, solange dieses Gen nicht mutiert. Tritt jedoch eine Störung ein, ob genetisch oder regulatorisch, kann dieser Flaschenhals nicht kompensiert werden und es kommt zu einer entsprechenden phänotypischen Ausprägung.

Im Fall von *CAD2/CAD3* würde dies bedeuten, daß bei fehlender Kompensation Cinnamylaldehyde nicht zu ihren korrespondierenden Alkoholen reduziert werden und sich damit die Zusammensetzung und Struktur der Lignane bzw. des Lignins als Endprodukt des Phenylpropanoidweges ändert. Letztendlich kann dies auch zu einer veränderten Menge der

akkumulierten Lignane und zu geänderten physikalischen Eigenschaften führen, beispielsweise der Stabilität und Hydrophobizität der lignifizierten Gewebe.

Das Gen *Cinnamoyl-CoA-Reduktase (BnCCR1)*

Cinnamoyl-CoA-Reduktase (CCR) katalysiert den Lignin-Syntheseschritt vor CAD und liefert somit das Substrat für CAD.

Chabannes et al. (2001) generierten *cad/ccr*-Doppelmutanten in Tabak und stellten fest, daß die Mutanten trotz einer Gesamtreduktion des (vaskulären) Lignins um 50% eine normale Entwicklung und Morphologie aufweisen. Möglicherweise liegt eine auf natürlichem Weg entstandene, ähnliche genetische Konstitution in den untersuchten Genotypen 1012-98 (geringer ADL-Gehalt) und V8 (mittlerer ADL-Gehalt) vor.

In den hier generierten Sequenzdaten fällt ähnlich wie bei *BnCAD2/CAD3* auf, daß die Anzahl der vorliegenden Genkopien nicht zweifelsfrei bestimmt werden kann. So zeigt Express 617, übereinstimmend mit der Situation für *BnCAD2/CAD3*, zwei Kopien von Exon 1 im A-Genom und eine Kopie im C-Genom. In 1012-98 hingegen kann nur die A-Genomkopie des Exons 1 nachgewiesen werden. Ebenfalls auffällig sind genetische Umordnungen in Exon 4 der beiden schwarzsamigen Genotypen, die zu chimären Sequenzen mit A- und C-Genomursprung führten. Eine wahrscheinliche Erklärung hierfür ist, daß es sich um das Ergebnis homöologer nicht-reziproker Translokationen handelt, bei denen ein Gensegment deletiert und durch sein dupliziertes Homöolog ersetzt wurde.

Pires et al. (2004) beschreiben am Beispiel des *FLOWERING LOCUS C (BnFLC)* in *B. napus*, daß eine Veränderung der Genkopienanzahl zu einem deutlichen phänotypischen Effekt führen kann. In jenem Fall blieb durch die Translokation und damit Duplizierung des kompletten Genlocus' der frühblühende Phänotyp erhalten, während umgekehrt eine Translokation des Chromosomenbereiches mit dem fehlenden *FLOWERING LOCUS* zur Ausbildung eines spät blühenden Phänotyps führte. Analog zu diesem Beispiel kann angenommen werden, daß die Kopienzahl der *BnCAD2/CAD3*- und *BnCCR1*-Loci in den hier untersuchten Genotypen ebenfalls die beschriebene phänotypische Variation bedingt.

Aminosäuresubstitutionen in *BnCAD2/CAD3* und *BnCCR1*

Bemerkenswert ist das verbreitete Auftreten von Aminosäuresubstitutionen in allen Translationsprodukten der Exons beider Kandidatengene. Einige Substitutionen stellen

wahrscheinlich Modifikationen dar, die *B. napus* von seinen Vorfahren unterscheiden und keine sortenspezifischen Veränderung sind. Als Beispiel hierfür sei der Austausch von Leucin (L) gegen Phenylalanin (F) an Position 151 in der von *BnCCR1* kodierten Aminosäuresequenz genannt. Andere Aminosäuresubstitutionen dagegen treten sehr spezifisch in einzelnen Genotypen auf und eignen sich dadurch als potentielle genetische Marker für die Untersuchung der phänotypischen Variabilität zwischen den Genotypen. Wichtig ist hier auch die Frage, ob sich die Aminosäuresubstitution phänotypisch manifestiert. Ohne Isolation des Enzyms und seine Aktivitätsbestimmung kann diese Frage nicht geklärt werden. Sicher ist jedoch, daß sich eine Ausprägung der veränderten Peptidsequenz auf die Enzymaktivität insbesondere dann ergeben wird, wenn Aminosäuresubstitutionen in oder nahe von sogenannten aktiven Stellen (*active sites*), also beispielsweise in den Substrat- oder Cofaktorbindungsstellen auftreten. Der Austausch einer einzelnen Aminosäure muß nicht zwangsläufig zu einer Konformationsänderung führen. Die Möglichkeit dafür ist jedoch gegeben, da nicht nur die Faltung des Peptids an der Stelle der jeweiligen Aminosäure betroffen ist, sondern auch die Wechselwirkung mit entfernteren Aminosäuren, die durch Faltung in die Nähe des Substitutionsortes gelangen. Auch die Bindung von Wasser kann beeinflußt sein. Inwieweit solche Änderungen das Substratspektrum oder die Reaktionskinetik ändern, kann von der Peptidsequenz allein nicht geschlossen werden.

Chromosomenmutationen und ihre Bedeutung für die Züchtung

Wie zuvor bereits erwähnt, kann durch chromosomale Umstrukturierungen erhebliche genetische und phänotypische Variation entstehen. Haben zum Beispiel homöologe Allele additive oder unterschiedlich starke Effekte auf ein Merkmal und wird ein Allel durch seine homöologe Form ersetzt, kommt es zu einem Dosiseffekt oder sogar zu einem Ausschalten des betreffenden Gens und damit zu einem starken phänotypischen Effekt (Gaeta et al. 2007). Umgekehrt betrachtet kann durch eben diesen Vorgang genetische Vielfalt verloren gehen, z. B. wenn durch nicht-reziproke Translokation (Insertion/Deletion)-die Unterschiede zwischen homöologen Loci verloren gehen. Translokationen sind dabei nicht als einmaliges Ereignis zu sehen, sondern als Teil eines dynamischen Prozesses. Viele *de novo* HNRTs wurden in Regionen beobachtet, die in einem Elter bereits HNRTs tragen (Udall et al. 2005). HNRTs, die zunächst nur kurze Segmente umfassen, können sich in weiteren meiotischen Teilungen weiter ausdehnen. Dabei startet die Chromosomenpaarung wahrscheinlich zwischen bereits bestehenden HNRTs und ihren ursprünglichen Homologen und setzt sich dann von den gepaarten homologen Segmenten in homöologe Regionen der beiden Chromatiden fort. Es entstehen dabei multivalente Chromosomenpaarungen, in denen

Rekombinationen stattfinden und durch die weitere *de novo* NRTs entstehen können (Attia und Röbbelen 1986, Udall et al. 2005). Sogar vollständige homöologe Chromatiden können so ausgetauscht werden. Dies bedeutet, daß bereits bestehende HNRTs die Assoziation von homöologen Chromosomen begünstigen, und damit letztlich eine Abweichung von disomischer Vererbung fördern. So ist es möglich, daß *B. napus* in Regionen mit starker homöologer Assoziation langfristig immer homologer wird und daß die Nachkommenschaft von Kreuzungen mit denselben homogenen Regionen in einer Weise aufspaltet, wie sie eher für Autopolyploide als für Allopolyploide typisch ist (Udall et al. 2005) und Neukombination oder Heterosis dann nicht mehr zu erwarten ist. Der Ertrag als heterotisches Merkmal ist jedoch das wichtigste Züchtungskriterium und daher besonders sensitiv in Bezug auf Genomhomogenisierung. Im Beispiel der Untersuchungen von Osborn et al. (2003) tragen Linien mit nicht-parenteralen Chromosomenkonfigurationen entweder nur *B. oleracea*- oder nur *B. rapa*-Allele und werden mit einer Ertragsreduktion assoziiert. Diese Linien verhalten sich dann analog zu partiellen Inzuchtlinien, verbunden mit einem reduziertem Heterosiseffekt, d.h. reduziertem Ertrag. Ein solches Phänomen kann auch auftreten, wenn auf Elite-Allele selektiert wird und diese in transponierten Segmenten liegen oder mit ihnen gekoppelt sind. Dann kann die betreffende chromosomale Umorganisation durch fortgesetzte Selektion erhalten bleiben und mit ihr die gekoppelten günstigen, aber gegebenenfalls auch ungünstigen phänotypischen Auswirkungen (Udall et al. 2005).

Homöologe Translokationen treten vor allem in frühen Stadien der Artbildung in der Folge von Polyploidisierung auf. In natürlichen Rapslinien spielen homöologe Translokationen jedoch eine untergeordnete Rolle, da sich das Genom bereits stabilisiert hat. In der Meiose verhalten sich Allopolyploide dann weitgehend wie Diploide, was darauf hinweist, daß die Chromosomenpaarung präzise kontrolliert und die homöologe Chromosomenpaarung gehemmt ist. Dies geschieht einerseits durch strukturelle Unterschiede der Homöologe und ihrer dadurch reduzierten Affinität, andererseits durch genetische Kontrollmechanismen zur Unterscheidung differenzierter Chromosomensätze (Jenczeswki et al. 2003). Hauptverantwortlich für die Chromosomenpaarung in Raps ist das Majorgen *PrBn* (*Pairing regulator in B. napus*) (Jenczeswki et al. 2003, Nicolas et al. 2009). Es gibt demnach Genotypen, die je nach ihrer genetischen Konstitution am *PrBn*-Locus mehr oder weniger ausgeprägt autosyndetische und allosyndetische Chromosomenpaarungen in der Meiose zeigen, also die Bildung von Bivalenten innerhalb des A- oder C-Genoms und zwischen den beiden Genomen. In Raps-Resynthesen fehlt dieser Locus offenbar mitunter, so daß hier homöologe Chromosomenpaarungen und Rekombinationen verstärkt auftreten. Genauere Kenntnisse über die Wirkungsweise des *PrBn*-Locus können einmal zu seiner Nutzung in der Züchtung führen, wenn es darum geht, genetische Variabilität zu erzeugen oder auch zu hemmen. In Weizen konnte beispielsweise durch Nutzung des *Ph1*-Locus (*pairing*

homeologous) - dem wichtigsten Faktor zur Kontrolle der homologen Chromosomenpaarung (Riley 1958) - gezielt Gelbrostresistenz durch induzierte homöologe Rekombination eingebracht werden (Riley et al. 1968). Dies zeigt, daß die markergestützte Selektion auch in diesem Zusammenhang ein hilfreicher Ansatz ist, um umfangreiches Zuchtmaterial bezüglich erwünschter Allele und deren Stabilität zu evaluieren. Neue Technologien für ein Sichten des Materials im Hochdurchsatz sind bereits auf dem Markt und mit zunehmender Weiterentwicklung der korrespondierenden Bioinformatik ein mächtiges Instrument zukünftiger und auch bereits angewandter Züchtung.

Neue DNA-Sequenzierungstechniken in Forschung und Züchtung

Die rasante Entwicklung der DNA-Sequenzierungstechnologien, angefangen von der Methode nach Sanger über die „*Next-Generation-Sequencing*“ Techniken zum aktuellen „*Third-Generation-Sequencing*“ ermöglicht heute die Sequenzierung ganzer Genome in kurzer Zeit und zu relativ günstigen Preisen (Edwards und Batley 2010). Sanger-Sequenzierung von überlappenden BACs ist sehr arbeits-, zeit- und kostenaufwendig, daher war noch vor 10 Jahren die Sequenzierung kleiner Genome, wie jene von Modellorganismen wie *Arabidopsis* und Reis vorrangiges Ziel von Genomsequenzierungsprojekten. Heute ermöglichen Hochdurchsatztechnologien die *de novo* Sequenzierung von Nutzpflanzen-genomen und auch die Resequenzierung von einzelnen Genotypen für Studien zur vergleichenden Genomik. Die Zielsetzung kann dabei sehr unterschiedlich sein: Die Genome wichtiger Nutzpflanzen können aufgeklärt und agronomisch wichtige Gene identifiziert werden. Auch können effiziente genomweite molekulare Marker (*single nucleotide polymorphisms*, SNP) zur Erforschung sortenspezifischer Variationen oder von Evolutionsvorgängen entwickelt werden (Feuillet et al. 2011).

Grundsätzlich ist für alle Sequenzierungsprojekte die Verfügbarkeit von Referenzgenomen von Vorteil, da diese bei der Assemblierung von Einzelsequenzen (sogenannten *reads*) als Matrize dienen können. Ist keine Referenzsequenz verfügbar, fällt die korrekte Assemblierung einzelner *reads* sehr schwer, insbesondere wenn es sich um große komplexe Genome mit einem hohen Anteil repetitiver DNA und von Transposons handelt. Auch polyploide - und insbesondere allopolyploide - Genome stellen mit ihren homologen bzw. homöologen Sequenzen besondere Herausforderungen an eine korrekte Assemblierung und Annotation. Darüber hinaus sind homozygote Genotypen heterozygoten vorzuziehen, sofern dies möglich ist. So stellte z. B. die Sequenzierung der Weinrebe (*Vitis vinifera*) wegen ihrer ausgeprägten Heterozygotie besondere Anforderungen an das Design der Sequenzierungsstrategie und der Datenauswertung (Velasco et al. 2007). Eine höherwertige

Referenzsequenz für die Weinrebe wurde daher für einen relativ homozygoten Genotyp erstellt (Jaillon et al. 2007).

In der vorliegenden Arbeit wurden *B. napus*-Sequenzen mit einer Referenzsequenz von *B. rapa* verglichen. Für *B. olearacea* stand keine genomische Referenzsequenz zur Verfügung, sondern lediglich annotierte Gensequenzen. Verschiedene *B. napus*-Sequenzen wurden zwar in den letzten Jahren durch Konsortien in Frankreich, Kanada sowie China und auch unter der Führung einer Privatfirma erstellt. Diese waren aber zum Zeitpunkt der Datenauswertung für diese Arbeit nicht öffentlich zugänglich. Die in der vorliegenden Arbeit angewendete Sequenzierungsstrategie hat daher zur Generierung einer Vielzahl von Einzelsequenzen geführt, die überwiegend dem A- oder C-Genom - entsprechend der bislang verfügbaren Teilsequenzen in öffentlichen Datenbanken - zugeordnet werden konnten. Für *BnCAD2/CAD3* konnten aber auch in einzelnen translatierten Contigs umfangreichere Aminosäuresubstitutionen beobachtet werden, die wahrscheinlich auf die Entstehung neuartiger Allele zurückgehen. Der Vergleich mit einer *B. napus*-Referenz würde hier sicherlich zur Aufklärung beitragen.

Die angewendete Methode der 454-Sequenzierung erlaubte die Ermittlung von maximal 500bp langen Sequenzstücken. Aus praktischen Erwägungen wurden daher nur Exonregionen aus genomischer DNA amplifiziert und sequenziert. Im Ergebnis konnten diese Amplikons nur exonweise zu Contigs assembliert werden, nicht jedoch so, daß die gesamte kodierende Gensequenz lückenlos abgedeckt wurde. Trotzdem konnten zahlreiche Informationen aus den Sequenzen gewonnen werden, wie die Beurteilung, ob beide erwarteten homöologen Kopien (A- und C-Genom) in den untersuchten Genotypen vorliegen. Überdies war es möglich, anhand von detektierten SNPs gen- und kopiespezifische Marker für *BnCAD2/CAD3* und *BnCCR1* in YE2-DH zu entwickeln. Ob diese Marker z.B. als Selektionsmarker geeignet sind, muß anhand von umfangreichem Material validiert werden. Wenigstens aber bekräftigen die Marker die Hypothese, daß *BnCCR1* und *BnCAD2/CAD3* an der Ausprägung des ADL-QTL beteiligt sind.

Selektionsmarker zur Anwendung in markergestützter Züchtung ermöglichen eine frühzeitige Selektion auf Genotypen mit günstigen Allelen und somit eine genotypbasierte Eingrenzung des Zuchtmaterials, was insbesondere bei aufwendigen und teuren Phänotypisierungen von Bedeutung ist. Dadurch kann die weitere Bearbeitung zur Verbesserung des Materials (z.B. hinsichtlich einer Ertragssteigerung) erheblich effektiver gestaltet werden. Neben der Entwicklung von Selektionsmarkern bieten Sequenzinformationen weitere Nutzungsmöglichkeiten. Eine genomische Referenzsequenz erlaubt z. B. die Klonierung von agronomisch wichtigen QTL und die Aufklärung ihres Gengehaltes. Für Reis wurden z. B. auf diese Weise QTL für Blühzeitpunkt, Krankheitsresistenz, Pflanzenarchitektur und

abiotischen Streß identifiziert (Yamamoto et al. 2009). Dichte genetische Karten mit sequenzbasierten Markern erlauben ferner die Integration genetischer und physischer Karten und damit auch die Übertragbarkeit von Informationen aus Modellorganismen und Kulturpflanzen (Edwards und Batley 2010). Modellgenome können Informationen über orthologe Kandidatengene in Nutzpflanzen liefern, vorausgesetzt, daß viele grundlegende Biosynthese- und Signalwege in Pflanzen gleich verlaufen. Das Beispiel der kartengestützten Klonierung von Blühzeitpunktgenen in Reis zeigt jedoch, daß dies nicht zwangsläufig der Fall sein muß. Hier weisen Reis und Arabidopsis zwar die gleichen regulatorischen Mechanismen auf, die differentielle Regulation beruht aber auf reversen Funktionen von zentralen Schlüsselregulatoren (Hayama et al. 2003). Folglich liefert der Sequenzvergleich zwar wertvolle Hinweise, kann aber nur der erste Schritt in der Charakterisierung von Genfunktionen sein. Funktionelle Analysen in den eigentlichen Nutzpflanzen bleiben daher essentiell, um artspezifische Besonderheiten zu identifizieren und nutzbar zu machen. In polyploiden Pflanzen ist die Genregulation zudem komplexer als in diploiden (Modell-) Pflanzen und führt möglicherweise zu einer Maskierung orthologer Zusammenhänge zwischen Modell und Nutzpflanze. Um also agronomisch wichtige Merkmale in den wichtigsten Nahrungs-, Futter- und Faserpflanzen zu verstehen, muß die gesamte Bandbreite genomischer Ressourcen ermittelt und genutzt werden. Neue Sequenzierungstechniken ermöglichen eine rasche Generierung umfangreicher Einzelsequenzen im Hochdurchsatzverfahren. Doch Einzelsequenzen (*reads* und *contigs*) zu einer durchgehenden Sequenz zu assemblieren, welche nach Möglichkeit ein komplettes Chromosom abdeckt, ist derzeit noch durch die maximal möglichen Sequenzlängen limitiert. Feuillet et al. (2011) sehen hierin eine Ursache für den Trend, daß heutzutage vorwiegend vorläufige Sequenzen (*draft sequences*) von möglichst vielen Spezies generiert werden und der Fokus nicht darauf liegt, umfassende und abgesicherte Genomsequenzen zu erstellen, wie es für Arabidopsis und Reis der Fall war. Vorläufige Sequenzen bieten mitunter nicht die nötige Präzision. Das Beispiel vergleichender Analysen zwischen Mais-Inzuchtlinien (Brunner et al. 2005) zeigt, daß eine einzelne Genomsequenz das Genom der Art nicht vollständig abbildet, sondern neben einem artspezifischen Kern an genomischer Sequenz auch einen Teil sortenspezifischer Sequenz enthält. Morgante et al. (2007) schlagen für dieses Phänomen den Begriff „Pangenom“ vor.

Methoden des *Third generation sequencing* lassen erwarten, daß mit größeren Sequenzlängen und ohne das Risiko fehlerbehafteter Amplifikation vor der Sequenzierung die Assemblierung komplexer Genome und Genregionen möglich wird (Feuillet et al. 2011).

Schlußbetrachtung

Züchtung von Elitematerial und neuen Sorten fußt häufig auf Kreuzungen zwischen leistungsstarken Elitelinien oder bereits registrierten Sorten. Ein Dilemma dieser Vorgehensweise, verbunden mit der gezielten Auslese auf u. U. wenige maßgebliche Merkmale führt fast zwangsläufig zu einer stetigen Einengung des genutzten Genpools. Bei allen Nachteilen, die sich aus der Verwendung genetisch entfernter Ressourcen ergeben, kann sie doch langfristig eine Stagnation des Züchtungsfortschritts verhindern. Neben dem Hauptzüchtungsmerkmal Ertrag sind Krankheitsresistenzen und Toleranzen gegenüber ungünstigen Klima- und Umweltbedingungen, aber auch die Anforderungen an die Qualität einem Wandel unterworfen, der nicht auf lange Sicht abschätzbar ist. Daher spielt neben der Neuzüchtung auch die Erhaltung der genetischen Vielfalt eine große Rolle. Raps als junge allopolyploide Pflanze in nahezu weltweitem Anbau bietet hierzu alle Möglichkeiten, von einem breiten internationalen Genpool bis zur Resynthese aus den diploiden Vorfahren.

Die Ergebnisse dieser Arbeit zeigen am Beispiel des aktuellen und wirtschaftlich interessanten Qualitätsmerkmals Fasergehalt im Rapskorn und -schrot molekulargenetische Wege, die Grundlagen für die Züchtung zu erweitern und zu verbessern. Angeregt wurde die Arbeit einerseits durch die rasche Entwicklung und Verfügbarkeit neuer technologischer Möglichkeiten und damit einhergehend der Verfügbarkeit von molekularen Markerdaten und Sequenzinformationen. Andererseits wurden komplexe Merkmale mit multigener Vererbung und plastischer Biosynthese bearbeitet. Das Objekt Raps stellt diesbezüglich mit seinem komplexen Genom eine zusätzliche Herausforderung dar. Für weiterführende Studien können vermutlich weitere methodische Fortschritte hilfreich werden: Zum einen wird ein *B. napus*-Referenzgenom, sobald es öffentlich verfügbar ist, viele molekulargenetischen Fragestellungen, wie die Bearbeitung allelischer Diversität, vereinfachen. Zum anderen werden metabolische Studien und Metabolom-Analysen eine genauere Definition der Qualitätsparameter ermöglichen und damit die Suche und Bereitstellung wichtiger Gene für die Züchtung erheblich vereinfachen.

Zusammenfassung

Die vorliegende Arbeit verfolgte das Ziel, Gene in der Kulturpflanze Raps (*Brassica napus*) zu identifizieren, die einen Einfluß auf den Fasergehalt im Schrot und auf die Samenfarbe haben. Kenntnisse über solche Gene und ihre Sequenzen ermöglichen die Entwicklung molekularer Marker, die in der markergestützten Selektion zur Züchtung hochwertiger Elite-Rapssorten mit verbesserter Schrotqualität eingesetzt werden können. Derartige Rapssorten bieten einen ökonomischen Mehrwert hinsichtlich der Verwertbarkeit des Nebenproduktes Rapsschrot als Futtermittel, da die bisher empfohlenen Anteile an den Futterrationen aufgrund einer besseren Verdaulichkeit insbesondere in der Geflügel- und Schweinefütterung erhöht werden können.

Basis für diese Untersuchungen waren Ergebnisse, nach denen auf dem *B. napus* Chromosom A9 *quantitative trait loci* (QTL) für die Merkmale Samenfarbe und Samenfasergehalt identifiziert wurden. Die Arbeit wurde in die folgenden aufeinander aufbauenden Arbeitspakete gegliedert:

1. Feinkartierung der QTL: Auf Basis einer vorliegenden genetischen Karte und der Verankerung der genetischen Marker auf den zur Verfügung stehenden Referenzgenomen von *Arabidopsis thaliana* und *Brassica rapa* wurden neue locusspezifische genetische Marker entwickelt. Mit diesen Markern wurde anschließend die Karte abgesättigt und der QTL-Bereich stärker eingegrenzt.
2. Sequenzierung des den QTL tragenden Chromosomenabschnitts: Mittels Hybridisierung einer *Bacterial-Artificial-Chromosome* (BAC) -Bibliothek wurden BAC-Klone der schwarzsamigen *B. napus*-Linie Express 617 identifiziert und isoliert. Ausgewählte Klone, die den QTL-Bereich abdecken, wurden sequenziert. Der potentielle Geninhalt konnte über eine öffentliche *B. rapa*-Datenbank ermittelt werden und führte so zur Identifizierung des Kandidatengens *Cinnamylalkoholdehydrogenase* (*Bra031216*) im QTL-Bereich.
3. Vergleichende Kartierung: Genetische Marker aus verschiedenen Kartierungspopulationen wurden vergleichend kartiert, was schließlich darauf hindeutete, daß es sich offenbar um den gleichen QTL-Bereich handelt, also folglich auch um dieselben Gene. Das Kandidatengen *Cinnamoyl-CoA-Reduktase* (*Bra026737*) wurde durch Vergleich der QTL-Region in der genetischen Karte in einer chinesischen Rapspopulation mit der physischen *B. rapa*-Karte von Chromosom A9 identifiziert.
4. Sequenzierung der Kandidatengene: Die beiden vorgenannten Kandidatengene wurden in den verwendeten Kartierungseltern 1012-98, Express 617, V8, P174 und GH06

sequenziert. Dazu wurden spezifische Primer für die Amplifizierung der Exons entwickelt.

5. Sequenzvergleich: Die Sequenzdaten aller Genotypen wurden gruppiert und die resultierenden Konsensus-Sequenzen gegen die jeweiligen Referenzsequenzen aus *B. rapa* und *B. oleracea* kartiert. Danach konnten die Sequenzen miteinander verglichen werden, um Unterschiede, die einen potentiellen phänotypischen Effekt haben könnten, zu identifizieren. Eine Reihe von *single nucleotide polymorphisms* (SNPs), Deletionen, Insertionen und resultierende Aminosäuresubstitutionen wurden hierbei identifiziert.

Im Ergebnis zeigen die Studien, dass die Gene *Cinnamylalkoholdehydrogenase* (*CAD*) und *Cinnamoyl-CoA-Reduktase* (*CCR*) gemeinsam für die Variation des ADL-Gehaltes in *B. napus*-Samen maßgeblich verantwortlich sind. Diese Gene kodieren für Enzyme, die die beiden letzten aufeinanderfolgenden Syntheseschritte für Monolignole katalysieren. In beiden Genen konnte eine hohe Sequenzvariation beobachtet werden.

Eine neuartige Genstruktur für *CAD* wurde sowohl in allen untersuchten Rapsgeotypen als auch in *B. rapa* identifiziert: Das Gen geht augenscheinlich auf die Fusion (Chimäre) zweier Kopien eines Gentandems zurück, wie es in *A. thaliana* vorliegt.

Der Genotyp 1012-98 mit niedrigem ADL-Gehalt besitzt keine intakte A-Kopie von *CAD*. Von *CCR* konnte eine chimäre Form aus den A- und C-Genom-Homöologen identifiziert werden. In GH06 wurde die C-Genom-Kopie des Gens *CCR* durch eine defekte Kopie des A-Genom-Locus' ersetzt. *CAD* liegt hier wahrscheinlich als eine chimäre Form aus den A- und C-Genom-Homöologen vor. V8, der Genotyp mittleren ADL-Gehalts trägt dagegen sowohl von *CAD* als auch von *CCR* je eine defekte Kopie und je eine möglicherweise nicht oder nur teilweise funktionale Kopie der Gene. Die hoch-ADL Genotypen Express 617 und P174 verfügen über intakte *CAD*- und *CCR*-Kopien.

Es wird aus dieser Studie also deutlich, daß ganz unterschiedliche Mutationen und ihre Kombinationen in den Kopien der beiden untersuchten Gene zu Variationen im Gehalt und vermutlich auch der Zusammensetzung des Parameters ADL in *B. napus*-Samen geführt haben. Dies zeigt, wie genetisch komplex die Lignin- bzw. Rohfaser-Fraktion in Rapssaaten determiniert ist. Aus den Erkenntnissen dieser Arbeit ergeben sich daher neue Ansätze für die Entwicklung nutzbarer molekularer Selektionsmarker.

Summary

This study aimed at identifying genes in the crop plant oilseed rape (*Brassica napus*, canola) which have influence on the fibre content and colour of the seeds. Knowledge about such genes and their sequences helps to develop molecular markers that can be used in marker-assisted selection (MAS) for breeding of new elite varieties with improved meal quality. Such varieties offer added economic value, because oilseed rape ratios in animal feeding, especially for poultry and pigs, can be increased due to improved digestibility of the meal.

The work was based on results identifying quantitative trait loci (QTL) on *Brassica napus* chromosome A9 for relevant seed colour and fibre traits. The study was divided into the following consecutive sub-projects:

1. Fine mapping: Based on an existing genetic map and anchorage of molecular markers to the available reference genomes of *Arabidopsis thaliana* and *Brassica rapa* new locus-specific molecular markers were developed. These markers were used to saturate the genetic map and to narrow down the QTL region on chromosome A9.
2. Sequencing of the chromosomal area containing the QTL region: Bacterial artificial chromosome (BAC) library hybridisation was used to identify and isolate large-insert genomic clones from the black-seeded *B. napus* line Express 617. Clones spanning the QTL region were sequenced. The potential gene content was identified using a publicly available *B. rapa* database. This led to identification of the candidate gene *Cinnamyl Alcohol Dehydrogenase* (*Bra031216*) in the QTL region.
3. Comparative mapping: Genetic markers developed in different mapping populations were comparatively mapped, suggesting that the QTL region investigated and the underlying genes are the same in both studies. The candidate gene *Cinnamoyl CoA Reductase* (*Bra026737*) was identified by comparison of the QTL region in a genetic map from Chinese *B. napus* with the corresponding physical *B. rapa* map of chromosome A9.
4. Sequencing of the candidate genes: The aforementioned candidate genes were sequenced in the mapping parents 1012-98, Express 617, V8, P174 and GH06. Specific primers for amplification of the exon sequences were developed.
5. Sequence comparison: Sequence data of all genotypes were grouped. The resulting consensus-sequences were mapped against the reference sequences of *B. rapa* and *B. oleracea*. Subsequently, the sequences were compared in order to identify differences with a potential phenotypic effect. A number of single nucleotide

polymorphisms (SNPs), deletions, insertions and resulting amino acid exchanges were thus identified.

In summary, the studies reveal that the genes *Cinnamyl Alcohol Dehydrogenase* (*CAD*) and *Cinnamoyl CoA Reductase* (*CCR*) both influence the variation of ADL content and composition in the *B. napus* seeds. These genes encode enzymes, which catalyze the two final steps in the monolignol biosynthesis. In both genes a high variability of the sequences was observed.

A novel gene structure for *CAD* was identified in all investigated *B. napus* genotypes and in *B. rapa*: The gene apparently originates from a fusion (chimera) of two gene copies belonging to a tandem repeat structure as it occurs in *A. thaliana*.

The genotype 1012-98, with low ADL content, completely lacks an A genome copy of *CAD*. A chimeric copy formed from parts of the A and C genome homoeologues of *CCR* was identified in this genotype. In GH06 the C-genome copy of *CCR* is replaced by a defect copy of the A-genome locus. Apparently, *CAD* occurs as a chimeric structure of the A and C genome homeologues. On the other hand, the medium-ADL genotype V8 carries one defect and one non-functional or only partially functional gene copy of both *CAD* and *CCR*. The high ADL genotypes Express 617 and P174 contain intact *CAD* and *CCR* gene copies.

The study makes clear that not only different mutations, but also combinations thereof in the copies of the two investigated genes lead to variation of the content and presumably composition of ADL in *B. napus* seeds. The results underline the genetically complex determination of this particular trait in rapeseed. The results of the present study revealed new approaches to develop molecular markers suitable for marker-assisted selection.

Literatur

Abbadi, A. und G. Leckband, Rapeseed breeding for oil content, quality, and sustainability, *European Journal of Lipid Science and Technology* 2011, 130:1198-1206

Abrahams, S., G. J. Tanner, P. J. Larkin, A. R. Ashton, Identification and biochemical characterization of mutants in the proanthocyanidin pathway in *Arabidopsis*, *Plant Physiology* 2002, 130:561-576

Anjou, K., B. Lönnerdal, B. Uppström, P. Åman, Composition of seeds from some Brassica cultivars, *Swedish Journal of Agricultural Research* 1977, 7:169-178

Atanassova, R., N. Favet, F. Martz, B. Chabbert, M.-T. Tollier, B. Monties, B. Fritig, M. Legrand, Altered lignin composition in transgenic tobacco expressing O-methyltransferase sequences in sense and antisense orientation, *The Plant Journal* 1995, 8:465-477

Attia, T. und G. Röbbelen, Meiotic pairing in haploids and amphidiploids of spontaneous versus synthetic origin in rape, *Brassica napus* L., *Canadian Journal of Genetics and Cytology* 1986, 28:330-334

Badani, A. G., R. J. Snowdon, B. Wittkop, F. D. Lipsa, R. Baetzel, R. Horn, A. De Haro, R. Font et al., Colocalization of a partially dominant gene for yellow seed colour with a major QTL influencing acid detergent fibre (ADF) content in different crosses of oilseed rape (*Brassica napus*), *Genome* 2006, 49:1499-1509

Baetzel, R., Stabilität und Genetik von wertbestimmenden Inhaltsstoffen gelbsamiger Rapslinien (*Brassica napus*) als Basis für die Züchtung von qualitativ hochwertigen Körnerölraps-Sorten, Dissertation im Institut für Pflanzenbau und Pflanzenzüchtung I der Justus-Liebig-Universität Gießen, 2004

Baxter I. R., J. C. Young, G. Armstrong, N. Foster, N. Bogenschütz, T. Cordova, W. A. Peer, S. P. Hazen, A. S. Murphy, J. F. Harper, A plasma membrane H⁺ ATPase is required for the formation of proanthocyanidins in the seed coat endothelium of *Arabidopsis thaliana*, *Proceedings of the National Academy of Sciences* 2005, 102:2649-2654

Beeckman, T., R. De Rycke, R. Viane, D. Inzé, Histological study of seed coat development in *Arabidopsis thaliana*, *Journal of Plant Research* 2000, 113:139-148

Bell, J. M., Factors affecting the nutritional value of canola meal: A review, *Canadian Journal of Animal Science* 1993, 73:679-697

Boerjan, W., J. Ralph, M. Baucher, Lignin biosynthesis, *Annual Review of Plant Biology* 2003, 54:519-546

Bonawitz, N. D. und C. Chapple, The genetics of lignin biosynthesis: connecting genotype to phenotype, *Annual Review of Genetics* 2010, 44:337-363

Brassica rapa Genome Sequencing Project Consortium: X. Wang, H. Wang, J. Wang, R. Sun, J. Wu, S. Liu, Y. Bai, et al., The genome of the mesopolyploid crop species *Brassica rapa*, *Nature Genetics* 2011, 43:1035-1039

Brunner, S., K. Fengler, M. Morgante, S. Tingey, A. Rafalski, Evolution of DNA sequence nonhomologies among maize inbreds, *The Plant Cell* 2005, 17:343–360

Campbell, M. M. und R. R. Sederoff, Variation in lignin content and composition, *Plant Physiology* 1996, 110:3–13

Carneiro, M.O., C. Russ, M. G. Ross, S. Gabriel, C. Nusbaum, M. A. DePristo, Pacific biosciences sequencing technology for genotyping and variation discovery in human data, *BMC Genomics* 2012, 13:375

Caño-Delgado, A., S. Penfield, C. Smith, M. Catley, M. Bevan, Reduced cellulose synthesis invokes lignification and defense responses in *Arabidopsis thaliana*, *The Plant Journal* 2003, 34:351-362

Chabannes, M., A. Barakate, C. Lapierre, J. M. Marita, J. Ralph, M. Pean, S. Danoun, C. Halpin et al., Strong decrease in lignin content without significant alteration of plant development is induced by simultaneous down-regulation of cinnamoyl CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD) in tobacco plants, *The Plant Journal* 2001, 28:257-270

Chen, F. und R. A. Dixon, Lignin modification improves fermentable sugar yields for biofuel production, *Nature Biotechnology* 2007, 25:759 - 761

Debeaujon, I., A. J. M. Peeters, K. M. Léon-Klosterziel, M. Koornneef, The Transparent testa 12 gene of *Arabidopsis* encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium, *The Plant Cell* 2001, 13:853-871

Edwards, D., J. Batley, Plant genome sequencing: applications for crop improvement, *Plant Biotechnology Journal* 2010, 8:2–9

Feuillet, C., J. E. Leach, J. Rogers, P. S. Schnable, K. Eversole, Crop genome sequencing: lessons and rationales, *Trends in Plant Science* 2011, 16:77-88

Freudenberg, K., Lignin: its constitution and formation from p-hydroxycinnamyl alcohols, *Science* 1965, 148:595-600

Friedt, W. und R. J. Snowdon, Oilseed rape, in J. Vollmann und I. Rajcun (eds) *Oil Crops, Handbook of Plant Breeding* 4, Springer Science + Business Media, LLC 2009

Gaeta, R. T., J. C. Pires, F. Iniguez-Luy, E. Leon, T. Osborn, Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype, *The Plant Cell* 2007, 19:3403-3417

Gaeta, R. T., J. C. Pires, Homoeologous recombination in allopolyploids: the polyploid ratchet, *New Phytologist* 2010, 186:18-28

Goujon, T., V. Ferret, I. Mila, B. Pollet, K. Ruel, V. Burlat, J.-P. Joseleau, Y. Barrière, C. Lapierre, L. Jouanin, Down-regulation of the *AtCCR1* gene in *Arabidopsis thaliana*: effects on phenotype, lignins and cell wall degradability, *Planta* 2003 (a), 217:218–28

Goujon, T., R. Sibout, A. Eudes, J. MacKay, L. Jouanin, Genes involved in the biosynthesis of lignin precursors in *Arabidopsis thaliana*, *Plant Physiology and Biochemistry* 2003 (b) 41:677–687

Halloin, J. M., Localization and changes in catechin and tannins during development and ripening of cottonseed, *New Phytologist* 1982, 90:651-657

Halpin, C., M. E. Knight, G. A. Foxon, M. M. Campbell, A. M. Boudet, J. J. Boon, B. Chabbert, M.-T. Tollier, W. Schuch, Manipulation of lignin quality by downregulation of cinnamyl alcohol dehydrogenase, *The Plant Journal* 1994, 6:339–350

Haughn G. und A. Chaudhury, Genetic analysis of seed coat development in *Arabidopsis*, *Trends in Plant Science* 2005, 10:472-477

Hayama, R., S. Yokoi, S. Tamaki, M. Yano, K. Shimamoto, Adaptation of photoperiodic control pathways produces short-day flowering in rice, *Nature* 2003, 422:719–722

- Hüsken, A., A. Baumert, D. Strack, H. C. Becker, C. Möllers, C. Milkowski, Reduction of sinapate ester content in transgenic oilseed rape (*Brassica napus*) by dsRNAi-based suppression of BnSGT1 gene expression, *Molecular Breeding* 2005, 16:127–138
- Jaillon, O., J.-M. Aury, B. Noel, A. Policriti, C. Clepet, A. Casagrande, N. Choisne, S. Aubourg et al. (The French-Italian Public Consortium for Grapevine Genome Characterization), The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla, *Nature* 2007, 449:463-468
- Jenczewski, E., F. Eber, A. Grimaud, S. Huet, M. O. Lucas, H. Monod, A. M. Chevre, PrBn, a major gene controlling homeologous pairing in oilseed rape (*Brassica napus*) haploids, *Genetics* 2003, 164:645-653
- Kim, S.-J., M.-R. Kim, D. L. Bedgar, S. G. A. Moinuddin, C. L. Cardenas, L. B. Davin, C. Kang, N. G. Lewis, Functional reclassification of the putative cinnamyl alcohol dehydrogenase multigene family in *Arabidopsis*, *Proceedings of the National Academy of Sciences* 2004, 10:1455–1460
- Kirk, T. K. und R. L. Farrell, Enzymatic „combustion“: the microbial degradation of lignin, *Annual Review of Microbiology* 1987, 41:465-505
- Leflon, M., L. Grandont, F. Eber, V. Huteau, O. Coriton, L. Chelysheva, E. Jenczewski, A.-M. Chèvre, Crossovers get a boost in *Brassica* allotriploid and allotetraploid hybrids, *The Plant Cell* 2010, 22:2253–2264
- Leplé, J.-C., R. Dauwe, K. Morreel, V. Storme, C. Lapierre, B. Pollet, A. Naumann, K.-Y. Kang, Downregulation of Cinnamoyl-Coenzyme A Reductase in poplar: multiple-level phenotyping reveals effects on cell wall polymer metabolism and structure, *The Plant Cell* 2007, 19:3669–3691
- Liang, M., E. Davis, D. Gardner, X. Cai, Y. Wu, Involvement of AtLAC15 in lignin synthesis in seeds and in root elongation of *Arabidopsis*, *Planta* 2006, 224:1185–1196
- Lindermayer, H., W. Preißinger, G. Propstmeier, G. Schweinefütterung mit heimischen Eiweißfuttermitteln. LfL-Information, Bayerische Landesanstalt für Landwirtschaft (LfL), Freising-Weihenstephan, Deutschland, 2010
- Marles, M. A. S. und M. Gruber, Histochemical characterisation of unextractable seed coat pigments and quantification of extractable lignin in the Brassicaceae, *Journal of the Science of Food and Agriculture* 2004, 84:251-262

- Moïse, J. A., S. Han, L. Gudynaite-Savitch L., D. A. Johnson, B. L. A. Miki, Seed coats: structure, development, composition, and biotechnology, *In Vitro Cellular and Developmental Biology - Plant* 2005, 41:620-644
- Moore, J. K. und H.-J. G. Jung, Lignin and fiber digestion, *Journal of Range Management* 2001, 54:420-430
- Morgante, M., De Paoli, E., Radovic, S., Transposable elements and the plant pan-genomes, *Current Opinion in Plant Biology* 2007, 10:149–155
- Mun, J.-H., S.-J. Kwon, T.-J. Yang, Y.-J. Seol, M. Jin, J.-A. Kim, M.-H. Lim, J. S. Kim et al., Genome-wide comparative analysis of the *Brassica rapa* gene space reveals genome shrinkage and differential loss of duplicated genes after whole genome triplication, *Genome Biology* 2009, 10:R111
- Mühmel, H., F. Pudel, L. Alpmann, C. Müller, H. Aniol, T. Korrman, Rapeseed morphology parameters influencing oil pressing, *Proceedings 13th International Rapeseed Congress* 2011, Prag, Tschechische Republik, 1097-1100
- Nicolas, S. D., M. Leflon, H. Monod, F. Eber, O. Coriton, V. Huteau, A.-M. Chèvre, E. Jenczewski, Genetic regulation of meiotic cross-overs between related genomes in *Brassica napus* haploids and hybrids, *The Plant Cell* 2009, 21:373-385
- Osborn, T. C., D. V. Butrulle, A. G. Sharpe, K. J. Pickering, I. A. P. Parkin, J. S. Parker, D. J. Lydiate, Detection and effects of a homeologous reciprocal transposition in *Brassica napus*, *Genetics* 2003, 165:1569–1577
- Parkin, I. A. P., A. G. Sharpe, D. J. Keith, D. J. Lydiate, Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape), *Genome* 1995, 38:1122-1131
- Parkin, I. A. P., A. G. Sharpe, D. J. Lydiate, Patterns of genome duplication within the *Brassica napus* genome, *Genome* 2003, 46:291-303
- Patten, A. M., C. L. Cardenas, F. C. Cochrane, D. D. Laskar, D. L. Bedgar, L. B. Davin, N. G. Lewis, Reassessment of effects on lignification and vascular development in the *irx4* *Arabidopsis* mutant, *Phytochemistry* 2005, 66: 2092–2107
- Piquemal, J., C. Lapierre, K. Myton, A. O'Connell, W. Schuch, J. Grima-Pettenati, A.-M. Boudet, Down-regulation of Cinnamoyl-CoA Reductase induces significant changes of lignin profiles in transgenic tobacco plants, *The Plant Journal* 1998, 13:71–83

- Pourcel, L., J.-M. Routaboul, V. Cheynier, L. Lepiniec, I. Debeaujon, Flavonoid oxidation in plants: From biochemical properties to physiological functions, *Trends in Plant Science* 2006,12:29-36
- Pires, J. C., J. Zhao, M. E. Schranz, E. J. Leon, P. A. Quijada, L. N. Lukens, T. C. Osborn, Flowering time divergence and genomic rearrangements in resynthesized *Brassica* polyploids (Brassicaceae), *Biological Journal of the Linnean Society* 2004, 82:675–688
- Rahman, M., Production of yellow-seeded *Brassica napus* through interspecific crosses, *Plant Breeding* 2001, 120:463-472
- Rahman, M. und P. B. E. McVetty, A review of *Brassica* seed color, *Canadian Journal of Plant Science* 2011, 91:437-446
- Ralph, J., C. Lapierre, J. M. Marita, H. Kim, F. Lu, R. D. Hatfield, S. Ralph, C. Chapple et al., Elucidation of new structures in lignins of CAD- and COMT-deficient plants by NMR, *Phytochemistry* 2001, 57:993–1003
- Ralph, J., K. Lundquist, G. Brunow, F. Lu, H. Kim, P. F. Schatz, J. M. Marita, R. D. Hatfield et al., Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids, *Phytochemistry Reviews* 2004, 3:29–60
- Rana, D., T. van den Boogaart, C. M. O'Neill, L. Hynes, E. Bent, L. Macpherson, J. Y. Park, Y. P. Lim, I. Bancroft, Conservation of the microstructure of genome segments in *Brassica napus* and its diploid relatives, *The Plant Journal* 2004, 40:725-733
- Riley, R. und V. Chapman, Genetic control of the cytologically diploid behaviour of hexaploid wheat. *Nature* 1958, 182: 713–715
- Riley, R., V. Chapman, R. Johnston, Introduction of yellow rust resistance of *Aegilops comosa* into wheat by genetically induced homoeologous recombination. *Nature* 1968, 217: 383–384
- Rogers, L. A. und M. M. Campbell, The genetic control of lignin deposition during plant growth and development, *New Phytologist* 2004, 164:17–30
- Rothberg, J. M., W. Hinz, T. M. Rearick, J. Schultz, W. Mileski, M. Davey, J. H. Leamon, K. Johnson et al., An integrated semiconductor device enabling non-optical genome sequencing, *Nature* 2011, 475:348-352

Schranz, M. E. und T. C. Osborn, Novel flowering time variation in the resynthesized polyploid *Brassica napus*, *Journal of Heredity* 2000, 91:242-246

Schranz, M. E. und T. C. Osborn, De novo variation in life-history traits and responses to growth conditions of resynthesized polyploid *Brassica napus* (Brassicaceae), *American Journal of Botany* 2004, 91:174–183

Schuster, W., Ölpflanzen in Europa, DLG-Verlag, 1992

Sharpe, A. G., I. A. P. Parkin, D. J. Keith, D. J. Lydiate, Frequent nonreciprocal translocations in the amphidiploid genome of oilseed rape (*Brassica napus*), *Genome* 1995, 38:1112-1121

Shirley, B. W., W. L. Kubasek, G. Storz, E. Bruggemann, M. Koornneef, F. M. Ausubel, H. M. Goodman, Analysis of *Arabidopsis* mutants deficient in flavonoid biosynthesis, *The Plant Journal* 1995, 8:659-671

Shirzadegan, M. und G. Röbbelen, Influence of seed color and hull proportion on quality properties of seeds in *Brassica napus* L., *Fette, Seifen, Anstrichmittel* 1985, 87:235-237

Sibout, R., A. Eudes, B. Pollet, T. Goujon, I. Mila, F. Granier, A. Séguin, C. Lapierre, L. Jouanin, Expression pattern of two paralogs encoding cinnamyl alcohol dehydrogenases in *Arabidopsis*. Isolation and characterization of the corresponding mutants, *Plant Physiology* 2003, 132: 848–860

Sibout, R., A. Eudes, G. Mouille, B. Pollet, C. Lapierre, L. Jouanin, A. Séguin, Cinnamyl alcohol dehydrogenase-C and -D are the primary genes involved in lignin biosynthesis in the floral stem of *Arabidopsis*, *The Plant Cell* 2005, 17:2059–2076

Simbaya, J., B. A. Slominski, G. Rakow, L. D. Campbell, R. K. Downey, J. M. Bell, Quality characteristics of yellow-seeded *Brassica* seed meals: protein, carbohydrates, and dietary fiber components, *Journal of Agricultural and Food Chemistry* 1995, 43:2062-2066

Slominski, B. A., L. D. Campbell, W. Guenter, Carbohydrates and Dietary fiber components of yellow- and brown-seeded canola, *Journal of Agricultural and Food Chemistry* 1994, 42:704-707

Slominski, B. A., J. Simbaya, L. D. Campbell, G. Rakow, W. Guenter, Nutritive value for broilers of meals derived from newly developed varieties of yellow-seeded canola, *Animal Feed Science and Technology* 1999, 78:249-262

- Snowdon, R. J. und F. L. I. Luy, Potential to improve oilseed rape and canola breeding in the genomics era, *Plant Breeding* 2012, 131:351—360
- Somers, D. J., G. Rakow, V. K. Prabhu, K. R. D. Friesen, Identification of a major gene and RAPD markers for yellow seed coat colour in *Brassica napus*, *Genome* 2001, 44:1077-1082
- Song, K., P. Lu, K. Tang, T. C. Osborn, Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution, *Proceedings of the National Academy of Sciences* 1995, 92:7719-7723
- Stam, P., Construction of integrated genetic linkage maps by means of a new computer package: Joinmap, *The Plant Journal* 1993, 3:739-744
- Stringam, A. M., D. I. McGregor, S. H. Pawlowski, Chemical and morphological characteristics associated with seedcoat color in rapeseed, *Proceedings 4th Intern. Rapeseed Congress* 1974, Giessen, Deutschland, 99-108
- Szadkowski, E., F. Eber, V. Huteau, M. Lodé, C. Huneau, H. Belcram, O. Coriton, M. J. Manzanares-Dauleux et al., The first meiosis of resynthesized *Brassica napus*, a genome blender, *New Phytologist* 2010 186:102–112
- Taherzadeh, M. J. und K. Karimi, Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: a review, *International Journal of Molecular Sciences* 2008, 9:1621-1651
- Theander, O., P. Åman, G. E. Miksche, S. Yasuda, Carbohydrates, polyphenols, and lignin in seed hulls of different colors from turnip rapeseed, *Journal of Agriculture and Food Chemistry* 1977, 25:270-273
- U, N., Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization, *Japanese Journal of Botany* 1935, 7:389–452
- Udall, J. A., P. A. Quijada, T. Osborn, Detection of chromosomal rearrangements derived from homeologous recombination in four mapping populations of *Brassica napus* L., *Genetics* 2005, 169:967-979
- Umezawa, T., The cinnamate/monolignol pathway, *Phytochemistry Reviews* 2010, 9:1–17
- Utz, H. F. und A. E. Melchinger, PLABQTL: A program for composite interval mapping of QTL, *Journal of Quantitative Trait Loci* 1996,
<http://wheat.pw.usda.gov/jag/papers96/paper196/utz.html>

- Van Soest, P. J., Symposium on nutrition and forage and pastures: new chemical procedures for evaluating forages, *Journal of Animal Science* 1964, 23:838-845
- Van Ooijen, J. W., JoinMap® 4: Software for the calculation of genetic linkage maps in experimental populations, Kyazma B.V., Wageningen, Netherlands, 2006
- Velasco, R., A. Zharkikh, M. Troggio, D. A. Cartwright, A. Cestaro A, D. Pruss, M. Pindo, L. M. FitzGerald et al., A high quality draft consensus sequence of the genome of a heterozygous grapevine variety, *Public Library of Science (PLoS) ONE* 2007, 2(12): e1326. doi:10.1371/journal.pone.0001326
- Wang, S., C. J. Basten, Z.-B. Zeng, Windows QTL Cartographer 2.5, Department of Statistics, North Carolina State University, Raleigh, NC, 2012
(<http://statgen.ncsu.edu/qtlcart/WQTLCart.htm>)
- Ward, A. K., H. L. Classen, F. C. Buchanan, Fishy-egg tainting is recessively inherited when brown-shelled layers are fed canola meal, *Poultry Science* 2009, 88:714–721
- Weier, D., J. Mittasch, D. Strack, C. Milkowski, The genes BnSCT1 and BnSCT2 from *Brassica napus* encoding the final enzyme of sinapine biosynthesis: molecular characterization and suppression, *Planta* 2008, 227:375–385
- Wendel, J. F., Genome evolution in polyploids, *Plant Molecular Biology* 2000, 42:225-249
- Winkel-Shirley, B., Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology, *Plant Physiology* 2001, 126:485-493
- Wittkop, B., R. J. Snowdon, W. Friedt, New NIRS calibrations for fiber fractions reveal broad genetic variation in *Brassica napus* seed quality, *Journal of Agricultural and Food Chemistry* 2012, 60:2248-2256
- Yamamoto, T., J. Yonemaru, M. Yano, Towards the understanding of complex traits in rice: substantially or superficially?, *DNA Research* 2009, 16:141-154
- Zhang, Y., X. Li, W. Chen, B. Yi, J. Wen, J. Shen, C. Ma, B. Chen, Identification of two major QTL for yellow seed color in two crosses of resynthesized *Brassica napus* line No. 2127-17, *Molecular Breeding* 2011, 28:335-342

Anhang

Eine CD-ROM mit den Sequenzdaten aller untersuchten Genotypen ist im Institut für Pflanzenbau und Pflanzenzüchtung I, Heinrich-Buff-Ring 26-32, 35392 Gießen, hinterlegt.

"The limit of man's knowledge in any subject possesses a high interest, which is perhaps increased by its close neighborhood of imagination."

Charles Darwin, *The voyage of The Beagle*, 1834

Danksagung

Ich bedanke mich bei Herrn Prof. Friedt für die Überlassung des Themas, hilfreiche Kommentare zum Manuskript, und nicht zuletzt dafür, daß er es mir ermöglichte, Promotion und Mutterschaft so gut wie möglich zu verbinden.

Ganz besonders danke ich Herrn Rod Snowdon für die stets freundliche und optimistische Betreuung, die kritischen Diskussionen, sowie für viele technische Anregungen und Kommentare zu den Manuskripten.

Herrn Liezhao Liu danke ich für die hervorragende Kooperation bei der genetischen Kartierung.

Herrn Christian Obermeier und Panji Basunanda danke ich für inspirierende Gespräche zur genetischen Kartierung und QTL-Berechnung.

Ich danke auch allen technischen und studentischen Hilfskräften, die mir bei verschiedenen Laborarbeiten hilfreich zur Seite standen, ganz besonders Stavros Tzigos, Liane Renno, Juliane Döhrel und Bashir Hosseini. Ich habe große Hilfsbereitschaft erfahren.

Danke auch an alle Kollegen, die mir vielfältigen Rat gaben und allzeit offene Ohren für mich hatten. Diese Hilfe war von sehr großem Wert für mich. Ich danke ganz besonders Birgit Samans, Frauke Urban, Sarah Hatzig und Sarah Schießl.

Ich danke allen, die Zeit in meine Arbeit investiert haben und somit an dieser Dissertation beteiligt sind.

Zuletzt danke ich der Justus-Liebig-Universität Gießen für die Gewährung des Promotionsabschlußstipendiums.

Eidesstattliche Erklärung

Ich erkläre:

Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.

Anna Stein

Oberursel, im Mai 2013