



Deposited via The University of York.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/219864/>

Version: Accepted Version

Article:

Niu, Yan, Li, Wenwen, Yang, Yinghui et al. (Accepted: 2024) Creation of rapeseed germplasm with high polyunsaturated fatty acid content by relative introgression of *Brassica carinata*. *Plant Communications*. ISSN: 2590-3462 (In Press)

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Journal Pre-proof

Creation of rapeseed germplasm with high polyunsaturated fatty acid content by relative introgression of *Brassica carinata*

Yan Niu, Wenwen Li, Yinghui Yang, Hao Wang, Zhesi He, Han Qin, Yikai Zhang, Dandan Hu, Jing Wang, Chunyu Zhang, Guangsheng Yang, Ian Bancroft, Jun Zou

PII: S2590-3462(24)00614-X

DOI: <https://doi.org/10.1016/j.xplc.2024.101193>

Reference: XPLC 101193

To appear in: *PLANT COMMUNICATIONS*

Received Date: 14 July 2023

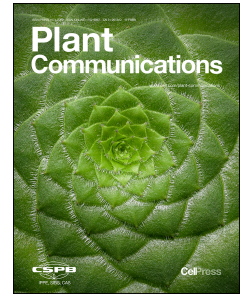
Revised Date: 30 August 2024

Accepted Date: 15 November 2024

Please cite this article as: Niu, Y., Li, W., Yang, Y., Wang, H., He, Z., Qin, H., Zhang, Y., Hu, D., Wang, J., Zhang, C., Yang, G., Bancroft, I., Zou, J., Creation of rapeseed germplasm with high polyunsaturated fatty acid content by relative introgression of *Brassica carinata*, *PLANT COMMUNICATIONS* (2024), doi: <https://doi.org/10.1016/j.xplc.2024.101193>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2024



1 **Creation of rapeseed germplasm with high**
2 **polyunsaturated fatty acid content by relative**
3 **introgression of *Brassica carinata***

4 Dear Editor,

5 *Brassica napus* L., commonly known as rapeseed, canola or oilseed rape,
6 is the world's third oilseed crop and accounts for ~12% of major worldwide oil
7 production (FAO., 2022). Rapeseed provides not only healthy and nutritionally
8 balanced edible oil for humans, but also protein-rich fodder for animals, and
9 renewable materials for biodiesel and industrial applications. Polyunsaturated
10 fatty acids (PUFAs), including omega-3 fatty acids and omega-6 fatty acids,
11 are mainly essential fatty acid and play a crucial role in the human diet, and
12 humans obtain these fatty acids through dietary sources. Most crop seeds and
13 vegetable oils, including canola, soybean, corn, and sunflower oils, are major
14 sources of omega-6 fatty acids in the form of linoleic acid with low proportions
15 of omega-3 fatty acids (Saini and Keum, 2018). There are several key genes
16 responsible for PUFAs, such as *FAD2*, which converts oleic acid to linoleic
17 acid and *FAD3*, which converts linoleic acid to linolenic acid (Okuley et al.,
18 1994; Yadav et al., 1993).

19 As a young allotetraploid species formed by interspecific crosses
20 between the diploid ancestors of *B. rapa* ($2n = 20$, AA) and *B. oleracea* ($2n =$
21 18 , CC), *B. napus* ($2n=38$, AACCC) has a limited genetic diversity due to the
22 short history of cultivation and domestication and long-term double-low
23 breeding (Allender and King, 2010; Friedt and Snowdon, 2010; Prakash et al.,
24 2011). Wild relatives and other related genetic resources represent a
25 favorable genetic variation that can be exploited to develop climate-resilient
26 modern crops (Wang et al., 2020; Zhou et al., 2021). During the process of
27 interspecific hybridization, exotic introgression occurs as a natural gene flow
28 and it not only introduces new genetic variations but also increases the

29 complexity of crop genomes. However, how to mine the elite or favorable
30 genetic variation with mixed genomic variations and introgression is a
31 challenge. Therefore, we aim to explore the novel germplasm harboring rich
32 exotic introgression with more efficiency and provide insights for explaining
33 the regulatory mechanisms for exotic introgression and genomics-assisted
34 wild relative introgression.

35 To broaden the genetic basis of rapeseed, through interspecific crosses
36 between 75 *B. carinata* accessions and 122 *B. rapa* accessions, we obtained
37 more than 20,000 lines of new-type *B. napus* gene pool (Hu et al., 2019; Zou
38 et al., 2018). From the gene pool of new-type *B. napus*, we select one line
39 called HLL45 with five large-scale homoeologous exchange (HE) including
40 C5/B1, A2/C2, A3/C3, A4/C4 and A9/C8 (Figure 1A, Supplemental Figure 1).
41 Interestingly, HE between C5/B1 resulted in an introgression fragment with
42 more than 5Mb on C5 from B1 chromosome of *B. carinata* (Supplemental
43 Figure 1). We investigated agronomic traits and quality traits for HLL45 and
44 found that it is a novel germplasm of high polyunsaturated fatty acid (PUFA)
45 content at 40%~45%, including linoleic acid (LEI) and linolenic acid (LEN)
46 content. Compared to conventional *B. napus* varieties, the PUFA content of
47 HLL45 was 15 percentage higher, mainly contributed by LEI, but its seed oil
48 content decreased (Figure 1B and Supplemental Table 1). Our lines should
49 therefore be considered pre-breeding material, from which new elite cultivars
50 can be developed. Using a genetic map constructed by HG DH population
51 (Hu et al., 2024), we identified a major QTL, *qFA-C5-4*, that explained 63.2%-
52 64.6%, 37.3%-73.3%, and 7.4% of the phenotypic variation for oleic acid
53 content (OLE), LEI and LEN, respectively (Figure 1B, Supplemental Figure 2
54 and Supplemental Table 2-5). Notably, the markers defining QTL *qFA-C5-4*
55 were located in the *B. carinata* B1 segment introgression region on
56 chromosome C5, which contains an orthologue of the *FAD2* gene, which
57 encodes the enzyme that converts OLE to LEI (Supplemental Figure 2).
58 Therefore, PUFA content in rapeseed appears to be associated with the B1

59 introgression from *B. carinata*.

60 To comprehensively study the characteristics and effect of exotic
61 introgression, we assembled four high-quality genomes including HLL45 and
62 its representational parents (Supplemental Figure 3 and 4) by a combination
63 of Illumina short reads, Oxford Nanopore Technology (ONT) long reads, Hi-C
64 reads and genetic map. The assembled genome for HLL45 comprised 344
65 contigs (N50, 18.0 Mb) and had a total size of 927 Mb, with high BUSCO
66 completeness at 99.0% and a mapping rate of 99.7% from short reads,
67 leading to a highly accurate and nearly complete assembly (Supplemental
68 Figure 4 and Supplemental Table 6). We successfully anchored 99.0% of the
69 contigs to 19 pseudochromosomes exhibiting good collinearity with HG
70 genetic map (Hu *et al.*, 2024), and annotated 101,495 protein-coding genes
71 for HLL45 (Supplemental Figure 5 and Supplemental Table 6). We identified a
72 total of 11,611 exotic genes with mainly on A05 and C07 in HLL45, including
73 4,538 genes from *B. rapa* and 7,073 genes from *B. carinata*. Among exotic
74 genes from *B. carinata*, 1,043 genes and 5,919 genes were exotic genes from
75 B and C sub-genome respectively, with the rest genes coming from
76 unanchored sequences. 80% of exotic genes from B sub-genome of *B.*
77 *carinata* were located on chromosome C5 and were mainly enriched in fatty
78 acid metabolism, unsaturated fatty acids biosynthesis and linoleic acid
79 metabolism (Figure 1D). For the major QTL *qFA-C5-4* located in the 1.34 Mb
80 region (50.0 Mb ~ 51.4 Mb) of C5 in HLL45 genome, there were 217 genes in
81 the region and 131 genes were different expressed genes in all of 10 different
82 seed development stages between HLL45 and ZS11 (Figure 1E and
83 Supplemental Figure 6). *BcaFAD2.B1* (BnaHLLC05G045860) had the highest
84 expression in seeds, and co-expression network analysis indicated that it was
85 the hub gene of PUFA (Supplemental Figure 7). We identified a six-Mb exotic
86 introgression on C5 (from 49.9 Mb to 55.9 Mb) in HLL45 from B1 of *B.*
87 *Carinata*, showing the *FAD2* on C5 was replaced by *BcaFAD2.B1* (Figure 1F).
88 Lines in HG population carrying the C5 SV contained 8.56% higher LEI than

89 those lines with no SV (Figure 1G and Supplemental Table 7). We further
90 confirmed SV on C5 was stably inheritable (Supplemental Figure 8). In the
91 HG population, the high-LEI lines carried the *BcaFAD2.B1* and C5 SV while
92 the low-LEI lines did not (Figure 1H and Supplemental Table 8). These results
93 implicate the six-Mb exotic introgression on C5 play an important role in
94 increasing PUFA.

95 We further surveyed the expression level of *FAD2* orthologues of six
96 species in *Brassica* and found that *BcaFAD2.B1* had the highest expression,
97 which had 89% of total orthologs expression in *B. carinata* (Supplemental
98 Figure 9). In HLL45, three *FAD2* orthologues was expressed and mainly
99 contributed by *FAD2.B1* with reaching its highest expression at 36 DAF, while
100 *FAD2.C5* was not expressed in HLL45 (Figure 1I). *FAD2.A5* and *FAD2.C1* in
101 HLL45 were both expressed at about the similar level as them in traditional *B.*
102 *napus* ZS11, while *FAD2.B1* was expressed more highly than *FAD2.C5* in
103 ZS11, which was confirmed by RT-PCR (Figure 1J). Except in seeds,
104 *FAD2.B1* was also expressed in root, leaf and stem (Supplemental Figure 10).
105 During the seed development, the PUFA content in HLL45 was higher than
106 that of ZS11, especially LEI. After 18 DAF, the LEI content in HLL45 was
107 higher than in ZS11, and the difference increased as the developmental time
108 extends (Figure 1J and Supplemental Figure 11). Therefore, the increased
109 PUFA content of HLL45 can be explained that *FAD2.C5* was replaced by
110 *BcaFAD2.B1* with highly expressed level.

111 *FAD2.B1* was the nearest clade to *FAD2* in *A. thaliana* and had many
112 sequences in promoter and enhancer within the 800bp upstream of gene,
113 while *FAD2.C5* did not have (Supplemental Figure 12 A, B and Supplemental
114 Table 9). Compared to *BcaFAD2.B1* and *BniFAD2.B1* with one exon,
115 *BnaFAD2.C5* and *BoIFAD2.C5* had one additional exon encoding 73 AA and
116 same conserved domain (Supplemental Figure 12C and 13). Besides,
117 *FAD2.C5* in HLL45 (BnaHLLC05G045860) and *BcaFAD2.B1*
118 (Bca4012B1G035280) had the same structure, but they both had almost

119 identical C-termini but distinct N-termini with *BnaFAD2.C5*
120 (BnaC05G0480500ZS) (Supplemental Figure 12D).

121 High PUFA content of rapeseed oil is a desirable trait which has been
122 unachievable using conventional germplasm. We have demonstrated the use
123 of wild relative introgression to achieve this, implicating high expression of
124 one of the *FAD2* orthologues from the B genome, which is not normally
125 present in oilseed rape, as causative. Our study provides both evidence for
126 how PUFA content in seed oil above the normally occurring range is controlled,
127 and an example for others to follow of how to use genomics to accelerate use
128 of wild relative introgression for crop improvement.

129 **Data availability**

130 The genome assembly have been deposited in the Genome Warehouse at
131 the National Genomics Data Center, Beijing Institute of Genomics, Chinese
132 Academy of Sciences/China National Center for Bioinformation, under
133 accession number PRJCA025800. The genomic sequencing data and
134 transcriptome sequencing data have been deposited in National Genomics
135 Data Center (GSA: CRA016277) that are publicly accessible
136 at <https://ngdc.cncb.ac.cn/gsa>.

137 **Acknowledgements**

138 This study was supported by the National Natural Science Foundation of
139 China (32171982, 31970564), the National Key Research and Development
140 Program of China (2021YFF1000100), the Fundamental Research Funds for
141 the Central Universities (2662023PY004) and the UK Department for
142 Environment, Food and Rural Affairs Oilseed Rape Genetic Improvement
143 Network (OREGIN) (CH0110).

144 Author Contributions

145 Y.N. performed the research, analyzed the data and wrote the paper. W.L.
146 performed field trials, genetic mapping and QTL analysis. Y.Y. performed
147 phenotyping, sampling and molecular marker analysis. H. W. and Y. Z.
148 contributed to the genome assembly. J.Z and H.Q contributed to plant
149 material development, cultivation, phenotyping, and QTL analysis. Z.H.
150 performed GDTPs analysis. D.H. contributed to QTL and genome
151 resequencing analysis. J. W. and C.Z. contributed to the design for RT-PCR
152 and fatty acid composition analysis. G.Y. contributed to the design of field
153 trials and revise the manuscript. J.Z. and I.B. designed the project, analyzed
154 the data, and wrote the manuscript. All authors edited and approved the final
155 manuscript.

156 Declaration of competing interest

157 The authors declare no competing interests.

158 **REFERENCES**

- 159 Allender, C.J., and King, G.J. (2010). Origins of the amphiploid species
160 *Brassica napus* L. investigated by chloroplast and nuclear molecular markers.
161 BMC Plant Biol 10:54. 10.1186/1471-2229-10-54.
- 162 FAO. (2022). WORLD FOOD AND AGRICULTURE STATISTICAL
163 YEARBOOK 2022 (FAO).
- 164 Friedt, W., and Snowdon, R. (2010). Oilseed rape. Oil crops:91-126.
- 165 Hu, D., Zhang, W., Zhang, Y., Chang, S., Chen, L., Chen, Y., Shi, Y., Shen, J.,
166 Meng, J., and Zou, J. (2019). Reconstituting the genome of a young
167 allopolyploid crop, *Brassica napus*, with its related species. Plant Biotechnol J
168 17:1106-1118. 10.1111/pbi.13041.
- 169 Hu, D., Lu, J., Li, W., Yang, Y., Xu, J., Qin, H., Wang, H., Niu, Y., Zhang, H.,
170 Liu, Q., et al. (2024). The occurrence, inheritance, and segregation of complex
171 genomic structural variation in synthetic *Brassica napus*. The Crop Journal
172 <https://doi.org/10.1016/j.cj.2024.01.002>.
- 173 Okuley, J., Lightner, J., Feldmann, K., Yadav, N., Lark, E., and Browse, J.
174 (1994). *Arabidopsis FAD2* gene encodes the enzyme that is essential for
175 polyunsaturated lipid synthesis. Plant Cell 6:147-158. 10.1105/tpc.6.1.147.
- 176 Prakash, S., Wu, X.i., and Bhat, S.R. (2011). History, Evolution, and
177 Domestication of *Brassica* Crops (Plant Breeding Reviews).
- 178 Saini, R.K., and Keum, Y.S. (2018). Omega-3 and omega-6 polyunsaturated
179 fatty acids: Dietary sources, metabolism, and significance — A review. Life
180 Sciences 203:255-267.
- 181 Wang, H., Sun, S., Ge, W., Zhao, L., Hou, B., Wang, K., Lyu, Z., Chen, L., Xu,
182 S., Guo, J., et al. (2020). Horizontal gene transfer of *Fhb7* from fungus
183 underlies *Fusarium* head blight resistance in wheat. Science
184 36810.1126/science.aba5435.
- 185 Yadav, N.S., Wierzbicki, A., Aegerter, M., Caster, C.S., Pérez-Grau, L., Kinney,
186 A.J., Hitz, W.D., Booth, J.R., Jr., Schweiger, B., Stecca, K.L., et al. (1993).

187 Cloning of higher plant omega-3 fatty acid desaturases. *Plant Physiol*
188 103:467-476. 10.1104/pp.103.2.467.

189 Zhou, Y., Bai, S., Li, H., Sun, G., Zhang, D., Ma, F., Zhao, X., Nie, F., Li, J.,
190 Chen, L., et al. (2021). Introgressing the *Aegilops tauschii* genome into wheat
191 as a basis for cereal improvement. *Nat Plants* 7:774-786. 10.1038/s41477-
192 021-00934-w.

193 Zou, J., Hu, D., Mason, A.S., Shen, X., Wang, X., Wang, N., Grandke, F.,
194 Wang, M., Chang, S., Snowdon, R.J., et al. (2018). Genetic changes in a
195 novel breeding population of *Brassica napus* synthesized from hundreds of
196 crosses between *B. rapa* and *B. carinata*. *Plant Biotechnol J* 16:507-519.
197 10.1111/pbi.12791.

198 **Figure 1 The exotic introgression on C5 in HLL45 from *B. carinata***
199 **increasing PUFA content.**

200 **(A)** Large-scale homoeologous exchanges (HEs) of new-type *B. napus* HLL45
201 identified by Genome Display Tile Plots. Red, black, orange, purple and green
202 arrows represent HEs of A2/C2, A3/C3, A4/C4, C5/B1 and A9/C8,
203 respectively.

204 **(B)** The upper figure represents the fatty acid component distribution for five *B.*
205 *napus* accessions. LEI, LEN, OLE and EA mean linoleic acid content, linolenic
206 acid content, oleic acid content and erucic acid content, respectively. The
207 bottom figure represents the QTL mapping for fatty acid traits using HG DH
208 populations (280 lines).

209 **(C)** The genomic features of exotic introgression in HLL45 (100 kb window). I-
210 V means gene density, TE density, exotic gene density from BraA (*B. rapa*),
211 BcaB and BcaC (B and C sub-genome of *B. carinata*).

212 **(D)** Pathway enrichment plot of exotic genes from BcaB in HLL45 (top 15).

213 **(E)** The volcano plot of 217 genes within the major QTL on C5 in seeds (40
214 days after pollination). Compared to ZS11, most candidate genes in HLL45

215 were significantly up-regulated, especially BnaHLLC05G045860
216 (*BcaFAD2.B1*).

217 **(F)** The collinearity of chromosome C5 among HLL45, *B. napus* Darmor-bzh
218 and *B. carinata* C4012.

219 **(G)** The generation, transmission process and effect of SV on C5. Difference
220 on linoleic acid (LEI) and oleic acid (OLE) content between lines with C5 SV
221 and no SV from HG population. Statistical significance was determined by
222 two-tailed unpaired Student's T-test.

223 **(H)** PCR validation for C5 SV and exotic gene *BcaFAD2.B1*. The left figure
224 represents PCR genotyping assay for lines with different linoleic acid (LEI)
225 content from HG populations on C5 SV markers. PCR amplicons of 211 bp
226 and 577 bp represent the presence and absence of the SV, respectively. The
227 left figure represents PCR genotyping assay for HLL45, HS3 and lines with
228 different linoleic acid (LEI) content from HG populations on *BcaFAD2.B1*
229 markers. PCR amplicons of 424 bp and 222 bp represent the presence and
230 absence of *BcaFAD2.B1*, respectively.

231 **(I)** The gene expression pattern of three *FAD2* orthologues in different tissues
232 and seed development in HLL45 and conventional *B. napus* ZS11 based on
233 FPKM.

234 **(J)** The upper figure means LEI and LEN content in different seed
235 development of new-type *B. napus* HLL45 and conventional *B. napus* ZS11.
236 LEI means linoleic acid content and LEN means linolenic acid content. The
237 bottom figure means the gene expression pattern of three *FAD2* orthologues
238 in different seed development of HLL45 and conventional *B. napus* ZS11
239 based on the relative expression level from RT-PCR.

240 Yan Niu¹, Wenwen Li¹, Yinghui Yang¹, Hao Wang¹, Zhesi He², Han

241 Qin¹, Yikai Zhang¹, Dandan Hu¹, Jing Wang¹, Chunyu Zhang¹,

242 Guangsheng Yang¹, Ian Bancroft^{2*}, Jun Zou^{1*}

243 ¹National Key Laboratory of Crop Genetic Improvement, College of Plant
244 Science and Technology, Huazhong Agricultural University, Wuhan, China
245 ²Department of Biology, University of York, Heslington, York, YO10 5DD, UK
246 ***Correspondence: Jun Zou (zoujun@mail.hzau.edu.cn) and Ian Bancroft**
247 **(ian.bancroft@york.ac.uk)**

Journal Pre-proof

