



**HAL**  
open science

## Why old duplicated genes are not thrown away in (paleo)polyploids? Example from the petC gene in *Brassica napus*

Maud Facon, Gwenaëlle Deniot, Maryse Lodé-Taburel, Heloise Archambeau, Emilie Montes, Younes Dello, Loeiz Maillet, Anne-Marie Chèvre, Mathieu Rousseau-Gueutin

### ► To cite this version:

Maud Facon, Gwenaëlle Deniot, Maryse Lodé-Taburel, Heloise Archambeau, Emilie Montes, et al.. Why old duplicated genes are not thrown away in (paleo)polyploids? Example from the petC gene in *Brassica napus*. *Plant Molecular Biology*, 2023, *Plant Molecular Biology*, 113, pp.323-327. 10.1007/s11103-023-01387-9 . hal-04338534

**HAL Id: hal-04338534**

**<https://hal.science/hal-04338534>**

Submitted on 29 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Title: Why old duplicated genes are not thrown away in (paleo)polyploids? Example from the *petC***  
2 **gene in *Brassica napus***

3 Maud Facon<sup>1</sup>, Gwenaelle Deniot<sup>1</sup>, Maryse Lode-Taburel<sup>1</sup>, Heloïse Archambeau<sup>1</sup>, Emilie Montes<sup>1</sup>, Younes  
4 Delloero<sup>1</sup>, Loeiz Maillet<sup>1</sup>, Anne-Marie Chèvre<sup>1</sup>, Mathieu Rousseau-Gueutin<sup>1\*</sup>

5 <sup>1</sup>IGEPP, INRAE, Institut Agro, Université de Rennes, 35650 Le Rheu, France

6 Maud Facon: [maud.facon@gmail.com](mailto:maud.facon@gmail.com);

7 Gwenaelle Deniot: [gwenaelle.deniot@inrae.fr](mailto:gwenaelle.deniot@inrae.fr);

8 Maryse Lode-Taburel: [maryse.taburel@inrae.fr](mailto:maryse.taburel@inrae.fr);

9 Heloïse Archambeau: [heloise.archambeau@argaly.com](mailto:heloise.archambeau@argaly.com);

10 Emilie Montes: [emilie.montes@ens-lyon.fr](mailto:emilie.montes@ens-lyon.fr);

11 Younès Delloero: [younes.delloero@inrae.fr](mailto:younes.delloero@inrae.fr);

12 Loeiz Maillet: [loeiz.maillet@inrae.fr](mailto:loeiz.maillet@inrae.fr);

13 Anne-Marie Chèvre: [anne-marie.chevre@inrae.fr](mailto:anne-marie.chevre@inrae.fr);

14 Mathieu Rousseau-Gueutin : [mathieu.rousseau-gueutin@inrae.fr](mailto:mathieu.rousseau-gueutin@inrae.fr);

15

16 \*Author for correspondence:

17 Mathieu Rousseau-Gueutin

18 Tel: +33 (0)2 23 48 51 31

19 ORCID: 0000-0002-1130-1090

20 Email: [mathieu.rousseau-gueutin@inrae.fr](mailto:mathieu.rousseau-gueutin@inrae.fr)

21 **Acknowledgement**

22 We thank F. Nogue, L. Chauvin, and F. Veillet for their help in vector and experiment design, as well as  
23 A. Bouchereau for providing the Open FluorCam.

24 **Abstract:** Duplication of genes at different time period, through recurrent and frequent polyploidization  
25 events, have played a major role in plant evolution, adaptation and diversification. Interestingly, some of  
26 the old duplicated genes (referred as paleologs) have been maintained for millions of years, and there is  
27 still a poor knowledge of the reasons of their retention, especially when testing the phenotypic effect of  
28 individual copies by using functional genetic approaches. To fill this gap, we performed functional genetic  
29 (Crispr-Cas9), physiological, transcriptomic and evolutionary studies to finely investigate this open  
30 question, taking the example of the *petC* gene (involved in cytochrome b6/f and thus impacting  
31 photosynthesis) that is present in four paleologous copies in the oilseed crop *Brassica napus*. RNA-Seq  
32 and selective pressure analyses highly suggested that all paleologous copies conserved the same function  
33 and that they were all highly transcribed. Thereafter, the Knock Out of one, several or all *petC* copies  
34 highlighted that all paleologous copies have to be Knocked out to suppress the gene function. In addition,  
35 we could determine that phenotypic effects in single and double mutants could only be deciphered in high  
36 light conditions. Interestingly, we did not detect any significant differences between single mutants  
37 Knocked Out for either the A03 or A09 copy (despite they were differentially transcribed), or even between  
38 mutants for a single or two *petC* copies. Altogether, this work revealed that *petC* paleologs have retained  
39 their ancestral function and that the retention of these copies is explained by their compensatory role,  
40 especially in optimal environmental conditions.

41 **Keywords:** (Paleo)polyploidy, duplicated gene evolution, fractionation, *Brassica napus*, Crispr-Cas9

42

### 43 Text File

44 Polyploidy or whole genome duplication, is a key process in species evolution, adaptation and  
45 diversification. It is frequent and recurrent in flowering plants, generating numerous copies of the same  
46 gene at different time periods. These copies may be subjected to different evolutionary fates: they may  
47 either lose (pseudogenization through fractionation), conserve their ancestral (subfunctionalization) or  
48 acquire a new function (neo-functionalization) (Birchler and Yang, 2022). **Despite much knowledge on the**  
49 **global evolutionary fates** of duplicated genes inferred from bioinformatic studies, there are so far relatively  
50 few functional genetic and phenotypic studies validating these predictions in polyploids as they are  
51 complex, labor and time-consuming. Nevertheless, such studies are essential to unequivocally decipher (i)  
52 why several copies of a given gene are retained over many millions of years (Mya), (ii) if they retained  
53 their ancestral function, (iii) if there exists a dosage effect or a dosage compensation in the case of the loss  
54 of one or several duplicated copies.

55 To fill this gap, we decided to work on the allotetraploid *Brassica napus* species (oilseed rape, AACC,  
56  $2n=4x=38$ ), which has been subjected to two ancient and one recent event of polyploidization. Indeed,  
57 diploid *Brassica* species have been subjected to two old Whole Genome Duplication events (also referred  
58 as a Whole Genome Triplication) that occurred about 22.5 million of years ago (Beilstein et al. 2010). In  
59 addition, the allotetraploid *B. napus* results from the recent hybridization and genome doubling (~7500  
60 years) between the two diploid species *Brassica rapa* (AA,  $2n=2x=20$ ) and *B. oleracea* (CC,  $2n=2x=18$ ),  
61 which diverge about 4 Mya (Chalhoub et al. 2014). Thus, according to the three WGD events that *B. napus*  
62 had been subjected since its divergence with Arabidopsis, it might be expected to present six copies of each  
63 gene derived from the past duplication events. In this paper, we wanted to investigate why gene copies  
64 duplicated by these old or recent polyploidization events are retained by thoroughly studying the duplicated  
65 *petC* gene in *B. napus* and by Knocking Out one to all of its copies using the Crispr-Cas9 technology. We  
66 chose this particular gene (encoding the Rieske protein of the cytochrome b6/f) as its phenotypic effect can  
67 be easily followed. Indeed, previous works realized in Arabidopsis revealed that the K.O. of this gene  
68 suppresses the photosynthetic electron flow and forms white leaves due to the absence of photosynthesis  
69 (not lethal to the plant when adding sucrose to the medium; Maiwald et al. 2003).

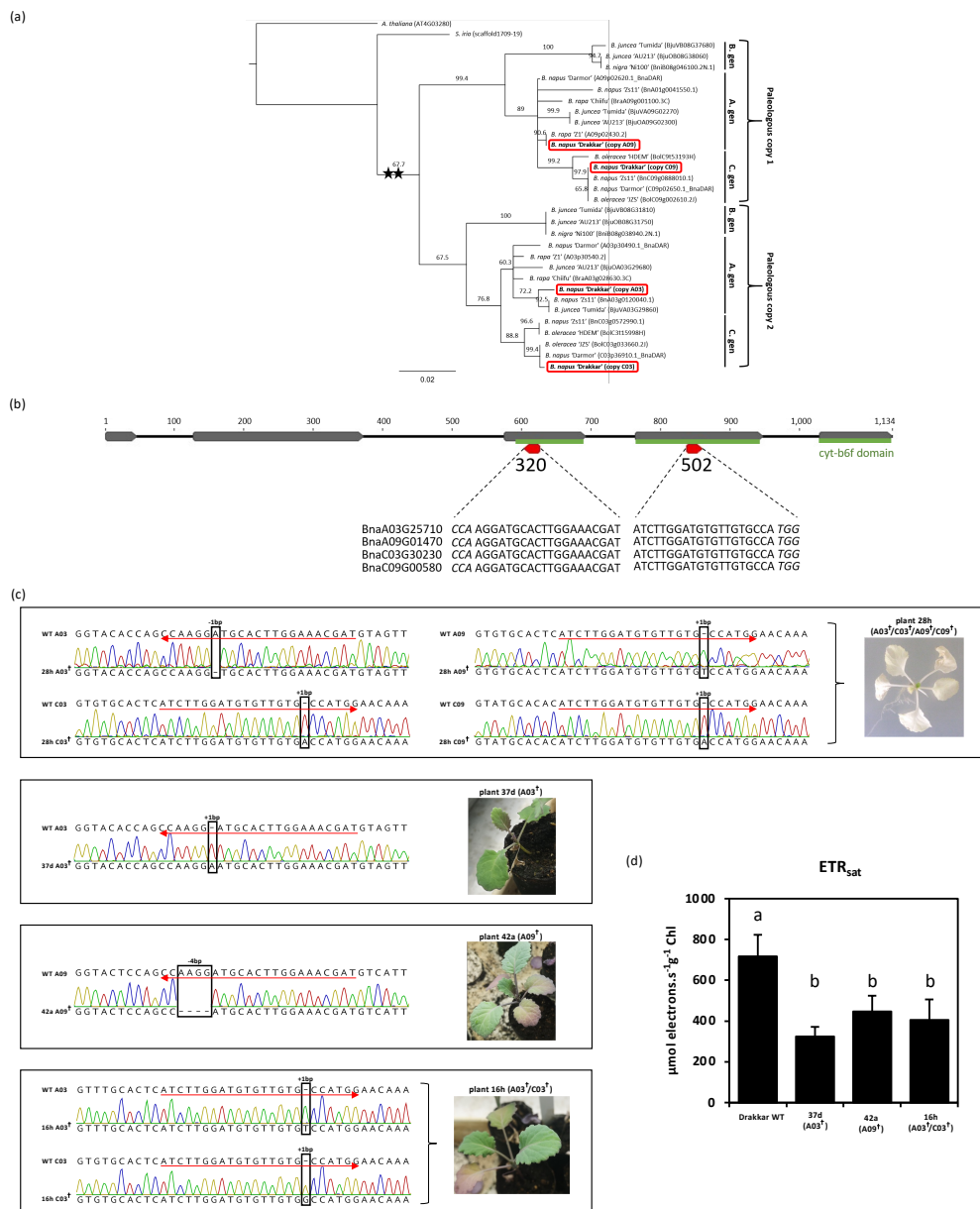
70 To understand the evolutionary dynamic of the *petC* gene in *Brassica*, we first performed a phylogenetic  
71 analysis taking advantage of various Brassicaceae genomes available (Table S1). This allowed to determine  
72 that it was in two copies in *Brassica* diploid species, indicating that one copy had been lost through the  
73 fractionation process, and in four paleologous copies in *B. napus*: BnaA03g25710D, BnaA09g01470D,  
74 BnaC03g30230D, BnaC09g00580D (referred as A03, A09, C03 and C09; Fig. 1a and Table S1). Thereafter,  
75 the analyses of the four *B. napus petC* coding gene sequences suggested that they all are functional as they  
76 all present an intact Open Reading Frame (ORF) and conserved their Rieske functional domain (according  
77 to Pfam: <https://pfam.xfam.org/>). In addition, we determined using PAML4 (Yang, 2007) that all these  
78 copies were under high negative selection (Ka/Ks values never exceeded 0.32; see Table S2), and we showed  
79 that they were all transcribed (Table S3) using RNA Seq data from *B. napus* cv. Darmor leaves (Chalhoub  
80 et al. 2014). These different results highly suggest that all copies conserved their ancestral function and are  
81 still functional.

82 To gain knowledge in the potential reasons that may explain the conservation under high negative selective  
83 pressure of these four paleologous *petC* copies, we used the Crispr-Cas9 technology, as previously realized  
84 in *B. napus* (reviewed in Saad et al. 2021), to perform the K.O. of one, several or all *petC* copies (4 *petC*  
85 paleologous copies corresponding to 8 paleoalleles: 2 alleles per gene copy) and examined their  
86 independent impact on the phenotype. As there is no genome assembly of *B. napus* cv. Drakkar (thereafter

87 referred as Drakkar) that was used for genetic transformation, we firstly developed primer pairs specific of  
88 each copy using *B. napus* cv Darmor reference genome, and then performed PCR, cloning and sequencing  
89 to obtain the different Drakkar *petC* full gene sequences (validated through phylogenetic analyses, see Table  
90 S4 and S5). Thereafter, we identified using CRISPOR (<http://crispor.tefor.net/>) two different gRNA, each  
91 targeting an identical exon region between the four *petC* copies (either exon 3 or 4) and that belong to the  
92 Rieske functional domain (Fig. 1b). We then synthesized a fragment that notably contained the two gRNAs  
93 (under U3 and U6 promoters, respectively, Table S6), and that was integrated by Gateway cloning into  
94 pDeCas9 vector containing the *BlpR* Basta-resistance gene (Addgene plasmid # 61433). This vector was  
95 used to transform 400 Drakkar hypocotyls using an adapted method from Cardoza and Stewart (2003).  
96 After Bialaphos selection, plants were grown on Murashige and Skoog (MS) medium supplemented with  
97 sucrose (30 g/l) to prevent potential photosynthetic deficiencies. Six Bialaphos resistant plants were  
98 regenerated and their transformation was validated by verifying the presence of the Cas9 using PCR. We  
99 then amplified each *petC* copy using a High fidelity Taq polymerase, and performed cloning/sequencing (8  
100 clones per copy, per region surrounding the guide and per transformed plant, totaling 64 clones per tested  
101 plant) to identify the potential Crispr-cas9 induced mutations within the transformed plants.  
102 Overall, five of the six transformed plants presented the K.O. of one to six out of the eight *B. napus* *petC*  
103 alleles (Table S5). Within these T0 plants, some chimerism was observed. We thereafter performed self-  
104 fertilisation of the five T0 mutated plants, and searched for T1 plants without the Cas9 (to prevent the  
105 presence of chimerism) and for which the mutation of one or several gene copies was fixed at the  
106 homozygous stage. From the T0 plant presenting 6 mutated alleles (2 alleles from A03, 1 allele from A09,  
107 2 alleles from C03 and 1 allele from C09, 25 T1 progenies (out of 154) were totally white (Fig. 1c, Table  
108 S5). Cloning and sequencing of several white T1 plants revealed that they were K.O. for all eight alleles,  
109 via the formation of a premature stop codon due to a frameshift within the coding region. We also sequenced  
110 some of the plants harboring green leaves and within these latter, we obtained some triple mutant plants  
111 that only harbored one functional copy (both wild type alleles) of either A09 or C09, highlighting that (i)  
112 the A09 and C09 copies kept their ancestral function, (ii) that the retention of a single functional copy (at  
113 least for A09 and C09) prevents to obtain a white leaves mutant. As these plants still contained the Cas9,  
114 they were not further investigated.  
115 From the other T0 plants (mutation of 1 to 2 *petC* copies), we tested a total of 603 T1, of which nine were  
116 Cas9 free, indicating that the Cas9 had been integrated at at least two different sites. Within these 9 T1 Cas9  
117 free plants, we obtained either single (A03 or A09) or double (A03 and C03) homozygous T1 mutant plants,  
118 which still harbored green leaves (Fig. 1c). To more finely examine the potential impact of the loss of one  
119 (either A03 or A09 copy) or several copies (A03/C03) on the phenotype, we thereafter produced T2 seeds  
120 from these 3 mutants and determined the saturating photosystem II electron transfer rate ( $ETR_{sat}$ ) under  
121 light-saturating conditions (four replicates per genotype) using a *Chlorophyll fluorescence Imaging*  
122 “Fluorcam” (Photon system Instruments) (Fig. S1), enabling to compare their maximal electron transfer  
123 efficiency. These  $ETR_{sat}$  values were normalized to chlorophyll content to account for possible variations  
124 independent from the mutations. Overall, the results showed significant differences (ANOVA *p-value*  
125 0.000528) between wild-type (718  $\mu\text{mol electrons}\cdot\text{s}^{-1}\cdot\text{g}^{-1}$  Chl) and mutated plants (324, 347 and 404  $\mu\text{mol}$   
126  $\text{electrons}\cdot\text{s}^{-1}\cdot\text{g}^{-1}$  Chl) (Fig. 1d), highlighting that (i) the A03 and A09 copies kept the ancestral *petC* function,  
127 (ii) but also that the phenotypic impact of the loss of these copies can only be visible under light-saturating  
128 and not under normal growth chamber conditions (low light irradiance). Interestingly, no statistical  
129 difference was observed between the single paleologous (A03 or A09) mutant plants (*post-hoc* Tukey HSD  
130 *p-value* 0.223), despite the A09 copy showed about a twice (1.74 times) more important transcription level

131 than the A03 copy. Similarly, we did not observe any significant difference between the double  
 132 homoeologous mutants (for A03 and C03) and the single mutant plants (*post-hoc* Tukey HSD *p-value*  
 133 0.549).

134 Overall, the work described here enabled to validate experimentally that old paleologous *petC* copies kept  
 135 their ancestral function (only inferred for C03) since 22.5 millions of years. It also revealed that the very  
 136 long-term conservation of functional *petC* paleologous copies under high negative selective pressure can  
 137 be explained by their compensatory role under optimal conditions. This work also highlighted the  
 138 importance to K.O. all paleologous gene copies to totally suppress a major gene function in *B. napus*.



139  
 140 **Figure 1** Phenotypic impact of the Knock Out of one to all four paleologous *petC* copies in *B. napus*. (a)  
 141 Copy number and origin of *petC* gene copies in *Brassicaceae* using Maximum likelihood phylogeny. (b)  
 142 Structure of the whole *petC* gene (grey rectangles: exons; thick black lines: introns) and position of the two  
 143 guideRNA (red arrows) targeting the cytb6f domain (green rectangles) of all four *petC* copies. (c) Fixed  
 144 mutations of different *petC* copies: plants 28H (quadruple mutant: K.O. of all four copies), 37d (A03

145 mutant), 42a (A09 mutant), 16h (A03/C03 mutant). (d) Comparison of the photosynthesis capacity (through  
146 ETRsat) in single and double *petc* mutants compared to a wild type Drakkar plant (\*: p-value<0.05).

#### 147 **Funding**

148 This work was funded by the Investissement d'Avenir program of the French National Agency of Research  
149 for the project GENIUS (ANR-11-BTBR-0001\_GENIUS) as well as from the ARN ChromCO (ANR-19-  
150 CE20-0018), which paid M. Facon salary.

#### 151 **Competing interests**

152 The authors declare no conflict of interest. Genetically engineered plants were conducted in accordance  
153 with national legislation.

#### 154 **Authors contribution**

155 MRG and AMC conceived the study. MRG performed the phylogeny. MRG and HA designed the gRNA.  
156 MLT, HA, and EM performed the genetic transformation, and produced progenies. GD performed the  
157 molecular biology experiments and sequencing analyses. MF, and YD performed the Chlorophyll  
158 Fluorescence imaging experiment. LM analysed RNA Seq data. MF and MRG wrote the manuscript.

#### 159 **References**

- 160 Beilstein, M. A., Nagalingum, N. S., Clements, M. D., Manchester, S. R., and Mathews, S. (2010) Dated  
161 molecular phylogenies indicate a Miocene origin for *Arabidopsis thaliana*. *PNAS USA*. **107**, 18724-18728.  
162 Birchler, J. A., and Yang, H (2022) The multiple fates of gene duplications: deletion, hypofunctionalization,  
163 subfunctionalization, neofunctionalization, dosage balance constraints, and neutral variation. *Plant Cell*.  
164 **34**, 2466-2474.  
165 Cardoza, V., and Stewart, C. N. (2003) Increased *Agrobacterium*-mediated transformation and rooting  
166 efficiencies in canola (*Brassica napus* L.) from hypocotyl segment explants. *Plant Cell Rep.* **21**, 599-604.  
167 Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., *et al.* (2014) Early allopolyploid evolution in  
168 the post-neolithic *Brassica napus* oilseed genome. *Science*. **345**, 950-953.  
169 Maiwald, D., Dietzmann, A., Jahns, P., Pesaresi, P., Joliot, P., *et al.* (2003) Knock-out of the genes coding  
170 for the rieske protein and the ATP-synthase  $\delta$ -subunit of *Arabidopsis*. Effects on photosynthesis, thylakoid  
171 protein composition, and nuclear chloroplast gene expression. *Plant Physiology* **133**, 191-202.  
172 Saad, N. S. M., Severn-Ellis, A., Pradhan, A., Edwards, D., Batley, J. (2021) Genomics armed with diversity  
173 leads the way in Brassica improvement in a changing global environment *Frontiers in Genetics* **12**, 600789.  
174 Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586-1591.

#### 175 **Supporting Figure**

176 **Figure S1.** Light and time-dependent saturation of PSII Electron Transfer Rate.

#### 177 **Supporting tables**

178 **Table S1:** Accession numbers used for phylogeny; **Table S2:** Evaluation of dN/dS using PAML4 (Yang,  
179 2007); **Table S3:** Transcription values of the four *petC* gene copies in *B. napus* cv Darmor leaves; **Table**  
180 **S4:** Drakkar *petC* sequences; **Table S5:** List of primer pairs; **Table S6:** Synthesized fragment inserted in  
181 pDONR207; **Table S7:** Mutations identified in T0 transformed plant.

182 **Supporting experimental procedures:** (i) Identification and evolution of the Brassicaceae *petC* genes; (ii)  
183 Selective pressure acting on the four *petC* nuclear genes; (iii) Evaluation of the four *petC* copies  
184 transcription level using RNA Seq data; (iv) Amplification, cloning and sequencing of the *B. napus* cv.  
185 Drakkar *petC* gene copies; (v) Design of a pair of guides simultaneously targeting all *B. napus* cv. Drakkar  
186 *petc* copies; (vi) Vector Construction; (vii) Plant transformation; (viii) Identification of putative mutated

187 copies in T0 transformed plants; (ix) Production of homozygous mutant plants for one or several copies  
188 and Cas9 free; (x) Photosystem II fluorescence measurements.