Genetic and molecular approaches to improve nutritional value of *Brassica napus* L. seed

Nathalie Nesi *, Régine Delourme, Michel Brégeon, Cyril Falentin, Michel Renard

INRA-Agrocampus Rennes-University of Rennes1 Joint Laboratory, UMR118, Plant Genetics and Biotechnologies, BP 35327, 35653 Le Rheu cedex, France

Available online 4 September 2008
Presented by Michel Caboche

**Abstract**

Oilseed rape (*Brassica napus* L.) is a major oil crop that also supplies proteins for the feed industry. In order to reduce total cost production, the objective is to increase oil yield while reducing crop inputs (especially nitrogen and pesticides). Concomitantly, it is necessary to anticipate specific uses (e.g., fatty acid composition) and to ensure the valorisation of the by-products (rapeseed meal). By the past, improvement of seed quality focused on fatty acid balance and low seed glucosinolate content. Current goals include the breeding of yellow-seeded rapeseed lines with high content of seed oil. The use of molecular tools and the exploitation of Arabidopsis knowledge will be presented and discussed.

**Résumé**

Approches génétiques et moléculaires pour l’amélioration de la qualité nutritionnelle des graines de colza (*Brassica napus* L.). Le colza (*Brassica napus* L.) est une plante oléagineuse de grande culture importante dont la graine fournit également une source de matières riches en protéines végétales, utilisable en alimentation animale sous forme de tourteau. L’objectif actuel de la filière est l’augmentation de la production d’huile pour une valorisation alimentaire et non alimentaire tout en limitant les intrants (nutrition azotée, pesticides) afin de garantir la rentabilité de la culture. En marge de ces enjeux, il convient également d’anticiper des besoins spécifiques (e.g., composition en acides gras de l’huile) et d’assurer un débouché aux coproduits (tourteau). Dans le passé, l’amélioration de la qualité de la graine de colza a porté sur la composition en acides gras et la réduction de la teneur en glucosinolates. Actuellement, les objectifs se focalisent sur la sélection de génotypes à graines jaunes à forte teneur en huile dans la graine. L’utilisation des outils moléculaires et l’exploitation des connaissances acquises chez Arabidopsis seront présentées et discutées. Pour citer cet article : N. Nesi et al., C. R. Biologies 331 (2008).

**Nomenclature of fatty acids:** 10:0, capric acid; 12:0, lauric acid; 14:0, myristic acid; 16:0, palmitic acid; 16:1, palmitoleic acid; 18:0, stearic acid; 18:1, oleic acid; 18:2, linoleic acid; 18:3, linolenic acid; 20:1, gadoleic acid; 22:1, erucic acid

**Abbreviations:** ACP, acyl carrier protein; AFLP, amplified fragment length polymorphism; CT, condensed tannins; EMS, ethyl methyl sulfonate; EST, expressed sequence tag; FA, fatty acid; FAD, fatty acid desaturase; Gly3P, glycerol-3-phosphate; GSL, glucosinolates; HEAR, High Erucic Acid Rapeseed; HO, high oleic; LEAR, Low Erucic Acid Rapeseed; LG, linkage group; LL, low linolenic; PFM, physical functional marker; QTL, quantitative trait locus; RFLP, restricted fragment length polymorphism; SNP, single nucleotide polymorphism; SSR, seed storage protein; SSR, single sequence repeat; TAG, triacylglycerol; TT, transparent testa; YS, yellow seed.

* Corresponding author.

E-mail address: nathalie.nesi@rennes.inra.fr (N. Nesi).

URL: http://w3.rennes.inra.fr/apbv/.

© 2008 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

1. Introduction

Oilseed rape (rapeseed; *Brassica napus* L., genome AACC, 2n = 38) arises from spontaneous hybridization between turnip (*Brassica rapa*) (AA, 2n = 20) and cabbage (*Brassica oleracea*) (CC, 2n = 18). It is the most important oilseed crop in Europe and the second one over the world after soybean (*Glycine max*). The crop is primarily used for food and feed, but has recently gained an increasing interest as a source for bio-products (e.g., biodiesel). Winter types are mostly grown in Western Europe, where winters are quite mild. They are sown in late summer and require a period of cold to set flowers. On the contrary, spring cultivars, which are sown at the end of winter, predominate in northern latitudes (e.g., Eastern Europe, Canada, Asia) and Australia. Wintersown varieties usually lead to higher seed yields than spring cultivars. Average seed yield of rapeseed crops depends on variety type (inbred line, hybrid or varietal association), environmental conditions, as well as agronomical practices (soil labor and fertility, fertilizer and pesticide inputs), and therefore vary from one country to the other, ranging from around 1.5–2 t/ha in Canada and Eastern Europe (extensive production of spring varieties) to 3.5 t/ha in Western Europe (intensive production of winter-type rapeseed). Average seed yield of rapeseed crop has shown an increase by up to 50% over the last five decades.

Dry mature seeds from oilseed rape primarily accumulate oil (45–50% w/w) and proteins (20–25% w/w). Lipids in the form of triacylglycerols (TAGs, Fig. 1) are stored in oil bodies within the cytoplasm of cotyledon cells (reviewed in [1]). Storage proteins are deposited in modified vacuoles called protein bodies. For the crushing industry, the main value of rapeseed is determined by the high oil content of the harvested seeds, in spite of some value of the protein fraction for the feed industry. Indeed, oil-free rapeseed meal contains 38–40% of crude proteins that display a well-balanced amino acid composition with high levels of essential sulfur-containing amino acids. Digestibility of rapeseed meal is greatly influenced by the presence of high amounts of fiber in residual hulls that are partially digested by pig, but not at all by poultry. In addition, the presence of phenolic compounds in the residual black seed hulls may contribute to certain undesirable properties of rapeseed meal, including dark color, bitter taste, and atrocity. These compounds, primarily condensed tannins (CTs) and sinapates are considered as anti-nutritive in animal feed. The concentration of phenolics in *Brassica* spp. is at least 30 times higher than that of soybean [2]. The presence of these compounds in rapeseed meal also prevents a move toward using rapeseed meal as a high-quality food-grade supplement.

Quality breeding of oilseed rape has been largely oriented by nutritional concerns, driven by consumer and food industry demands. The main goals to improve nutritional value of oilseed rape are as follows: 1) reduce content of undesirable fatty acids; 2) improve oil stability to diversify applications; 3) increase seed oil content; and 4) improve meal energy value for feed.

The present paper overviews the advances toward the goal of improving seed quality in oilseed rape and it presents some of the future prospects.

2. Sources of genetic variability and strategies to improve seed quality

Variability used in breeding for seed quality in oilseed rape can arise from either natural or induced mutations among the *Brassica* species. For instance, mutants with altered fatty acid composition were found in *B. napus* germplasm collections (e.g., low 22:1; see Section 3.1) or within mutagenized rapeseed populations (e.g., low 18:3, high 18:1; see Section 3.2). In addition, the use of interspecific hybridization between the two highly polymorphic *B. rapa* and *B. oleracea* genomes to resynthesize “novel” *B. napus* germplasm offers the opportunity to get a higher genetic variability than the one expected by solely exploring oilseed rape collections [3]. Novel genetic variations in rapeseed could also originate from gene transfer.

Breeding elite varieties with desirable traits through backcrossing programs could take years. The development of both molecular markers (e.g., RFLP, AFLP, SSR, SNP) as well as high-throughput biochemical methods (e.g., gas chromatography, near-infrared spectrometry) to monitor seed quality features in progeny have greatly improved efficiency of plant breeding. More recently, the completion of the Arabidopsis genome sequence has brought new sources of molecular markers (e.g., the physical functional markers (PFM) [4]) that were directly designed on the gene coding sequences and allowed to draw comparative alignment of
Brassica spp. and Arabidopsis (Arabidopsis thaliana) genomes [5,6]. In addition, the availability of genomic resources in Arabidopsis (e.g., mutants) provided new insights to unravel functions underlying genomic regions involved in seed quality in B. napus.

However, there are seed-specific features that differ from one oilseed crop species to another. To overcome this problem, several initiatives reported the development of rapeseed genomics tools to identify genes expressed during seed development in this species. More than 40,000 cDNA clones have been produced from seed and anther transcripts and sequenced through a Génoplante project (the French plant genomics program [7]). The corresponding ESTs were clustered into 13,083 putative unigenes. Recently, 49,025 high quality ESTs were released by a Canadian consortium and stood for 10,642 putative unique sequences [8]. Analysis of the gene expression pattern during seed development will open new opportunities to identify the critical functions involved in the control of the seed quality in oilseed rape.

3. Modification of the fatty acid (FA) balance of seed oil

The nutritional value and usefulness of vegetable oils depend on their respective FA balance.

3.1. Reduction of erucic acid content

Oilseed rape accumulates naturally TAGs containing erucic acid (22:1) esterified at the sn-1 and sn-3 positions of the glycerol backbone to a level of 45–50% of the total fatty acid (FA) mixture in seeds of the old varieties (the so-called double high seed quality varieties) (Fig. 1). Erucic acid content in rapeseed is genetically controlled by two additive loci (\(E^A\) and \(E^C\)) located on both A- and C-genomes [9], which together explain 90% of the total variation in 22:1 although they do not contribute equally to the 22:1 final content [10]. The two loci have been mapped in rapeseed using a QTL approach [10].

Feeding rats with High Erucic Acid Rapeseed (LEAR) oil promoted myocardial lesions as well as abnormal fat accumulation in animal tissues and at the same time led to a reduction of body weight gain [11–13]. Although this detrimental nutritional effect has never been observed with human beings, 22:1 became undesirable in edible oils and selection of Low Erucic Acid Rapeseed (LEAR) cultivars was achieved by the introduction of recessive alleles (\(e^A\) and \(e^C\)) at both loci involved in 22:1 content. A spontaneous mutant of the German “Liho” spring rapeseed is the only source of LEA in rapeseed and was used to breed the first LEAR variety in Canada and then worldwide [14,15]. Oil from modern LEAR (so-called single zero [“0”] seed quality varieties) lacks nutritionally undesirable long chain FA (less than 2% of 20:1 and 22:1) and is highly appreciated due to its FA profile (5–7% of saturated FA, 58–60% of monounsaturated FA and 30–35% of polyunsaturated FA) that almost fits diet recommendations.

More recently, the two genes \(BnFAE1.1\) and \(BnFAE1.2\), encoding \(\beta\)-ketoacyl-CoA synthases, have been cloned in rapeseed, and were shown to co-segregate with \(E^A\) and \(E^C\) respectively [16,17]. The LEA trait of rapeseed was attributed to a four-nucleotide deletion in the \(BnFAE1\) gene encoding the fatty acid elongase 1 [18].

3.2. Toward high oleic and low linolenic (HOLL) rapeseed varieties

Reduced level of polyunsaturated FA (especially linolenic acid, 18:3) and increased content of monounsaturated FA (oleic acid, 18:1) (Fig. 1) provide higher oil stability and the resulting product can be used for salad dressings as well as for food applications requiring high cooking and frying temperatures. Therefore, breeding rapeseed with high 18:1 (HO) and low 18:3 (LL) content is a major goal. HO and LL rapeseed mutants have been identified in mutagenized rapeseed populations [19]. Genetic analyses revealed that one [20] or two [21] major loci, depending on the mutants and corresponding to the \(FA D_2\) (fatty acid desaturase) genes, controlled 18:1 content, and two loci corresponding to the \(FA D_3\) genes controlled 18:3 content [22,23]. Several studies identified molecular markers tightly associated with both traits [23,24]. More recently, \(FA D_2\) and \(FA D_3\) genes were mapped to the QTL controlling 18:1 and 18:3 contents respectively [25]. In addition, mutations corresponding to single nucleotide polymorphism (SNP) were identified in the sequences of both \(FA D_2\) and \(FA D_3\) genes and used to develop markers for direct selection of desirable alleles for breeding HOLL varieties [21,25].

3.3. Other fatty acids

Most edible oils, like Brassica oil, predominantly consist of unsaturated FA and are therefore hydrogenated to increase the level of saturated FA, which is of interest for particular food uses (e.g., margarines,
confectionery products). However, hydrogenation promotes the production of trans-FA isomers that have been associated with coronary heart diseases. Several studies reported the production of transgenic rapeseed lines with higher levels of saturated FA in seeds than from conventional varieties. For instance, introduction of the cuphea Ch FatB1 gene encoding a palmitoyl-ACP thioesterase resulted in transgenic rapeseeds with more than 35% of palmitic acid (16:0) in total FA instead of 5–10% in conventional lines [26]. High stearic (18:0) lines of oilseed rape were obtained either by reducing the activity of stearoyl-ACP desaturase [27] or through introduction of the Garm FatA1 gene encoding an acyl-ACP thioesterase from mangosteen (Garcinia mangostana) [28].

4. Progress to increase seed oil content in oilseed rape

4.1. Genetic control of seed oil content

Increase in seed oil content is a major goal for oilseed rape breeding. However, seed oil content is under a complex genetic determinism that is still poorly understood. A better knowledge of genetic determinism of oil content will be relevant for the breeders to control the genetic advance of the crop. By using different segregating rapeseed populations, recent studies reported the identification of numerous QTL (7 to 14 regions per study) involved in the control of oil content, which is consistent with the polygenic determinism of the trait [29–34]. Each of the QTL accounted for less than 10% of the total oil content variance [33,34]. Some of these QTL coincided with loci controlling erucic acid content, suggesting that it is a major determinant for oil content in oilseed rape [29,30,34]. Additive effects were shown to be the main factors controlling oil content [29,33,34], with individual additive effect of the different alleles ranging from 0.2 to 1.2% [33,34]. In addition, strong environmental effects underlie variations in oil content [31,33,35].

In silico comparison of the genetic locations for the published oil content QTL was conducted by Debourme and coworkers [33] and led to the preliminary conclusion that the region mapped on linkage group (LG) A3 was revealed across the different genetic backgrounds studied so far, suggesting that the same molecular determinant underlies these QTL. In addition, oil QTL located on LGs A1, A8 and C3 were identified in three studies out of the five published. On the contrary, some other QTL were specifically found in only one population. To address the question whether these QTL cover the same regions across the different genetic backgrounds, a consensus genetic map will be built by increasing the number of common genetic markers between the existing oilseed rape crosses. Then,
combining genotyping and phenotyping data will allow to establish a consolidate map of QTL for oil content.

In Arabidopsis, Hobbs and coworkers [36] reported the identification of four QTL that control seed oil content in the Ler x Cvi recombinant inbred line population and their localization on chromosomes 1 (top and bottom), 2 (bottom) and 3 (top). In an attempt to resolve oilseed rape QTL, we used Arabidopsis genomic data to derive markers in rapeseed and showed that the rapeseed oil QTL on A1 was collinear to the Arabidopsis QTL3t (N. Nesi et al. unpublished results). These preliminary results demonstrate the feasibility to use the Arabidopsis genome sequence to increase the marker density within the vicinity of a rapeseed QTL. In addition, using comparative genomics between Arabidopsis and rapeseed would offer opportunities to infer the likely presence of candidate genes for the control of seed oil content.

4.2. Transgenic approaches to increase seed oil content

Attempts to improve final lipid content of the seed have primarily focused on metabolic steps involved in FA biosynthesis or TAG assembly. Over-expression of a lysophosphatidate acyltransferase gene from yeast (Saccharomyces cerevisiae) in oilseed rape significantly promotes seed oil content under controlled conditions [37]. When assayed under field conditions, these transgenic lines displayed increases of about 10% in oil content [38]. By contrast, engineering individual FA biosynthesis genes did not significantly increase lipid accumulation in seeds (reviewed in [39]). More recently, it has been proposed that the glycerol-3-phosphate (Gly3P) supply is a limiting factor for lipid synthesis [40]. Over-expression of a yeast Gly3P dehydrogenase gene in rapeseed resulted in a three- to fourfold increase in Gly3P content that leads to a 40% increase in lipid content [41].

5. Attempts to improve the rapeseed meal value

5.1. Seed storage protein (SSP) content and composition

Seed protein content is generally negatively correlated with seed oil content [32] and improvement of the seed quality in oilseed rape has been conducted with little attention paid to the protein fraction. Cruciferins (12S globulins), napins (2S albumins) and oleosins (oil body proteins, reviewed in [1]) are the major proteins in oilseed rape seeds. Taken together, SSP (cruciferins and napins) accounted for up to 70% of seed proteins and display specific features. Napins contain higher levels of sulfur and aromatic residues (essential amino acids) than cruciferins, which make them the most important targets for the improvement of seed protein composition. Several promising attempts for the genetic engineering of 2S albums have been conducted in rapeseed through introduction of a Brazil nut (Bertholletia excelsa) 2S gene [42,43] or expression of a cruciferin antisense gene [44]. In all cases, transgenic plants accumulate more napins in their seeds, which led to an increase in cysteine, methionine and lysine contents, the two latter amino acids being essential. In addition, the increase in napins was counter-balanced by a decrease in cruciferin content, suggesting that the 12S/2S balance was tightly controlled.

Surprisingly, only two studies reported the existence of variability for SSP composition in rapeseed [45,46]. The authors showed that the 12S/2S ratio can vary from 0.7 to 2, with the total SSP content remaining stable, consistent with the existence of a regulator controlling the SSP balance. In addition, analysis of a wide set of rapeseed varieties for their relative SSP composition suggested that the 12S/2S ratio has increased in modern grown double low varieties (zero 22:1 and low seed glucosinolate [GSL] content) while old rapeseed cultivars (double high and “0” varieties) display higher levels of napins [46]. At least two hypotheses can be proposed to address this difference between old and modern oilseed rape varieties. First, the reduction of aliphatic GSL (methionine-derived) in double low varieties (see Section 5.2) could have perturbed amino acid metabolism and subsequent accumulation of SSP. Such association between GSL metabolism and amino acid biosynthesis has been observed in Arabidopsis [47]. Second, the presence of residual chromosomal segments of the low seed GSL cultivar in modern double low varieties could account for modifications in seed quality as previously suggested [48]. More recently, Devouge and coworkers [49] reported the differential proteomic analysis of four near-isogenic rapeseed varieties chosen upon their 22:1 and GSL contents. Optimization of 2-D electrophoresis conditions and protein identification by mass spectrometry analysis allowed identifying 69 spots that were differentially detected between the four lines. However, accumulation of cruciferins was found to be only slightly affected, whereas most of the significant differences were observed for proteins involved in plant defense or carbohydrate metabolism.
5.2. Reduction of antinutritive compounds

5.2.1. Glucosinolates (GSLs)

GSLs are secondary metabolites synthesized by Brassicaceae. Three groups of GSLs designated aliphatic, aromatic (phenylalanine-derived) or indole (tryptophan-derived) are found in oilseed rape. Intakes of high amounts of GSLs in rapeseed meal can result in goitrogen-induced hypertrophy. The low seed GSL trait was identified in the Polish spring rapeseed “Bronowski” and subsequently used for backcrossing programs that resulted in the double low (“00”) oilseed rape varieties. The modern grown “00” varieties display 10–15 µmol GSL/g seed instead of 60–100 µmol/g seed in old varieties. Genes involved in GSL metabolism have been associated with QTL for seed GSL content and will be helpful to monitor GSL content in the seed without modifying GSL level in other plant organs [50].

5.2.2. Condensed tannins (CTs)

Oilseed rape seed is brown due to the accumulation of CTs in the integuments. Yellow-seeded (YS) rapeseed lines that are deprived of CTs have been developed through interspecific introgression of yellow seed coat color genes from related species (B. rapa, B. carinata, B. juncea; Fig. 2). Recently, YS variants were identified in an EMS-mutagenized rapeseed population (N. Nesi et al., unpublished data). YS rapeseed lines display low fiber content and offer improved quality of the seed as well as of the derived meal (higher proportions of oil and proteins [51]). Despite important research efforts during the last 20 years, attempts to develop a true breeding rapeseed that consistently yields pure and bright yellow seeds under a wide range of environmental conditions have not been successful [52]. Genetic analyses of YS from independent origins showed that at least three loci were involved in seed pigmentation with partial dominance effect [53,54]. Molecular markers linked to seed coat color trait were developed in rapeseeds [53]. Therefore, a better knowledge of the CT metabolic network in this species and especially of the biosynthetic and/or regulatory key steps should help to breed YS rapeseed lines. The flavonoid metabolism, and in particular the CT sub-pathway, has been extensively studied in the model crucifer Arabidopsis (reviewed in [55]). To date, 26 independent loci involved in seed coat pigmentation (the so-called Transparent Testa [TT] genes) have been identified. Orthologs of the TT genes in rapeseed have been identified ([56,57]; Auger, Nesi, unpublished results) and some of the BnTT genes were proposed to co-localize with QTL for seed color and fiber content [54].

5.2.3. Sinapine

Sinapate esters, like CTs, can form complexes with meal proteins thus reducing their bio-availability and digestibility. Seed-specific genes involved in the final steps of sinapate ester metabolism have been identified, such as those encoding UDP-glucose:sinapate
glucosyltransferase (SGT1) or sinapoyl-glucose:choline sinapoyltransferase (SCT1 and SCT2). Silencing these target genes in transgenic rapeseed plants resulted in substantial reductions (60–70%) in the total sinapate ester content in seeds [58,59].

6. Concluding remarks

Breeding of rapeseed for improved seed quality has led to the development of double low cultivars with zero 22:1 and low seed GSL contents that are grown worldwide. Intensive and dynamic research on seed quality in *B. napus* has also allowed the identification of numerous genomic regions involved in seed quality within several mapping populations. One future challenge is therefore to integrate these data together in order to establish a consolidate map of QTL for seed quality. This will definitely point out the genomic regions critical for elaboration of seed features. In addition, the achievement of the Arabidopsis genome sequence can help in resolving QTL for seed quality. Finally, the complete sequence of *B. rapa* (A genome) will be soon released [60] and will bring new perspectives for identification and characterization of functions underlying seed quality QTL.

Current goals for improving seed quality in rapeseed deal with the development of yellow-seeded cultivars that are high yielding and display higher levels of oil and proteins. In addition, future prospects will also include the identification of genotypes able to grow under low input farming regimes (especially low nitrogen input). Therefore, elaboration of seed quality (especially C/N balance) will need to be revisited in the context of low N nutrition, with the objective to improve N mobilization without altering oil accumulation in the seed.

Major advances in improving nutritional quality in other oleaginous plants have also been conducted with similar approaches in sunflower (*Helianthus annuus*), soybean or linseed (*Linum usitatissimum*). For instance, oleic sunflower varieties (18:1 > 82%) have been set up and are now grown [61]. This underlies the plasticity of seed metabolism that can overcome drastic modifications. However, questions about the limits for modification of the seed features remain to be addressed. For instance, what are the limits in increasing oil content in terms of C/N balance in the seed? What are the impacts of such modifications on seed development and seedling vigor? Attempts to produce Arabidopsis and rapeseed transgenic plants with very low levels of SSP are underway in our laboratory and would provide clues to answer these questions.

References


