The Plant Cell, Vol. 24: 4875–4891, December 2012, www.plantcell.org

A Dominant Point Mutation in a RINGv E3 Ubiquitin Ligase Homoeologous Gene Leads to Cleistogamy in *Brassica napus*[C][W][OA]

Yun-Hai Lu,[a] Dominique Arnaud,[a] Harry Belcram,[a] Cyril Falentin,[b] Patricia Rouault,[b] Nathalie Piel,[b] Marie-Odile Lucas,[b] Jérémy Just,[a] Michel Renard,[b] Régine Delourme,[b] and Boulos Chalhoub[a,1]

[a] Unité de Recherche en Génomique Végétale (Institut National de la Recherche Agronomique, Centre National de la Recherche Scientifique, Université d'Evry Val d'Essonnes), Organization and Evolution of Plant Genomes, 91057 Evry cedex, France
[b] Institut National de la Recherche Agronomique, Unité Mixte de Recherche 1349 Institut de Génétique, Environnement et Protection des Plantes, F-35653 Le Rheu, France

In the allopolyploid *Brassica napus*, we obtained a petal-closed flower mutation by ethyl methanesulfonate mutagenesis. Here, we report cloning and characterization of the Bn-*CLG1A* (*CLG* for cleistogamy) gene and the Bn-*clg1A-1D* mutant allele responsible for the cleistogamy phenotype. Bn-*CLG1A* encodes a RINGv E3 ubiquitin ligase that is highly conserved across eukaryotes. In the Bn-*clg1A-1D* mutant allele, a C-to-T transition converts a Pro at position 325 to a Leu (P325L), causing a dominant mutation leading to cleistogamy. *B. napus* and *Arabidopsis thaliana* plants transformed with a Bn-*clg1A-1D* allele show cleistogamous flowers, and characterization of these flowers suggests that the Bn-*clg1A-1D* mutation causes a pronounced negative regulation of cutin biosynthesis or loading and affects elongation or differentiation of petal and sepal cells. This results in an inhibition or a delay of petal development, leading to folded petals. A homoeologous gene (Bn-*CLG1C*), which shows 99.5% amino acid identity and is also constitutively and equally expressed to the wild-type Bn-*CLG1A* gene, was also identified. We showed that P325L is not a loss-of-function mutation and did not affect expression of Bn-*clg1A-1D* or Bn-*CLG1C*. Our findings suggest that P325L is a gain-of-function semidominant mutation, which led to either hyper- or neofunctionalization of a redundant homoeologous gene.

INTRODUCTION

The rapid explosion in diversity that followed the emergence of flowering plants in the early Cretaceous, ~130 million years ago, may be linked to the evolution of their reproductive organs, the flowers (Carroll, 2001). Cleistogamy, characterized by closed flowers that are self-pollinated and can produce fruits and seeds as a result of autogamy, contrasts with chasmogamy, characterized by opened flowers that can be cross-pollinated and produce fruits and seeds by allogamy (Uphof, 1938; Lord, 1981; Culley and Klooster, 2007). Darwin (1877) first reported cleistogamy as a common phenomenon widely distributed among the angiosperms and likely to have evolved from chasmogamy to ensure seed set by selfing. Cleistogamy is widespread and has been reported in at least 693 angiosperm species from 228 genera and 50 families (Culley and Klooster, 2007). There is no unique feature of cleistogamy, but it can be classified into three categories (reviewed in Culley and Klooster, 2007): dimorphic cleistogamy, in which prominent floral differences in cleistogamy and chasmogamy floral morphology result from divergent developmental pathways; complete cleistogamy, in which individual plants produce only cleistogamous flowers; and induced cleistogamy, in which the environment arrests the development of chasmogamous flowers prior to anthesis and results in a mechanical failure of the flowers to open.

Although cleistogamy leads to inbreeding depression, it also may be advantageous for plants (reviewed in Culley and Klooster, 2007). The advantages include reproductive assurance (e.g., when pollinators are rare) and reduced resources needed for reproduction (Schemske, 1978; Waller, 1984). From a practical point of view, the cleistogamy trait could be useful in maintaining genetic purity (Saxena et al., 1993) and in developing genetically modified (GM) cultivars with low risk of gene flow to non-GM varieties (Kwon et al., 2001; Daniell, 2002; Lu, 2003).

Genetic control of cleistogamy has been studied in several species. Lord (1981) suggested that the chasmogamy/cleistogamy might be under the control of polygenic inheritance and potentially readily modified by the environment in some studied species. Two or more independent genes were found to be responsible for cleistogamy in sorghum (*Sorghum bicolor*) (Merwin et al., 1981), soybean (*Glycine max*) (Takahashi et al., 2001), and barley (*Hordeum vulgare*) (Kurauchi et al., 1993; Turuspekov et al., 2004). Single recessive genes were found to be responsible for the cleistogamous trait in pigeon pea (*Cajanus cajan*; Saxena et al., 1992, 1993), durum wheat (*Triticum durum*; Chhabra and Sethi, 1991), and rice (*Oryza sativa*; Maeng et al., 2006). Recently, the

[1] Address correspondence to chalhoub@evry.inra.fr.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Boulos Chalhoub (chalhoub@evry.inra.fr).
[C] Some figures in this article are displayed in color online but in black and white in the print edition.
[W] Online version contains Web-only data.
[OA] Open Access articles can be viewed online without a subscription.
www.plantcell.org/cgi/doi/10.1105/tpc.112.104315

barley *Cleistogamy1* gene (*Cly1*) was isolated by positional cloning (Nair et al., 2010). *Cly1* was found to be a homolog of the *Arabidopsis thaliana* transcription factor *APETALA2*, and the cleistogamous flowering phenotype was suggested to be caused by a mutation at the microRNA (miR172) targeting site that suppressed miR172-guided mRNA cleavage (Nair et al., 2010).

Oilseed rape (*Brassica napus*) is an economically important crop, cultivated worldwide for human consumption (third most important vegetable oil source after soybean and palm oil), animal feeding, and biodiesel. The genome of oilseed rape (AACC, 2n = 38) is a recent amphitetraploid, derived from interspecific hybridization between *Brassica oleracea* (CC, 2n = 18) and *Brassica rapa* (AA, 2n = 20). The vast majority of cultivated oilseed rape cultivars are self-compatible and allogamous, with an outcrossing rate ranging between 12 and 55% (average 30%), depending on the genotype and environmental conditions (Beckie et al., 2003), although both of its diploid progenitors are self-incompatible (Silva and Goring, 2001). *B. napus* can disseminate its pollen as well as receive pollen from quite distantly related plants, including wild related species (Jørgensen and Andersen, 1994; Halfhill et al., 2002; Warwick et al., 2003; Ammitzbøll and Bagger Jørgensen, 2006; Ford et al., 2006; Song et al., 2010).

An oilseed rape mutant phenotype with flowers that do not open, leading to cleistogamy, was obtained in the early seventies by ethyl methanesulfonate (EMS) mutagenesis of seeds of the 'Primor' cultivar (Renard and Tanguy, 1997; Figure 1). The mutant plants 'Primor-*Clg*' (*Clg* stands for cleistogamy) develop and reproduce normally while having closed flowers (Figures 1B and 1E). Cleistogamous oilseed rape lines bred from this mutant were shown to have an autogamy rate as high as 94% and to emit 10 times less pollen than an open-flowered oilseed rape in the same growing conditions (Fargue et al., 2006). No differences in seed production were observed between cliestogamous mutants and wild-type plants, indicating that fertility is not affected in cleistogamous oilseed rape (Renard and Tanguy, 1997; Fargue et al., 2006). This induced cleistogamous trait would be useful in oilseed rape breeding as it decreases the risk of genetic contamination and gene flow from GM varieties through outcrossing to other oilseed rape cultivars or to related *Brassica* species.

Here, we report the positional cloning and characterization of the gene responsible for this trait. We show that a single amino acid mutation in a highly conserved region of a RINGv E3 ubiquitin ligase homoeologous gene copy led to the cleistogamy phenotype in oilseed rape.

RESULTS

Phenotyping, Genetic Mapping, and Positional Cloning of the Cleistogamy Gene in *B. napus*

Cleistogamous oilseed rape showed a normal development compared with noncleistogamous plants. Cleistogamous flowers did not normally open at anthesis and remained closed, appearing like a big yellow bud until petal abscission (Figure 1). Other floral organs, such as carpel and stamens, develop normally. Notably, the cleistogamous character was shown to be variable where, depending on the genotype tested and environmental conditions, some flowers partially open, either showing a hole at the top of the bud permitting a partial extrusion of stamens

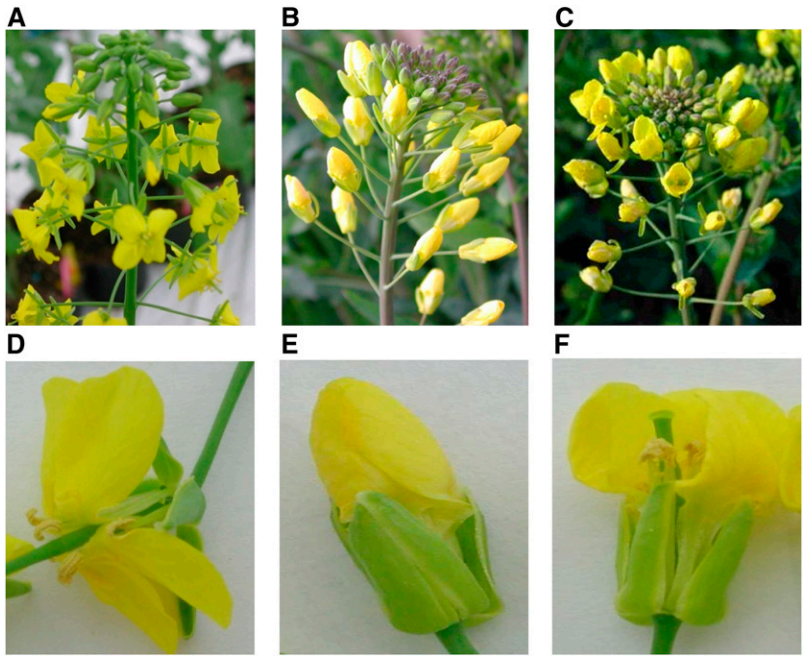


Figure 1. Oilseed rape (*Brassica napus*) flower morphology.

Normal noncleistogamous plants with open petals of cv Yudal (**[A]** and **[D]**), cleistogamous plants with closed petals of line B001-*Clg* (**[B]** and **[E]**), and intermediate-closed petals F1 plants (**[C]** and **[F]**).

or having one petal laterally malformed creating a slight gap between petals (Fargue et al., 2006; Figure 1; see Supplemental Figure 1 online). The cleistogamy trait has been successfully introgressed in several varieties where several pairs of cleistogamous/non-cleistogamous near-isogenic lines were produced (Renard and Tanguy, 1997).

Early genetic investigations showed that the cleistogamy phenotype in *B. napus* is controlled by a major locus called *Clg1* (Renard and Tanguy, 1997). F1 heterozygous plants show an intermediate phenotype with partially closed flowers (Renard and Tanguy, 1997; Figure 1C). In this study, we followed the international *Arabidopsis* and *Brassica* gene nomenclature and named the wild-type allele *CLG1A* and the mutant allele leading to cleistogamy as *clg1A-1D.* The letter A designates the genome on which the gene was mapped and the letter D the semidominant nature of the mutation. The prefixes Bn-, Br-, or Bo- were added to designate species names (*B. nap*us, *B. rapa*, and *B. oleracea*, respectively) to avoid confusion when describing the different homologous copies of the gene (see below). We initiated mapping of Bn-*CLG1A*/Bn-*clglA-1D* using a segregating population of double haploid (DH) lines of a cross between the French cleistogamous line 'B001-*Clg*' and the genetically distant Korean cultivar 'Yudal.' The different steps of positional cloning are summarized in Figure 2 and detailed in Supplemental Results 1 online.

Briefly, the phenotyping of an initial population of 255 DH lines showed a segregation of 53% of noncleistogamous DH lines to 47% intermediate-cleistogamous and cleistogamous ones (see Supplemental Figure 1 online), fitting a monogenic model with the Bn-*clg1A-1D* allele leading to cleistogamy or partial cleistogamy. Initial molecular mapping was done with amplified fragment length polymorphism (AFLP) (Vos et al., 1995; Chalhoub et al., 1997) and allowed identification of five linked markers (Figure 2A). Sequencing of these AFLP markers and comparison with *Arabidopsis* genome sequence allow identification of a putative *Arabidopsis* orthologous genomic region on chromosome IV, based on high sequence homology of one AFLP marker sequence to the *At4g33760* gene of *Arabidopsis*. This was followed

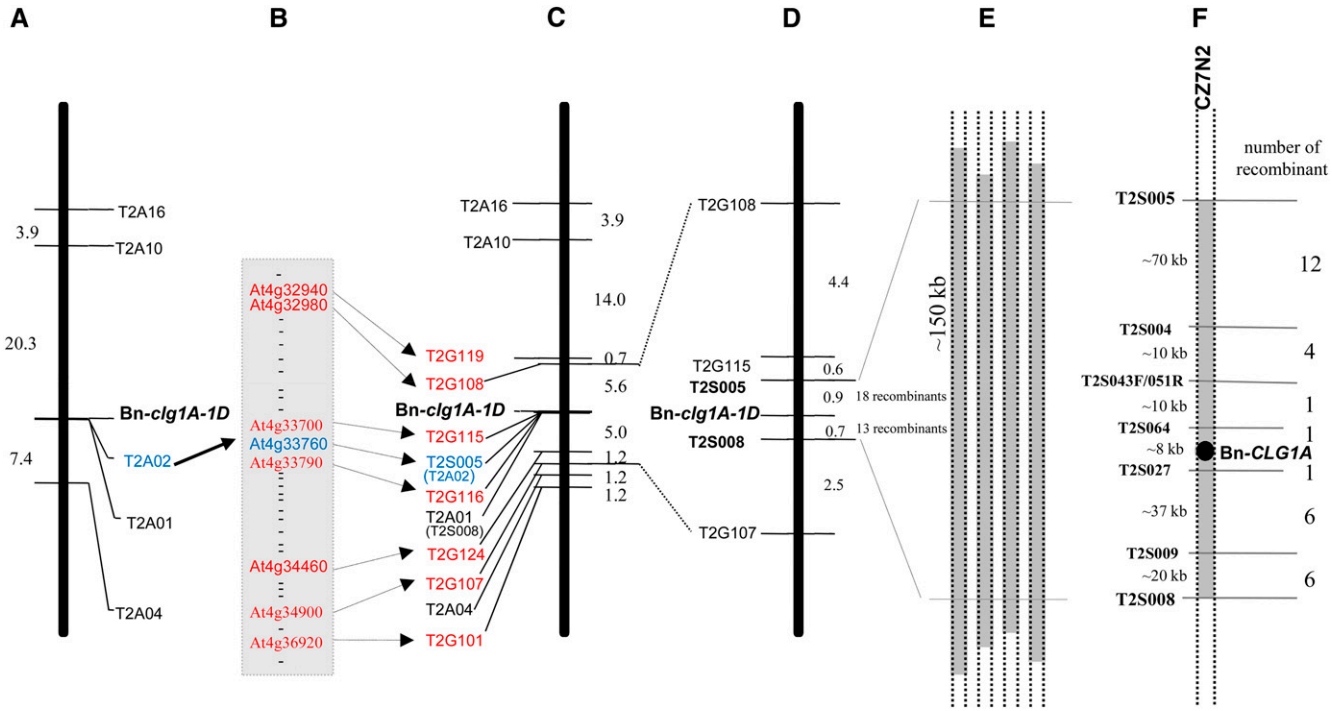


Figure 2. Mapping and Positional Cloning of Bn-*CLG1A*.

(A) Initial mapping of five AFLP markers on a population of 255 segregant DH lines derived from a cross between B001-*Clg* and cv Yudal. The genetic distance (centimorgans) between markers is shown on the left.

(B) Thirty-two *Arabidopsis* genes from a region of 2030 kb of chromosome IV around the *At4g33760* gene were used for generation of additional PCR markers.

(C) Seven genes (in addition to *At4g33760* gene) produced polymorphic PCR markers and were mapped around Bn-*clg1A-1D* on the same 255 DH segregating population.

(D) Finer genetic mapping of Bn-*clg1A-1D* on a larger DH population of 2158 lines.

(E) Identification of four BAC clones from Darmor-*bzh* BAC library, containing the closest flanking genetic markers: T2S005 and T2S008 and Bn-*CLG1A*.

(F) The BAC clone CZ7N2 was sequenced (237,660 bp) and shows that T2S005 and T2S008 delimit a region of 150 kb, whose sequence were used to generate five additional PCR markers. These were used to precisely locate Bn-*CLG1A*, using the 31 lines that show recombination between T2S005 and T2S008, to an ~8-kb interval containing a single candidate gene, encoding a putative RINGv E3 ubiquitin ligase.

by the use of 32 *Arabidopsis* gene sequences from this region to generate and map additional PCR markers (Figures 2B and 2C). Finer genetic mapping of Bn-*clg1A-1D* on a larger DH population of 2158 lines was then used to identify closely flanking markers (Figure 2D). The closest flanking genetic markers (T2S005 and T2S008) were used to screen a Darmor-*bzh* BAC library and four BAC clones that contain both markers, and thus the Bn-*CLGA1* wild-type allele, were identified (Figure 2E). One BAC clone (CZ7N2) was sequenced (237,660 bp) and compared with the *Arabidopsis* chromosome IV orthologous regions (See Supplemental Figure 2 and Supplemental Table 2 online). Analysis of the CZ7N2 BAC clone sequence showed that T2S005 and T2S008 delimit a region of 150 kb that was then used to generate five additional PCR genetic markers (Figure 2F). Mapping of these on 31 DH lines, which had a recombination between T2S005 and T2S008, delimited the Bn-*CLG1A* candidate region to an ~8-kb interval containing a single candidate gene that encodes a putative RINGv E3 ubiquitin ligase (Figure 2F).

Identification of the Causal Mutation by Comparison of the Mutant and Wild-Type Sequences

Sequence comparison of the 8-kb region spanning the Bn-*CLG1A* candidate gene between the wild and the EMS mutant cleistogamous lines was used to identify the causal mutation. As the *B. napus* genome is highly duplicated, to avoid artifacts related to simultaneous PCR amplification of homologous fragments and to directly access the genomic region of the mutant allele *Bn-clg1A-1D*, we constructed a pooled BAC library of the line 'Primor-*Clg*' as described by Isidore et al. (2005). Screening of the library with PCR markers linked to Bn-*clg1A-1D* identified three BAC clones that contain this mutant allele. After verification, one of the *clg1A*-carrying BAC clones (Clg_H1B_P13a_3_J1) was sequenced. The original *B. napus* cv Primor containing the wild-type Bn-*CLG1A* allele was used in the pedigree of the cv Darmor-*bzh* for which we have sequenced the BAC clone CZ7N2. Direct sequence comparison of Bn-*CLG1A* and *Bn-clg1A-1D* (CZ7N2 from Darmor-*bzh* and Clg_H1B_P13a_3_J1 from 'Primor-*Clg*,' respectively) revealed three single nucleotide polymorphisms (SNPs) in a region of ~85 kb (see Supplemental Figure 3 online). Two of these were found in intergenic noncoding regions at ~10 kb (C-to-T) and ~35 kb (G-to-A), 5′- of the Bn-*CLG1A* candidate sequence (see Supplemental Figure 3 online). The third SNP mutation (C-to-T) was found inside the Bn-*CLG1A* candidate sequence, delimited by the closest flanking markers, T2S064 and T2S027. This SNP mutation occurs in the second exon of the predicted RING-finger protein gene and causes a substitution of Pro with a Leu at the position 325 (hereafter called P325L) in the protein sequence (see Supplemental Figure 3 online). Thus, comparative sequencing suggests that the P325L mutation in the gene encoding a putative RING-finger protein is the causal mutation leading to the cleistogamy trait.

Comapping between the Causal SNP and the Cleistogamy Phenotype

To confirm that the P325L mutation is responsible for the cleistogamous phenotype, we developed a cleaved amplified polymorphic sequence (CAPS) marker (dCAPS_*Clg1*) which allows a clear distinction between wild-type and mutant alleles (see Supplemental Figure 4 online). The dCAPS_*Clg1* marker was completely linked to the cleistogamous phenotype with zero recombinant plants when analyzing the 'B001-*Clg*' × 'Yudal' DH lines segregating population of 255 lines and the 31 recombinant lines revealed from the larger 2158 segregating DH population. In particular, noncleistogamous DH lines (note 1) were shown to have the Bn-*CLG1A*/Bn-*CLG1A* genotype, whereas intermediate cleistogamous (noted 2 and 3) and cleistogamous DH lines (noted 4 and 5) have the genotype Bn-*clg1A*/Bn-*clg1A.* The dCAPS_*Clg1* marker also revealed a perfect association between the SNP type (T/C) and the corresponding cleistogamy phenotype on 19 pairs of near-isogenic lines for cleistogamy, as well as on 20 *B. napus* varieties (see Supplemental Figure 4B and Supplemental Table 3 online). No oilseed rape lines, in which cleistogamy trait was not introduced, were shown to carry this mutation (data not shown).

Bn-*CLG1A* Gene Structure

Sequence analysis predicted that Bn-*CLG1A* is 4144 bp in size and composed of nine exons interspersed by eight introns (Figure 3A). The predicted exon-intron structure was experimentally supported by sequencing of RT-PCR products obtained from flower buds as well as by the *B. napus* EST sequence data available in the National Center for Biotechnology Information public database.

The putative encoded protein is 1110 amino acids and has a predicted size of 123.05 kD (Figure 3A). Analysis of the Bn-CLG1A predicted protein sequence revealed the presence of several functional domains (Figure 3B): a RINGv protein domain at the N terminus (72 to 120, InterPro IPR011016); an adenylate cyclase-associated CAP_1 signature (293 to 305, InterPro IPR001837) located in a coiled-coil domain (294 to 321); a Leu zipper domain (800 to 821); and 14 transmembrane regions/helices (163 to 183, 202 to 222, 352 to 372, 464 to 484, 529 to 549, 575 to 595, 633 to 651, 703 to 723, 796 to 816, 844 to 864, 885 to 914, 928 to 948, 983 to 1003, and 1017 to 1037). The Bn-CLG1A protein sequence exhibits feature similarities with previously described RINGv E3 ubiquitin ligase protein in *Arabidopsis* (Kraft et al., 2005; Stone et al., 2005; Vierstra, 2009). A RING domain is required in interactions with substrate protein for E3 ligase activity by binding E2 (Lovering et al., 1993; Borden and Freemont, 1996; Borden, 2000). The encoded protein also contains other diverse protein–protein interaction domains that can serve as additional specific substrate binding sites of the E3 ligase. Internal domains, like the coiled-coil domain/CAP_1 signature and Leu zipper domain (Figure 3B), might function as a substrate binding domain of the Bn-CLG1A E3 ligase for specific substrate/target recognition (Lorick et al., 1999; Smalle and Vierstra, 2004). The point mutation at position 325 in the Bn-clg1A protein is located near the coiled-coil domain/CAP_1 domain and could thus modify the specificity of the substrate protein recognition or cofactor binding of the Bn-clg1A protein.

Further analysis of the predicted protein sequences (see Methods) shows that Bn-CLG1A is predicted to localize in the plasma membrane (WoLFPSORT and MultiLoc) and in plastids/chloroplasts (TargetP1.1, ChloroP 1.1, and TargetLoc), with

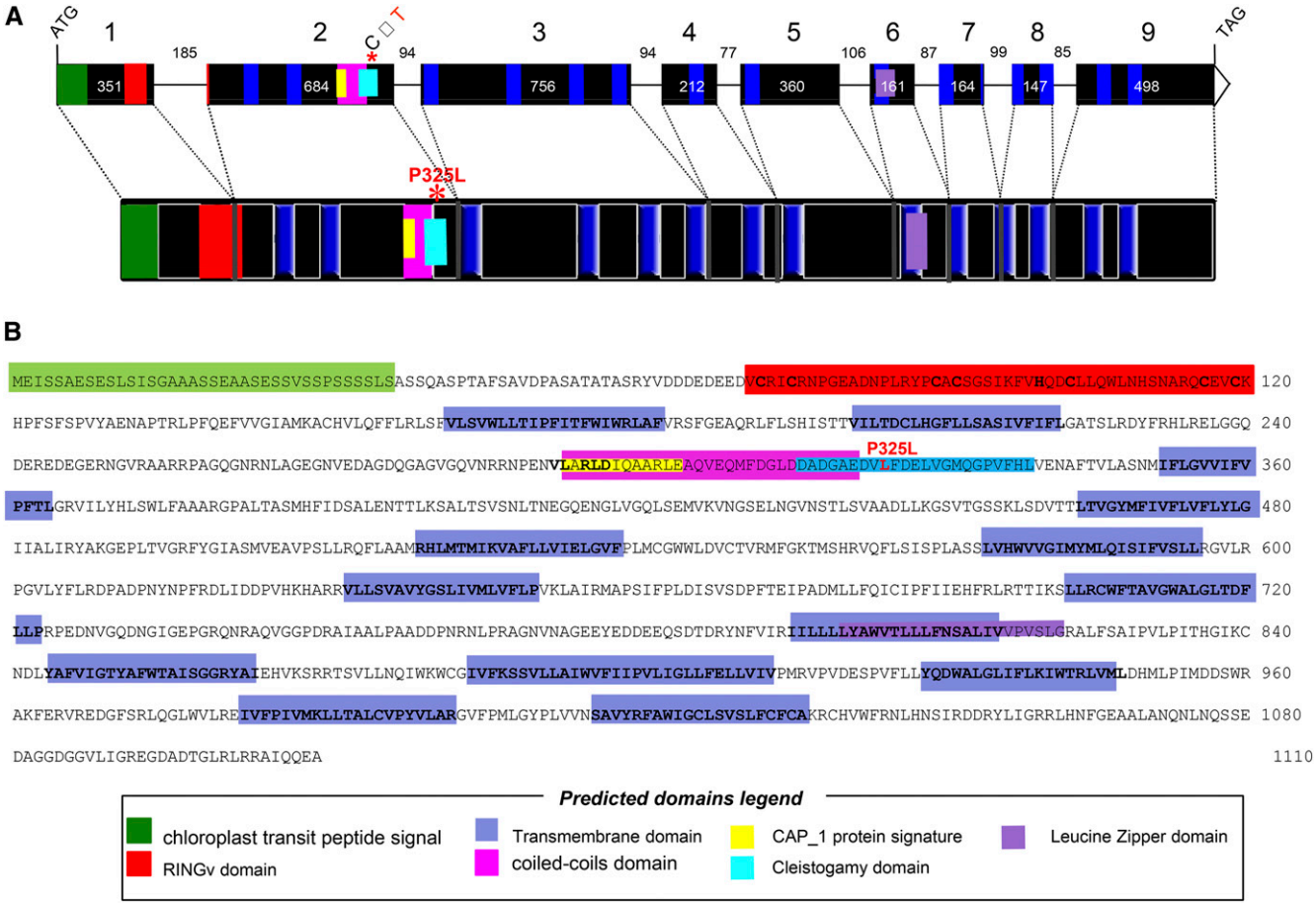


Figure 3. *B. napus* Cleistogamy (Bn-*CLG1A*) Gene Structure.

(A) Exon/intron structure of Bn-*CLG1A* with nine exons (black boxes) and eight introns with their corresponding sizes (in base pairs).
(B) Deduced *Bn-CLG1A* mutant amino acid sequence.
The different domains predicted by comparison with structure of homologs from different species are highlighted by same colored boxes in **(A)** and **(B)**, and their legend is presented in the "predicted domains" box. The P325L cleistogamy mutation is highlighted in red.

a typical chloroplastic Ser-rich domain (von Heijne et al., 1989; Zybailov et al., 2008) at its N terminus (Figures 3A and 3B).

Identification of a Highly Conserved Redundant Homoeolog in *B. napus* and Identical Corresponding Orthologs in the Parental Species

Because *Brassica* species genomes are highly duplicated through recurrent polyploidy, it was important to understand the evolution of Bn-*CLG1A* duplicated copies for a better functional characterization.

The predicted Bn-CLG1A protein shares a 90.5% amino acid identity and has similar intron/exon structure with the protein encoded by the *Arabidopsis* orthologous gene *At4g34100* (Figure 4), also predicted to be a RINGv E3 ubiquitin ligase family member (Stone et al., 2005). Interestingly, a recent study demonstrated that a mutant form of *At4g34100* corresponds to the wax mutant *eceriferum9*, which confers drought tolerance to *Arabidopsis* (Koornneef et al., 1989; Lü et al., 2012) (see Discussion). Analysis of the *Arabidopsis* genome did not identify another homolog of *At4g34100* on the *Arabidopsis* region that resulted from the last alpha genome duplication (Schranz and Mitchell-Olds, 2006) located on chromosome 2. A highly divergent paralogous copy (*At4g32670*), sharing only 23% amino acid similarity, was identified on chromosome 4, located at 576 kb from *At4g34100*. These two gene copies are both expressed in *Arabidopsis* (EST available in public databases). They both possess a conserved predicted RINGv domain at their N-terminal region but present different predicted membrane topology patterns. Like At4g34100, the At4g32670 protein is predicted to be localized at the plasma membrane, but it does not contain a chloroplast targeting signal like At4g34100 at its N terminus. The At4g32670 protein is predicted by TargetP1.1 to have no chloroplast and no mitochondrial localization. So the two proteins should have diverged functions/roles in the cell and might have different proteins/substrates as targets for degradation.

Brassica species have undergone genome triplication, shortly after the divergence from *Arabidopsis*. Accordingly, compared with *Arabidopsis*, three homologous regions are expected in the genome of each of the *Brassica* diploid species (for example, *B. rapa* and *B. oleracea*) and up to six in the amphiploid *B. napus*

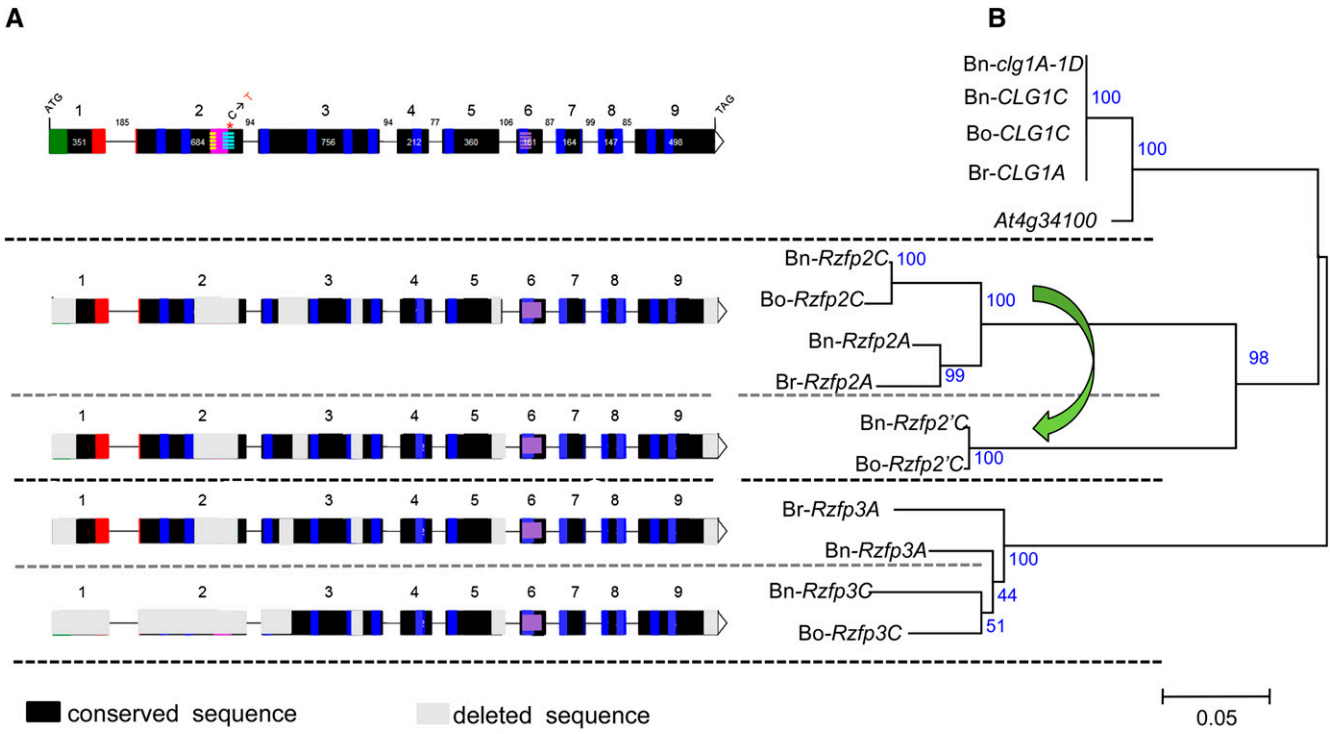


Figure 4. Gene Structure Comparison and Phylogenetic Analysis between Bn-*CLG1A* Homologs Identified in *B. napus*, *B. oleracea*, *B. rapa*, and *Arabidopsis*.

Gene structure comparison **(A)** and phylogenetic analysis **(B)**. Deduced amino acid sequences of these homologs were aligned with ClustalX (Thompson et al., 1997). The phylogeny analysis was performed with MEGA version 4.1 using the neighbor-joining method (Tamura et al., 2007). The percentage of replicates trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The legend of gene structure in **(A)** is the same as in Figure 3. The green arrow indicates tandem duplication of the *Rzfp2C* gene copy into *Rzfp2'C* in the C genome of *B. napus* (Bn-*Rzfp2'C*) and *B. oleracea* (Bo-*Rzfp2'C*). Rzfp, RING zinc finger protein.

(Parkin et al., 2005; Lysak et al., 2007). Direct comparison of the Bn-*CLG1A* sequence with the recently released genome sequence of *B. rapa* (Wang et al., 2011) allows identification of three homologous copies that resulted from the *Brassica* genome triplication (see Supplemental Figure 5A online). In comparison, four copies were found in the complete genome assembly of *B. oleracea* (http://brassicadb.org/brad/). Three of these result from the *Brassica* genome triplication and were orthologous to those of *B. rapa*, whereas the fourth copy located at 31 kb from one of the *B. oleracea* triplicated copies is derived from tandem duplication that occurred only in *B. oleracea* (Figure 4). As expected, seven copies were identified in *B. napus*, three on the A genome, orthologous to those identified in *B. rapa* and four on the C genomes, orthologous to those identified in *B. oleracea*.

Of these different *Brassica* homologs, only the orthologous copies from *B. rapa* (Br-*CLG1A*) and *B. oleracea* (Bo-*CLG1C*) and the homoeologous copy found on the C genomes of *B. napus* (Bn-*CLG1C*), are highly conserved where all protein domains described above can be found (Figure 4A). The *B. rapa* Br-*CLG1A* orthologous copy shows identical amino acid sequences to the Bn-*CLG1A* wild-type gene on the A genome of *B. napus* (Figure 4B). Similarly, the *B. oleracea* Bo-*CLG1C* copy shows identical amino acid sequences to the Bn-*CLG1C* orthologous copy identified on the C genome of *B. napus* (Figure 4B). Comparisons between the A and C genomes show also that Bn-*CLG1A*/Br-*CLG1A* and Bn-*CLG1C*/Bo-*CLG1C* orthologous and homoeologous genes are highly similar (99.5 amino acid identity) (Figure 4B), with only two amino acid deletions and three other transitions in the C genome copies compared with the A ones. The other homologous copies were truncated in several of their exons and protein domains (Figure 4A), indicating important gene fractionations that have followed *Brassica* species genome triplication as previously reported (Wang et al., 2011; Cheng et al., 2012; Tang et al., 2012) and were grouped far from the Bn-*CLG1A*/Bo-*CLG1A* and Bn-*CLG1C*/Bo-*CLG1C* genes and their *Arabidopsis At4g34100* ortholog (Figure 4B). Importantly, the cleistogamy domain found in Bn-*CLG1A* and its homoeologous and orthologous genes were also deleted in the remaining homologous copies (Figure 4A). Therefore, they were designated as RING zinc finger protein (*Rzfp*) coding gene homologs (Figure 4). The close evolutionary relationship of *Arabidopsis* orthologous genes with one of the *Brassica* homologs, derived from the genome triplication, which was closer than between the homologs themselves, was previously reported (Wang et al., 2011; Cheng et al., 2012; Tang et al., 2012). It is important to note that this close relationship between *Arabidopsis* and one of the three *Brassica* subgenomes does not constantly favor a specific subgenome, but rather switches among the three

subgenomes, depending on the considered gene, as we showed in the comparison of five genes that surround Br-*CLG1A* in *B. rapa*, which are conserved as triplicates in *B. rapa* and are syntenic to *Arabidopsis* (see Supplemental Figure 5 online). Notably, this close relationship is relative because comparisons show that overall triplicated genes from *Brassica* and their *Arabidopsis* orthologs remain closer to each other than to Bn-*CLG1A* homologs from any other given species (data not shown). These observations may confirm that, in comparison to *Arabidopsis*, the three *Brassica* subgenomes reunited together through two steps of ancient allopolyploidy (Wang et al., 2011) would have undergone an accelerated and divergent fractionation of redundant genes, as suggested in previous studies (Wang et al., 2011; Cheng et al., 2012; Tang et al., 2012).

Bn-*CLG1A* Homologs in Other Plant and Eukaryote Species

In addition to *Arabidopsis* and other *Brassica* species, we identified, by BLASTp, tBLASTx, or tBLASTn, Bn-*CLG1A* homologous genes in various plant species (see Supplemental Table 4 online). Homologs were also identified in all other analyzed eukaryote species, including protists, fungi, algae, nematode, insects, fish, mouse, human, and yeast (*Saccharomyces cerevisiae*) (see Supplemental Table 4 online). No homologous sequence was found in prokaryotes (*Bacteria* or *Archaea*).

A phylogenic tree of Bn-*CLG1A* homologous genes in plants was constructed using the protein sequences (Figure 5A). Another phylogenetic tree including all eukaryotes and showing the separation between plant and nonplant Bn-*CLG1A* homologs is presented in Supplemental Figure 6 online (See also Supplemental Data Set 2 online for sequence alignment.). In the rare cases where more than one copy was found in a specific genome, the copy most similar to Bn-*CLG1A* was selected for tree construction.

In the plant kingdom, the eudicot species were clearly distinguished from monocots and from the nonvascular plants, such as *Physcomitrella patens* (Figure 5A). Bn-*CLG1A* homologs in all studied eudicots contain a chloroplast/plastid targeting polypeptide signal in its N terminus (Figure 5B). By contrast, this domain was not found in the six grass monocot species and the nonvascular plant *P. patens* (Figure 5B). The corresponding N-terminal sequences of the Bn-*CLG1A* homologs in the two noneudicot angiosperm species (*Aquilegia coerulea* and *Amborella trichopoda*) and the two nongrass monocots (*Musa acuminata* and *Phoenix dactylifera*) did not align well with the chloroplast/plastid targeting polypeptide signal of eudicots (Figure 5B). Nevertheless, the *A. coerulea* noneudicot species and the *M. acuminata* nongrass monocots species are Ser rich (Figure 5B). Ser enrichment was suggested as an important, but not sufficient, feature of the chloroplast/plastid targeting polypeptide signal (von Heijne et al., 1989; Zybailov et al., 2008).

All identified Bn-*CLG1A* homologs share a similar transmembrane domain topology with 14 predicted transmembrane regions, as well as a highly conserved RINGv domain (C4HC3 type) with characteristic metal ligands (Figure 5C; see Supplemental Figure 7 online).

The Bn-CLG1A protein shares 24.3% amino acid sequence similarity with human TEB4p and 13.1% with yeast DOA10p, and TEB4p and DOA10p have 11.9% amino acid similarity with each other (Swanson et al., 2001; Hampton, 2002; Carvalho et al., 2006; Kreft et al., 2006). Despite their sequence divergence, comparisons show a similar predicted transmembrane topology of Bn-CLG1A, TEB4p, and DOA10p (see Supplemental Figure 8 online). Both the N-terminal RINGv domain and the C terminus are predicted to be localized to the inner side of the membrane (~60% of protein is predicted to face the inner side and ~20% of protein to the outer side of the membrane) as in yeast and human homologs.

The cleistogamy-inducing point mutation P325L in Bn-clg1A-1D occurred in a 23–amino acid domain that is highly conserved in all analyzed plant species as well as in the nonvascular plant *P. patens* subsp *Patens* (Figure 5D). We named this the *cleistogamy* domain (Figure 5D).

Bn-*CLG1A* and Bn-*CLG1C* Homoeologous Genes Are Constitutively and Equally Expressed in *B. napus*

The expression pattern of Bn-*CLG1A* and Bn-*CLG1C*, highly conserved homoeologous gene copies, was first analyzed in *B. napus* by quantitative PCR using gene-specific primers on a wide panel of tissues and organs collected from the oilseed rape cultivar Darmor-*Bzh*. Both Bn-*CLG1A* and Bn-*CLG1C* homoeologs were constitutively and equally expressed in young leaves, petioles, stems, floral buds, and young siliques and immature seeds (Figure 6A). We observed a slight increase in the steady state level of Bn-*CLG1A* and Bn-*CLG1C* transcripts in roots (Figure 6A).

Expression of the Bn-*clg1A-1D* Mutant Allele Is Similar to the Wild-Type Bn-*CLG1A* and the Homoeologous Bn-*CLG1C*

To determine whether the expression of the Bn-*clg1A-1D* allele is affected, we compared the expression of wild-type Bn-*CLG1A* and Bn-*clg1A-1D* as well as the Bn-*CLG1C* homoeolog using quantitative PCR in young 3- or 5-mm flower buds of the two isogenic *B. napus* cultivars Primor and Primor-*Clg*. We found that the Bn-*clg1A-1D* mutant allele is expressed in Primor-*Clg* equally to the Bn-*CLG1A* wild-type allele in Primor (Figure 6B). Expression of the Bn-*CLG1C* homoeolog was also similar to Bn-*clg1A-1D* and Bn-*CLG1A* in both isogenic cultivars (Figure 6B).

We also analyzed Bn-*CLG1A* promoter activity in floral tissues using a reporter gene fusion construct. Thus, the 893-bp promoter fragment of *CLG1A* was inserted into the pBI101-GUS vector to produce the *pBn-CLG1A*:*GUS* construct (see Supplemental Figure 9 online), and β-glucuronidase (GUS) histochemical assays were performed on young and old floral buds of *B. napus* transgenic plants. Plants transformed with *pBn-CLG1A*:*GUS* showed that the transcriptional activity was high in the sepals of young floral buds (Figure 6C). In mature flowers, Bn-*CLG1A* promoter activity was stronger in sepals and anthers than in the style of pistils and the base of petal blades (Figure 6C).

B. napus Plants Transformed with the Bn-*clg1A-1D* Gene Exhibit a Cleistogamy Phenotype

To further validate the cleistogamy candidate gene, next recapitulated the cleistogamy mutant phenotype by transforming wild-type *B. napus* with the *clg1A* mutant genomic fragment. To

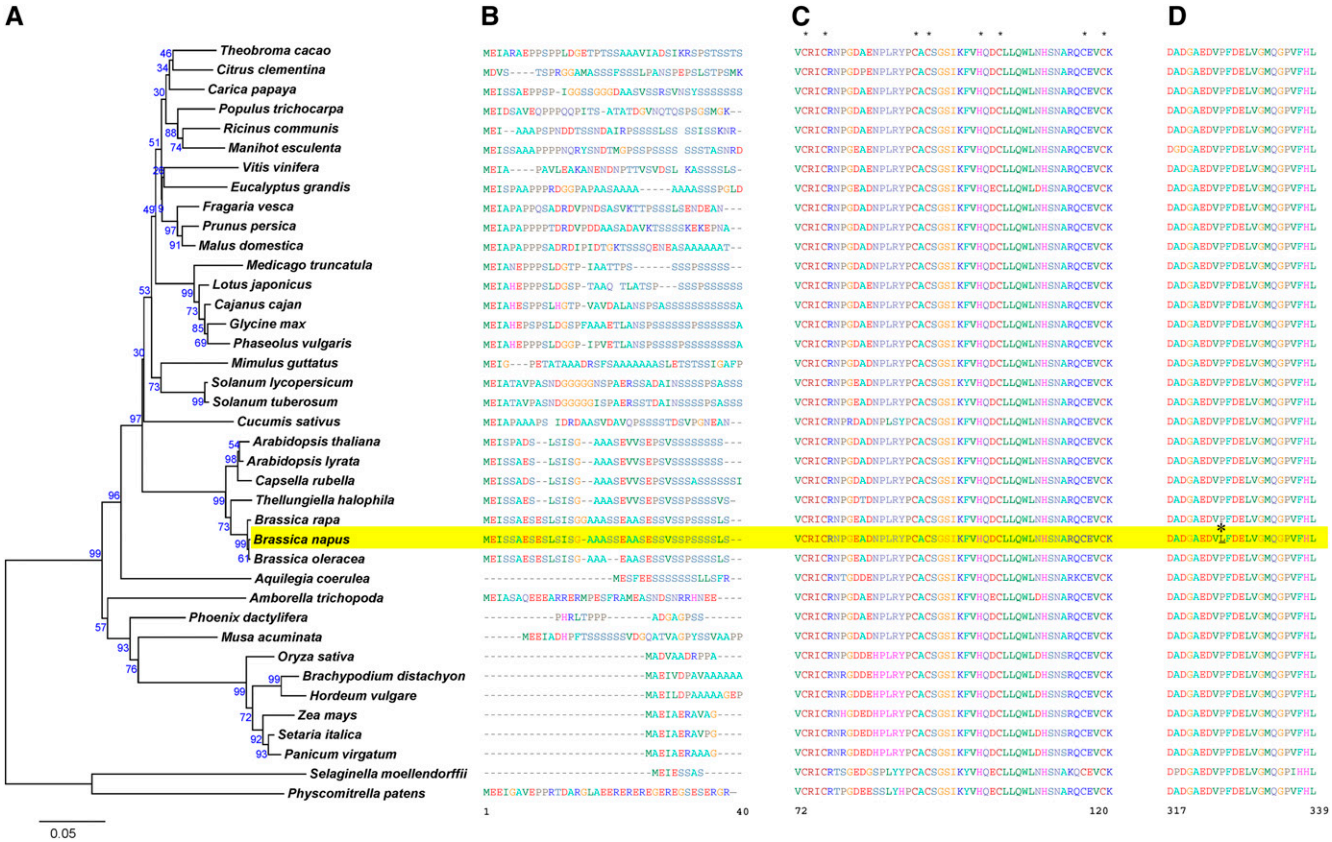


Figure 5. Phylogenetic analysis of plant homologs of Bn-*CLG1A* and amino acid sequence alignment of important domains.

Phylogenetic analysis was done with homologs of Bn-*CLG1A* from 39 plant species. **(A)**. The chloroplastic domain characterized by enrichment of the Ser amino acid **(B)**. The RINGv domain characterized by conserved eight metal ligands marked with an asterisk **(C)**. The highly conserved cleistogamy domain where the cleistogamy mutation P325L (asterisk) occurred in the *B. napus* Bn-*clg1A-1D* allele **(D)**. Amino acid sequences were aligned using ClustalX (Thompson et al., 1997). The phylogeny analysis was performed with MEGA version 4.1 using the neighbor-joining method (Tamura et al., 2007). The percentage of replicates in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. Amino acid numbering is presented according to the Bn-clg1A-1D sequences of *B. napus* (highlighted in yellow). (See Supplemental Data Set 1 online for sequence alignment.)

this end, we conducted *Agrobacterium tumefaciens*–mediated transformations of a noncleistogamous oilseed rape line cv Westar as well as an *Arabidopsis* line (Columbia-0 [Col-0]) with a Bn-*clg1A-1D*–derived construct to express the Bn-*clg1A-1D* allele under the control of its native promoter (see Methods and Supplemental Figure 9 online).

Of 18 kanamycin-resistant T1 transgenic *B. napus* plants, theoretically hemizygous for the transgene and homozygous for the wild-type Bn-*CLG1A* allele, six lines exhibited the wild-type opened flower phenotype and 12 lines showed an intermediate-cleistogamous phenotype. Three representative intermediate-cleistogamous *B. napus* transgenic (T1) plants are shown in Figure 7A. These transgenic plants develop normally and flower at a similar time to the control non-transgenic *B. napus* cv Westar. Compared with the control, all the intermediate-cleistogamous lines analyzed exhibited the same floral phenotype with apical floral buds completely closed, while opened or partly opened flowers were observed at the middle and the base of the stems, confirming that the cleistogamy phenotype was caused by the induced P325L mutation (Figure 7A). Notably, few flowers were also observed and the cleistogamous phenotype observed for the *B. napus* Bn-*clg1A-1D* transgenic lines was less pronounced than the phenotype of the original Bn-*clg1A-1D* mutant line Primor-Clg or to the 'B001-*Clg*' line and more resembled the intermediate-cleistogamous phenotype described above (cf. Figures 1 and 7A).

Arabidopsis Plants Transformed with the Mutated Bn-*clg1A-1D* Gene Exhibit a Pronounced Cleistogamy Phenotype and Reduced Fertility

To get more insights into the function of the Bn-*clg1A-1D* allele, we next recapitulated the phenotype in *Arabidopsis* by expressing the same Bn-*clg1A* construct that worked in *B. napus*. We carefully analyzed the phenotype of three *Arabidopsis* independent T1 lines transformed with the Bn-*clg1A-1D* gene–derived construct as was used in *B. napus* (under the control of its native promoter). At the flowering stage, flowers of Bn-*clg1A-1D*

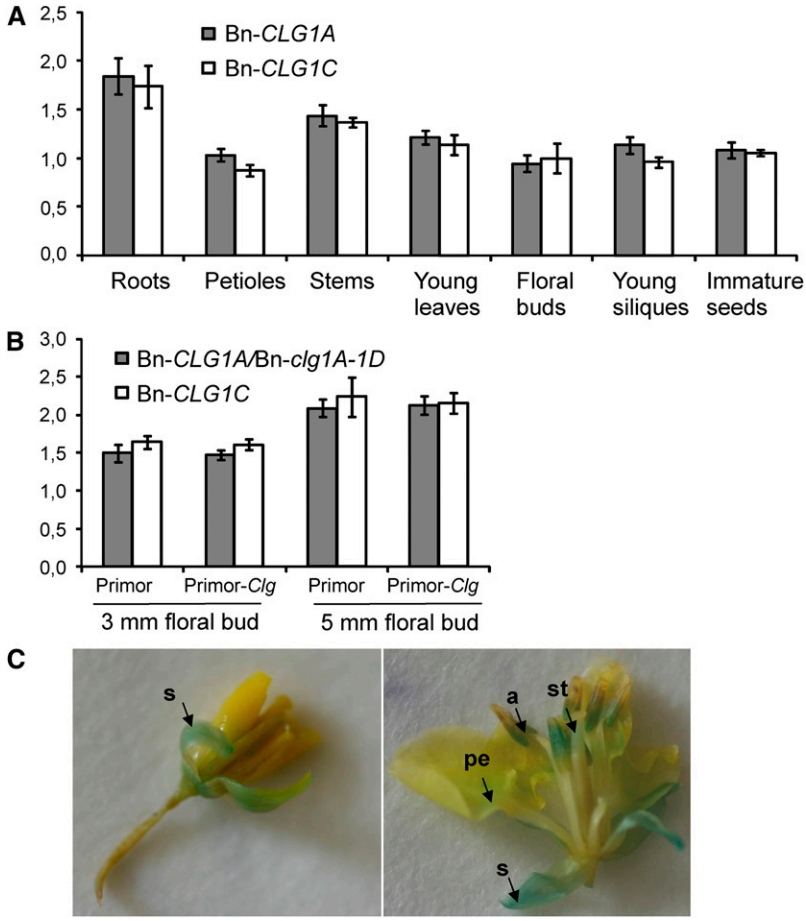


Figure 6. Bn-*clg1A-1D*, Bn-*CLG1A*, and Bn-*CLG1C* Homoeologs Are Constitutively and Similarly Expressed in *B. napus*.

(A) Bn-*CLG1A* and Bn-*CLG1C* expression levels in different tissues of *B. napus* cv Darmor. Transcript levels were determined by quantitative RT-PCR and normalized to both *EF-1* and *UBQ2* genes. Bars indicate SD ($n = 3$).

(B) Bn-*clg1A-1D* and Bn-*CLG1C* expression levels in *B. napus* cv Primor (the wild type) and Primor-*Clg* (Bn-*clg1A-1D* mutant) 3- and 5-mm floral buds. Transcript levels were determined by quantitative RT-PCR and normalized to both *EF-1* and *UBQ2*. Bars indicate SD ($n = 3$). The experiments have been repeated twice with similar results.

(C) GUS assay on young (left) and old (right) flowers of transgenic *B. napus* cv Westar transformed with the *p*Bn-*CLG1A*:*GUS* construct. Arrows indicate the accumulation of GUS product (blue staining) in sepal (s), the base of petal (pe), anther (a), and style (st) of a developing flower.

[See online article for color version of this figure.]

transgenic *Arabidopsis* were strongly cleistogamous with unopened floral buds (Figure 7B). Indeed, compared with the cleistogamous *B. napus* 'B001-*Clg*' line, the Bn-*clg1A-1D* transgenic *Arabidopsis* lines exhibited a more pronounced cleistogamous phenotype with more strongly closed flowers (Figures 7A and 7B). Terminal inflorescences of Bn-*clg1A-1D* transgenic *Arabidopsis* produced more floral buds and were more compact compared with the wild type (Figure 7B; see Supplemental Figures 10A and 10B online). In addition, apical floral buds appear glued between them.

In the Bn-*clg1A-1D* transgenic *Arabidopsis*, most of the siliques did not develop and the ones that developed were semi-sterile (see Supplemental Figure 10B online). Each line produced only one to five fully developed siliques that were curved, shorter, and wider compared with wild-type siliques (see Supplemental Figures 10D to 10F online). Notably, petals of Bn-*clg1A-1D* transgenic *Arabidopsis* remain attached to siliques, whereas those of wild-type plants abscised normally (see Supplemental Figures 10D to 10F online).

Bn-*clg1A-1D* Transgenic *Arabidopsis* Plants Are Affected in Petal Epidermal Cell Elongation and Cuticular Nanoridge Deposition

To further dissect the cleistogamous phenotype of T1 *Arabidopsis* plants transformed with Bn-*clg1A-1D*, floral buds and flowers were analyzed with a microscope or using a scanning electron microscope. Unopened floral buds of Bn-*clg1A-1D* transgenic *Arabidopsis* displayed organ fusions between sepals and between sepals and petals (Figures 8A and 8B; see Supplemental Figure 11 online). Compared with the wild type, petals of Bn-*clg1A-1D* *Arabidopsis* were folded, thicker, and

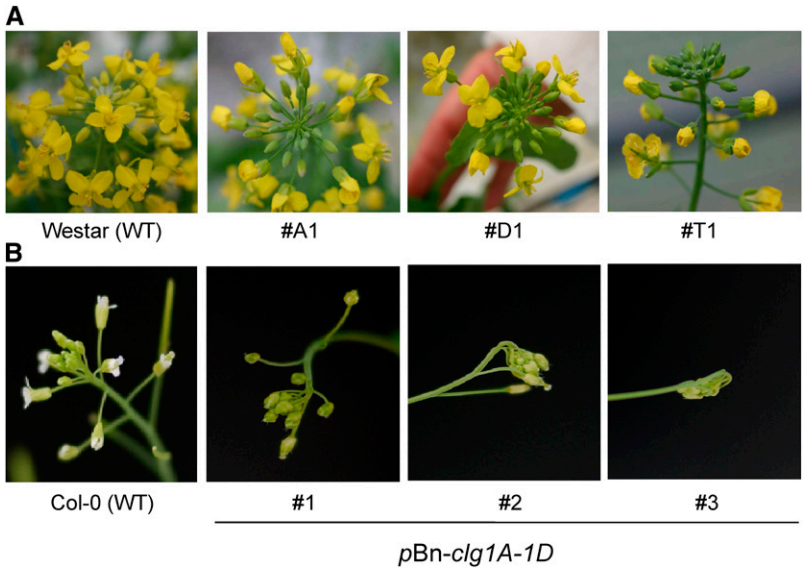


Figure 7. Cleistogamous Phenotypes of *B. napus* and *Arabidopsis* Plants Transformed with Bn-*clg1A-1D*.

(A) Inflorescence of *B. napus* cv Westar and three representative T0 transgenic lines (#T1, #D1, and #A1) of same cultivar transformed with the *p*Bn-*clg1A-1D* construct. WT, the wild type.

(B) Inflorescence of *Arabidopsis* ecotype Col-0 (the wild type) and three independent *p*Bn-*clg1A-1D* transgenic T0 lines (#1, #2, and #3) in the same ecotype (see Supplemental Figure 9 online for details of construct *p*Bn-*clg1A-1D*).

cannot develop inside the closed flower (Figure 8; see Supplemental Figures 11A to 11D online). Elongation of carpel was restricted and they were curved or twisted (see Supplemental Figures 11C and 11D online).

In Bn-*clg1A-1D* transgenic *Arabidopsis*, epidermal cells at the base of petals were much less elongated than in wild-type plants, whereas the size of epidermal cells of the upper petal blade did not change compared with the wild type (see Supplemental Figures 11E and 11F online). As previously observed (Pyke and Page, 1998), these two types of epidermal cells each occupied about half of the petal length in the wild-type plant. Interestingly, in Bn-*clg1A-1D* transgenic *Arabidopsis*, the proportion of conical epidermal cells in the upper part of petals, especially the ones at the adaxial surface, was higher than the proportion of the thin and elongated epidermal cells at the base of petals (see Supplemental Figures 11E and 11F online). This result suggests that in Bn-*clg1A-1D* transgenic *Arabidopsis*, epidermal cells at the base of petals are defective in cell elongation or delayed in their differentiation. On the opposite, petal mesophyll cells were more elongated in Bn-*clg1A-1D* transgenic *Arabidopsis* than in wild-type plants (see Supplemental Figures 11E and 11F online). Similarly, epidermal cells on the abaxial surface of the sepal were in general slightly more elongated and wider than their wild-type counterpart (Figures 8C and 8D; see Supplemental Figures 12G and 12H online).

No significant differences were observed in the number of epidermal cells at the abaxial surface of sepal and both abaxial and adaxial surfaces of petals. In square areas of 10,000 μm^2, wild-type plants of *Arabidopsis* show 7.33 $\pm$ 1.17 and 9.23 $\pm$ 0.76 cells, whereas Bn-*clg1A-1D* transgenic *Arabidopsis* showed 9.07 $\pm$ 1.17 and 10.67 $\pm$ 1.55 cells for the abaxial and adaxial of surfaces, respectively, of petal blades. Nevertheless, we observed that cells of the Bn-*clg1A-1D* transgenic plants were strongly misshapen (Figures 8E to 8H;see Supplemental Figure 12 online). Abaxial epidermal cells of the petal blade were more prominent and less inflated than wild-type cells (Figures 8E and 8F; see Supplemental Figures 12A and 12B online).

Interestingly, cuticular nanoridges, typically displayed on wild-type petal and sepal epidermis (Li-Beisson et al., 2009; Shi et al., 2011), were absent (petals) or severely reduced (sepals) at the abaxial surface of epidermal cells in the Bn-*clg1A-1D* transgenic *Arabidopsis* (Figures 8E to 8H; see Supplemental Figures 12G and 12H online). On the adaxial surface of petals, epidermal cells of Bn-*clg1A-1D* transgenic *Arabidopsis* were less conical and their flattened tips were deprived of nanoridges (see Supplemental Figures 12E and 12F online). Remarkably, abaxial surface of sepals appeared cracked or wrinkled because most of epidermal cells were hardly distinguishable and completely lost their tubular shapes (Figures 8G and 8H). Some surface areas exhibit wild-type cells with regular shapes, notably at the base of sepals.

DISCUSSION

In this study, we showed that a point mutation in a RINGv E3 ubiquitin ligase encoding gene (Bn-*CLG1A*) leads to cleistogamy phenotype in *B. napus*. The existence of Bn-*CLG1A* homologs in all eukaryotes suggests that their function is evolutionarily conserved. Biochemical and cellular characterization have shown that human (TEB4p) and yeast (DOA10p) homologs play a role in endoplasmic reticulum–associated degradation where ubiquitinylation is essential for both their retrotranslocation into the

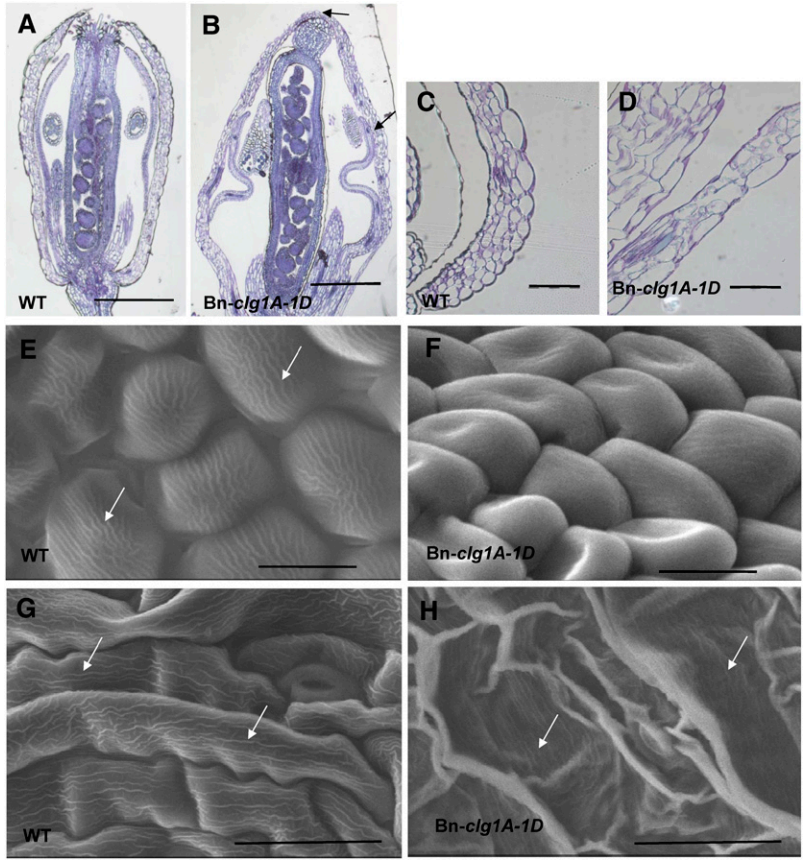


Figure 8. Changes in Flower Morphology and Surface Characteristics of *Arabidopsis* Plants Transformed with Bn-*clg1A-1D* as Observed by Light and Electron Microscopy.

(A) to **(D)** Light microscopy images of Toluidine Blue–stained longitudinal sections of flowers at developmental stage 12 (**[A]** and **[B]**) and sepals at developmental stage 13 (**[C]** and **[D]**). Col-0 wild-type (WT) (**[A]** and **[C]**) and transgenic Bn-*clg1A-1D* plants (**[B]** and **[D]**) are represented. Arrows indicate sepal-to-sepal and sepal-to-petal fusions.

(E) to **(H)** Scanning electron microscopy images of abaxial surfaces of petal (**[E]** and **[F]**) and sepal (**[G]** and **[H]**) epidermis in Col-0 wild-type (**[E]** and **[G]**) and transgenic Bn-*clg1A-1D* plants (**[F]** and **[H]**). Arrows indicate nanoridges.

Bars = 500 μm in **(A)** and **(B)**, 100 μm in **(C)** and **(D)**, 10 μm in **(E)** and **(F)**, and 30 μm in **(G)** and **(H)**.

[See online article for color version of this figure.]

cytosol and degradation (Swanson et al., 2001; Hwang et al., 2010). The Cri-du-chat syndrome, a neurodegenerative disorder in humans, was found to be associated with the chromosome region containing the *TEB4* gene (Hassink et al., 2005). The high protein structure conservation suggests that Bn-CLG1A should maintain a similar biochemical and cellular function in endoplasmic reticulum–associated degradation and ubiquitin-mediated protein degradation.

Bn-*CLG1A* and its homologs in plants exhibit specific evolutionary features: (1) The putative chloroplast/plastid targeting polypeptide signal in the N terminus of Bn-*CLG1A* homologs is highly conserved in eudicot plant species and was not found in fungi and animals and more interestingly not evidenced in the monocot grass plant species. (2) Moreover, the occurrence of a 23–amino acid domain is highly conserved among all the analyzed vascular and nonvascular plant species and that we called the cleistogamy domain. The P325L cleistogamy-inducing point mutation in *B. napus* Bn-*clg1A-1D* occurred in this domain.

Interestingly, *B. napus* and *Arabidopsis* plants transformed with the Bn-*clg1A* gene showed cleistogamous flowers. The cleistogamy phenotype in these transgenic plants is most likely caused by the P325L mutation and not by other possible effects that would be independent from the mutation, such as overexpression of the Bn-*clg1A-1D* gene: (1) Transformation experiments were done in this study using a construct where the Bn-*clg1A-1D* gene is under the control of its endogenous promoter and not, for example, under the control of the 35S promoter, which favors overexpression. However, it could not be completely excluded that more or overexpression of the Bn-*clg1A-1D* transgene occurred under its endogenous promoter, as its expression was not analyzed in this study. (2) Nevertheless, overexpression is unlikely to cause the cleistogamy phenotype. The transformation and overexpression in *Arabidopsis*

of the *Arabidopsis* orthologous wild-type allele (*cer9*), which does not contain the P325L mutation, was done in a separate study (Lü et al., 2012), and no effect on flower morphology or cleistogamy phenotype was noticed. (3) Finally, *B. napus* noncleistogamous cultivars, having thus the two wild-type homoeologs, Bn-*CLG1A* and Bn-*CLG1C*, show normal flower phenotype.

During flower bud opening, various events take place, representing all aspects of plant development, such as cell division, cellular differentiation, and cell elongation or expansion. Flower opening depends on many factors, including internal ones, such as plant growth regulators (phytohormones), water supply, carbohydrate metabolism, cell wall/membrane metabolism, as well as external factors, such as light, temperature, and humidity (Kumar et al., 2008). The mechanism of flower opening varies considerably in the angiosperms (reviewed in van Doorn and van Meeteren, 2003). For some plant species, flower opening could be related to the development of adjacent tissues. For example, it may require growth of the pedicel or forced separation or abscission of covering parts. Flower opening can also depend on petal movements, which might be due to (1) reversible ion accumulation (independent from elongation growth), (2) cellular death in a specific area of the petal, (3) loss of water during the day and refilling during the night, and (4) differential growth rate of the two sides of the petal (as reviewed by van Doorn and van Meeteren, 2003). In most species, petal movements are due to a difference in growth rate of the two sides. In *Arabidopsis*, the opening of flower buds appears to be the result of rapid cell expansion (instead of cell division) in petals (primarily of the basal epidermal and mesophyll cells) (Pyke and Page, 1998). In grasses, like rice and barley, flower opening is achieved by a mechanism of lodicule (equivalent to petals in dicotyledons) swelling (Yoshida et al., 2007; Nair et al., 2010).

In this study, we have characterized several features of *Arabidopsis* Bn-*clg1A-1D* transgenic cleistogamous flowers that highlight important insights about the mechanism by which the P325L mutation leads to the cleistogamy phenotype. (1) We showed that expression of Bn-*clg1A-1D* causes petal-to-sepal and sepal-to-sepal fusions. Interestingly, *Arabidopsis* mutant lines, reported in other studies as affected in the flower phenotype, show also similar flower organ fusions (Li-Beisson et al., 2009; Panikashvili et al., 2009, 2011; Shi et al., 2011). Also, flower organ fusions, consisting of the formation of lodicule-glume mosaic organs, have been observed in cleistogamous flowers of rice (Yoshida et al., 2007). (2) We observed that Bn-*clg1A-1D* expression perturbed petal and sepal cell elongation, cellular expansion, and/or differentiation. Compared with wild-type plants, sepals epidermal and petal mesophyll cells of *Arabidopsis* Bn-*clg1A-1D* transgenic cleistogamous flowers were more elongated while epidermal cells at the base of petal were less elongated. (3) More importantly, we observed reduction of cuticular nanoridges on sepal and petal epidermal cells in cleistogamous *Arabidopsis* flowers, indicating defects in cutin loading or synthesis in flowers. Strikingly, similar defects in the synthesis or loading of cutin monomers was previously observed for other *Arabidopsis* lines that were independently reported as being affected in flower morphology (Li-Beisson et al., 2009; Panikashvili et al., 2009; Panikashvili et al., 2011; Shi et al., 2011).

Interestingly, a recent study has suggested that the orthologous *At4g34100* (*CER9*) wild-type gene of *Arabidopsis* acts as a negative regulator of cutin biosynthesis or loading (Lü et al., 2012). This was based on the fact that *cer9* recessive and loss-of-function mutants, corresponding to wax mutant *eceriferum9* (Koornneef et al., 1989), showed an elevated amount of cutin monomer (Lü et al., 2012). On the opposite side, the P325L cleistogamy mutation in *B. napus* Bn-*clg1A-1D* might cause a defect or reduction in cutin loading or synthesis in flowers as suggested by the observed reduction in petal and sepal cuticular nanoridges. These data strongly suggest that the mutant gene acts as a pronounced negative regulator of cutin biosynthesis or loading. Notably, no defect in flower morphology was reported for the *cer9* recessive and loss-of-function mutants (Koornneef et al., 1989; Lü et al., 2012). This was also the case for one *Arabidopsis* knockout line, mutated by T-DNA in the *At4g34100* gene, and two other lines transformed with RNA interference constructs that we have phenotyped (data not shown). Measurement of cuticular wax composition in Bn-*clg1A-1D* flowers will better shed light on how the P325L mutation affects the Bn-*CLG1A* function.

Together, these observations and comparisons with other studies strongly suggest that the P325L cleistogamy mutation Bn-*clg1A-1D* causes a pronounced negative regulation of cutin biosynthesis or loading and affects elongation or differentiation of petal and sepal cells. This probably results in inhibition or delay of the petal development as well as petal-to-sepal and sepal-to-sepal fusions, leading to folded petals and the cleistogamy phenotype. Accordingly, results from promoter-GUS analysis indicate that the Bn-*CLG1A* gene might function preferentially in petal, sepal, and stamen during flower development. No defect in leaf morphology or plant development was observed (data not shown). Intriguingly, quantitative RT-PCR analyses showed that Bn-*CLG1A* is constitutively expressed in all plant organs. Differences in protein–protein interactions depending on tissue-specific expression of the Bn-clg1A protein may explain this discrepancy. Alternatively, the P325L mutation at the cleistogamy domain may specifically affect flower development.

The P325L cleistogamy mutation in Bn-*clg1A-1D* did not affect its expression or that of its homoeologous Bn-*CLG1C* gene copy, suggesting that it acts at the protein interaction level. Moreover, heterozygous oilseed rape lines as well as transgenic *Arabidopsis* and oilseed rape plants expressing Bn-*clg1A-1D* exhibited a cleistogamous phenotype (Figures 1 and 3; Renard and Tanguy, 1997). These observations suggest that the P325L mutation is semidominant and not a loss-of-function mutation.

It is unlikely that the P325L mutation would correspond to a dominant-negative form. As discussed above, the *Arabidopsis* orthologous *CER9* gene was shown to act as a negative regulator of cutin loading or biosynthesis, (Lü et al., 2012). Therefore, if the P325L mutation is a dominant-negative form and alters this function (of negative regulation of cutin loading), we would expect to have more cutin monomer (and cuticular wax) in *Arabidopsis* transgenic plants. In contrast with this, we observed a marked and a more pronounced reduction of cuticular nanoridges in petals and sepals. Moreover, we didn't observe brighter leaves, stem, or silique phenotypes characteristic of *cer9* knockout

mutants (Koornneef et al., 1989). Therefore, the P325L mutation is functional and would rather correspond to a gain-of-function and semidominant positive form.

Various interpretations could be advanced to explain functionality of the gain of function and semidominant mutation of Bn-*clg1A-1D* in relation to the homoeologous Bn-*CLG1C*. A primary interpretation is that Bn-*CLG1A* and Bn-*CLG1C* wild-type homoeologous genes had established different and independent functions in *B. napus*, prior to the cleistogamy mutation. In this case, the semidominant P325L EMS mutation in the Bn-*CLG1A* homoeologous gene copy has led to the gain of function of the Bn-*clg1A-1D* cleistogamy allele, without any interaction with the Bn-*CLG1C* homoeologous gene. Sequence comparisons showed that Bn-*CLG1A* or Bn-*CLG1C* homoeologous genes in *B. napus* are 100% identical at the amino acid level to their corresponding orthologs in the *B. rapa* (Br-*CLG1A*) and *B. oleracea* (Bo-*CLG1C*) parental species. This suggests that no specific mutations or changes have occurred in *B. napus* to diverge the function of these homoeologous genes. This leaves a possibility that different functions preexisted and would have evolved in the Br-*CLG1A* or Bo-*CLG1C* orthologous genes of the parental species, before the recent allopolyploidy reunited them in the *B. napus* genome. This is also unlikely because the Br-*CLG1A* gene of *B. rapa* and its orthologous Bo-*CLG1C* gene of *B. oleracea*, which are highly conserved (99.5% identical at the amino acid sequence level), represent the unique genes of these parental species whose structure is highly conserved with Bn-*CLG1A* and Bn-*CLG1C* homoeologous genes of *B. napus*.

Therefore, the most plausible interpretation is that Bn-*CLG1A* and Bn-*CLG1C* of *B. napus* are functionally redundant. Functionally redundant duplicated genes can follow one of many possible evolutionary fates, including nonfunctionalization (deletion or pseudogenization of one duplicated copy), neofunctionalization (evolution of novel functions among alleles or homoeoalleles), or subfunctionalization (evolution of partitioned ancestral functions among alleles or homoeoalleles) (Lynch and Force, 2000; Prince and Pickett, 2002). Subfunctionalization and neofunctionalization were suggested as mechanisms leading to the preservation of duplicated genes after polyploidization (Lynch and Force, 2000; Prince and Pickett, 2002), and several interesting examples have been recently reported (Hovav et al., 2008; Chaudhary et al., 2009; Zhang et al., 2011). Although the cleistogamy mutation studied here is EMS induced and didn't occur during natural evolution as originally defined for sub- and neofunctionalization (Lynch and Force, 2000), our study provides an interesting example of functional and concerted interaction between two homoeologous genes in the relatively recent allopolyploid *B. napus*. The possible interaction fates are as follows: (1) The fact that the P325L cleistogamy mutation is not a loss-of-function or a recessive but rather a gain-of-function and semidominant positive mutation (see above) excludes the possibility that it led to loss of function of Bn-*CLG1A-1D*. (2) It is possible that this mutation may have caused an increased contribution of the Bn-*clg1A-1D* gene to the original function. We previously suggested this type of subfunctionalization, where a homoeologous gene contributes more to the function than the other ones, and defined it as hyperfunctionalization in the case of the domestication *Q* gene homoeologs in polyploid wheat (Zhang et al., 2011). The hyperfunctionalization suggested here would act at the protein level and is supported by the finding that compared with the *At4g34100* wild-type (*CER9*) orthologous gene of *Arabidopsis*, which acts as a negative regulator of cutin biosynthesis or loading (Lü et al., 2012), the Bn-*clg1A-1D* cleistogamy mutant acts as a stronger negative regulator of cutin biosynthesis or loading. (3) On the other hand, neofunctionalization could also represent an alternate model for the interaction between the Bn-*clg1A-1D* gain-of-function semidominant mutation and its homoeologous Bn-*CLG1C* gene. The mutation didn't affect expression of the *clg1A* gene or of Bn-*CLG1C*. Moreover, the cleistogamy phenotype was observed independently from the copy number of homoeologs or homologs of the *clg1A* gene, present in the mutated *B. napus* as well as the transgenic *B. napus* and *Arabidopsis* in which Bn-*clg1A-1D* was introduced. These features characterize neofunctionalization (Lynch and Force, 2000).

Notably, whether Bn-*clg1A-1D* has been hyper- or neofunctionalized or has been subjected to a combination of both phenomena, it is not demonstrated in this study that the P325L mutation has impacted preservation of duplicated gene copies (Lynch and Force, 2000), although reproductive fitness and fertility of cleistogamous *B. napus* were not affected (Renard and Tanguy, 1997; Fargue et al., 2006).

In conclusion, our study reports the cloning of a cleistogamy gene in dicots that may cause pronounced negative regulation of cutin biosynthesis or loading in flower organs and corresponds to a gain of function in a homoeologous gene. Further detailed analyses are needed to determine the cutin monomer composition in cleistogamous flowers as well as to identify Bn-CLG1A and Bn-clg1A interacting proteins. Analysis of the Bn-CLG1A and Bn-clg1A E3 ligase activity also would be interesting. The study opens new perspectives for future characterizations of the biological functions of Bn-*CLG1A* as well as the fate of duplicated genes and the evolution of new functions of duplicated gene copies in allopolyploid species.

METHODS

Plant Materials

The *Brassica napus* cleistogamous mutant was originally obtained in the seventies at Institut National de la Recherche Agronomique Rennes by EMS mutagenesis on seeds of the cv Primor (Renard and Tanguy, 1997). The cleistogamous trait was then introduced into the cultivar 'Samourai' by successive backcrossing, and this gave rise to the cleistogamous line 'B001-*Clg*' used as the cleistogamous parent in this study.

A population of 255 DH lines was produced by microspore culture from the F1 plants of a cross between 'Yudal' (open flower type) and 'B001-*Clg*' (cleistogamous type) as described by Polsoni et al. (1988) and used for initial mapping of Bn-*clg1A-1D*. All the DH lines were first phenotyped twice in greenhouse for cleistogamy. These were also evaluated in the experimental field conditions of Institut National de la Recherche Agronomique Le Rheu (Rennes, France) and used for the initial AFLP mapping. For fine mapping, a total of 2158 DH lines were genotyped with linked molecular markers, and 31 recombinant lines were then evaluated for the cleistogamy phenotype in the field and greenhouse conditions.

Arabidopsis thaliana plants were grown on soil in a greenhouse at 22°C/16-h day and 18°C/8-h night cycle. *B. napus* plants were grown either in

the field or in a greenhouse with a thermo/photoperiod of 25°C/16-h days and 18°C/8-h nights.

For expression analysis, plants of *B. napus* cv Darmor were individually planted into 2-liter pots filled with perlite. Plants were regularly watered with Hoagland nutritive solution. Root samples were collected from 6-week-old plants. Leaf blades and petioles separated from the three younger leaves and stem sections at the base of plants were sampled on 2-month-old plants. Young floral buds as well as young siliques and immature seeds at 11 and 19 d after manual pollination, respectively, were collected on 3.5-month-old plants. For *B. napus* cv Primor and Primor-*Clg* winter cultivars, 3 weeks after planting in soil, plants were vernalized for 2 months at 8°C with a 16-h/8-h day/night conditions. After 1 month of growth in the greenhouse, 3- to 5-mm floral buds were collected. All samples were immediately frozen in liquid nitrogen upon collection. For each sample, at least three individual plants were pooled together.

Methods of phenotyping, molecular marker development, DNA extraction, and fine genetic mapping of Bn-*clg1-1D* are detailed in Supplemental Methods 1 online.

Construction and Screening of Ordered and Pooled BAC Libraries

A BAC genomic DNA library was previously constructed from cv Darmor-*Bzh* (Chalhoub et al., 2004). This was used for physical mapping of Bn-*CLG1A* where the BAC clone CZ7N2 containing Bn-*CLG1A* was identified.

To isolate the mutant allele Bn-*clg1A-1D*, we constructed a pooled large insert library for Primor-*Clg* cultivar following the procedures described by Isidore et al. (2005). Specific PCR markers tightly linked to Bn-*clg1A-1D* were then used to screen the pools of BAC clones of the library to identify the BAC clone "Clg_H1B_P13a_3_J1" containing the mutant allele.

Sequence Alignment and Phylogenetic Analysis of Bn-*CLG1A* Homologs

BAC clone sequencing and annotation were done as essentially described by Chantret et al. (2005). Multiple sequence alignments were performed with ClustalX (Thompson et al., 1997). The protein sequences of Bn-*CLG1A* homologs in other species were retrieved by TBLASTp search from the "nonredundant protein sequences" databases at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/BLAST/). The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling, 1965). Phylogenetic analyses were conducted with MEGA4 software (Tamura et al., 2007), where the software option relying on the neighbor-joining method (Saitou and Nei, 1987) was applied. A bootstrap test inferred from 1000 replicates was applied (Felsenstein, 1985).

Analysis and Comparison of Protein Sequence Domains

Subcellular localization of the different protein sequences were predicted by the online software WoLFPSORT (http://wolfpsort.org/), MultiLoc (http://www-bs.informatik.uni-tuebingen.de/Services/MultiLoc/), TargetP1.1 (http://www.cbs.dtu.dk/services/TargetP/), ChloroP 1.1, (http://www.cbs.dtu.dk/services/ChloroP/), TargetLoc (http://www-bs.informatik.uni-tuebingen.de/Services/MultiLoc/), PredictNL (http://cubic.bioc.columbia.edu/cgi/var/nair/resonline.pl), and LOCtree (http://cubic.bioc.columbia.edu/cgi-bin/var/nair/loc). Membrane topology of Bn-CLG1A, TEB4p, and DAO10p was predicted by TMHMM2.0 (http://www.cbs.dtu.dk/services/TMH/) (Krogh et al., 2001).

Gene Expression Analysis

Total RNA was extracted using Trizol reagent (Invitrogen) for Darmor samples or using the SV Total RNA isolation kit (Promega) for Primor and Primor-Clg samples according to the manufacturers' instructions. Genomic DNA was removed by digestion with DNase I (RNase-free; Ambion) according to the manufacturer's protocol except that DNase I was heat inactivated by 5-min incubation at 75°C. RNA concentration and integrity were determined by OD measurement at 260 and 280 nm and with an Agilent 2100 Bioanalyzer (Agilent Technologies). Two micrograms of total RNA were reverse transcribed using 500 ng of oligo(dT)12-18 and 200 units of SuperScriptII RT in a total volume of 20 μL following the manufacturer's instructions (Invitrogen, Life Technologies). After first-strand cDNA synthesis, RNA complementary to the cDNA was removed by incubating the reaction with 2 units of *Escherichia coli* RNase H at 37°C for 20 min. cDNA was diluted fivefold before real-time PCR. Real-time PCR reactions were performed on a CFX384 Touch Real-Time PCR detection system (Bio-Rad) using 9 μL of SsoAdvanced SYBR Green Supermix (Bio-Rad), 2 μL of cDNA, and 0.56 μM primers in a total volume of 18 μL per reaction. The cycling conditions were composed of an initial 30-s denaturation step at 95°C, followed by 40 cycles of 95°C for 5 s and 60°C for 30 s. Melting curve was run from 65 to 95°C to ensure the specificity of product. Data were analyzed using CFX Manager Software version 2.1 (Bio-Rad). *B. napus* elongation factor 1 (*EF-1*) and ubiquitin 2 (*UBQ2*) were used as reference genes for normalization of gene expression levels in all samples. Quantitative RT-PCR forward and reverse primer sequences for each gene are listed in Supplemental Table 1 online. Primer efficiencies for each couple of primers were 95.57% for Bn-*CLG1A*, 101.52% for Bn-*CLG1C*, 91.79% for *EF-1*, and 91.07% for *UBQ2*.

Gene Constructs, Transformations, and Complementation Analysis

To validate function of the cleistogamy candidate gene, two constructs were used. The pBn-clg1A:Bn-clg1A construct (see Supplemental Figure 9 online) contains an *Xba*I-generated 5999-bp genomic fragment of the mutant allele cloned into the binary vector pBINPLUS (van Engelen et al., 1995). The 5999-bp genomic fragment is composed of a 895-bp native promoter region, the 4144-bp predicted coding Bn-*clg1A-1D* allele, and the 960-bp downstream 3′-untranslated region (see Supplemental Figure 9 online). The pBn-clg1A:GUS construct was made using the 893-pb fragment of the promoter region of the candidate Bn-*CLG1A* gene (see Supplemental Figure 9 online) inserted at the beginning of GUS gene, into a binary vector pBI101-GUS with Gateway cloning technology (Invitrogen). We conducted *Agrobacterium tumefaciens*–mediated transformations in oilseed rape using methods developed by Moloney et al. (1989) and Sparrow et al. (2004) and in *Arabidopsis* by floral dip according to Clough and Bent (1998). Transgenic T1 plants were selected on half-strength Murashige and Skoog media containing 50 μg/mL kanamycin. pBn-clg1A:GUS construct was used for GUS coloration assays (Jefferson et al., 1987) to check specificity of expression in the different tissues of transformed *B. napus*.

Microscopy

For Toluidine Blue-O staining, the inflorescences of 7-week-old plants were fixed in a 1% glutaraldehyde, 2% paraformaldehyde, and 1% caffeine solution and embedded in Kulser resin. Sections were cut to a thickness of 5 μm on a microtome before Toluidine Blue-O staining. For scanning electron microscopy analysis, flowers from 7-week-old plants were collected, and samples were slowly frozen at –30°C under a partial vacuum on the Peltier stage before observation using a Hirox SH-1500 microscope at 15 kV.

Accession Numbers

Sequence data from this article can be found in the Arabidopsis Genome Initiative or GenBank/EMBL databases under the following accession

numbers: BAC clone CZ7N2, KC206509; BAC clone Clg_H1B_P13a_3_J1, KC206510; Bn-*CLG1A*, KC206501; Bn-*clg1A-1D*, KC206502; Bn-*CLG1C*, KC206503; Bn-*Rzfp2A*, KC206504; Bn-*Rzfp2C*, KC206505; Bn-*Rzfp3A*, KC206506; Bn-*Rzfp3C*, KC206507; Bn-*Rzfp2Cb*, KC206508; *EF1*, *FJ529180*; *UBQ2*, *EE426322*; *CER9*, *At4G34100*. The other *Brassica* sequence data are from the BRAD database (http://brassicadb.org/brad/) and have the following accession numbers: Bo-*CLG1C*, BGIContig_196373_816_7400_rv; Bo-*Rzfp2C*, BGIscf7180014794361_6919537_6925790_rv; Bo-*Rzfp2Cb*, BGIScf7180014794480_102891_108258_rv; Bo-*Rzfp3C*, BGIScf7180014794480_67182_72615_rv; Br-*CLG1A*, Bra011487 or BGIScaffold000011_2395792_2402090; Br-*Rzfp2A*, BGIScaffold000023_175528_181823; Br-*Rzfp3A*, BGIScaffold000097_438462_443628. From GenBank (http://www.ncbi.nlm.nih.gov/), accession numbers of sequence data for Bn-*CLG1A* homologs of other species used in this study are supplied in Supplemental Table 4 online.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure 1. Scoring and Segregation of the Cleistogamy Phenotype in a Population of 255 Doubled Haploid Lines of *B. napus*.

Supplemental Figure 2. Dot Plot Comparison between the *Brassica napus* Region Spanning the Bn-*CLG1A* Wild-Type Gene and the Orthologous Genomic Sequence from *Arabidopsis thaliana*.

Supplemental Figure 3. Physical Positions of the Genetic Markers Closely Linked to Bn-*CLG1A* and Sequence Comparison between Bn-*CLG1A* and Bn-*clg1A-1D*.

Supplemental Figure 4. Design of a dCAPS Marker for the Detection of the T/C SNP Differentiating Bn-*clg1A-1D* and Bn-*CLG1A*.

Supplemental Figure 5. Chromosomal Colinearity and Phylogenetic Relationships between Genes from the *Arabidopsis* Chromosome 4 Genomic Region Spanning the *At4g34100* Gene and Their Triplicated Orthologs Found in *Brassica rapa*.

Supplemental Figure 6. Phylogenetic Analysis of Homologs of Bn-*CLG1A*.

Supplemental Figure 7. Alignment of RINGv Domains of Bn-CLG1A Homologous Protein Sequences Identified in All Eukaryote Species.

Supplemental Figure 8. Comparison of Membrane Topology between Bn-CLG1A TEB4p and DOA10p Proteins.

Supplemental Figure 9. Schematic Presentation of Genomic DNA Promoter Fragments Used for the *p*Bn-*CLG1A*:*GUS* Construct, as Well as the Entire Gene Sequence Fragment Used for Bn-*clg1A-1D* Gene-Derived Construct (*p*Bn-*clg1-1D*).

Supplemental Figure 10. Reproductive Organ Morphology of *Arabidopsis* Plants Transformed with Bn-*clg1A-1D*.

Supplemental Figure 11. Light Microscopy of Toluidine Blue–Stained Sections of Wild-Type and Bn-*clg1A-1D* Flowers.

Supplemental Figure 12. Scanning Electron Microscopy of *Arabidopsis*Bn-*clg1A-1D* Transgenic Plants.

Supplemental Table 1. Primers Used for Genetic Mapping and Quantitative PCR.

Supplemental Table 2. Comparison of Gene Content between the *Brassica napus* BAC Clone CZ7N2 and the Corresponding Region of *Arabidopsis* Chromosome 4.

Supplemental Table 3. Bn-*clg1A-1D*/ Bn-*CLG1A* Near-Isogenic Lines and Representative Oilseed Rape Cultivars Analyzed by the dCAPS_Clg1 Marker.

Supplemental Table 4. Bn-*clg1A-1D* Orthologs Used for Phylogenetic Tree Construction in Figure 5 and Supplemental Figure 6.

Supplemental Data Set 1. Test File of Alignment Used to Generate Phylogenetic Tree in Figure 5.

Supplemental Data Set 2. Test File of Alignment Used to Generate Phylogenetic Tree in Supplemental Figure 6.

Supplemental Results 1. Detailed Description of the Positional Cloning of Bn-*CLG1A*.

Supplemental Methods 1. Phenotyping, Molecular Marker Development, DNA Extraction, and Fine Genetic Mapping of Bn-*CLG1A*.

ACKNOWLEDGMENTS

This work was supported by the French Genoplante (GOP-T2 and GOP-T7) projects (http://www.genoplante.com/). D.A. was support by a Genopole-Evry postdoctorate fellowship. We thank Halima Morin and Nero Borrega from the Cytology and Cell Imaging Platform (Institut National de la Recherche Agronomique Versailles) for help with microscopy observations. We also thank France Denoeud from the Commissariat à l'Energie Atomique-Genoscope (Evry, France) for help obtaining homologs of the Bn-*CLG1A* gene from different species.

AUTHOR CONTRIBUTIONS

Y.-H.L., D.A., R.D., and B.C designed the project, conducted the experiments, and wrote the article. H.B., P.R., N.P., M.-O.L., and C.F. conducted experiments. J.J. conducted sequence analysis. M.R. contributed to the design of the project and to writing.

Received August 22, 2012; revised November 11, 2012; accepted December 5, 2012; published December 31, 2012.

REFERENCES

Ammitzbøll, H., and Bagger Jørgensen, R. (2006). Hybridization between oilseed rape (*Brassica napus*) and different populations and species of *Raphanus*. Environ. Biosafety Res. **5:** 3–13.

Beckie, H.J., Warwick, S.I., Nair, H., and Seguin-Swartz, G. (2003). Gene flow in commercial fields of herbicide-resistant canola (*Brassica napus*). Ecol. Appl. **13:** 1276–1294.

Borden, K.L. (2000). RING domains: Master builders of molecular scaffolds? J. Mol. Biol. **295:** 1103–1112.

Borden, K.L., and Freemont, P.S. (1996). The RING finger domain: A recent example of a sequence-structure family. Curr. Opin. Struct. Biol. **6:** 395–401.

Carroll, S.B. (2001). Chance and necessity: The evolution of morphological complexity and diversity. Nature **409:** 1102–1109.

Carvalho, P., Goder, V., and Rapoport, T.A. (2006). Distinct ubiquitin-ligase complexes define convergent pathways for the degradation of ER proteins. Cell **126:** 361–373.

Chalhoub, B., Belcram, H., and Caboche, M. (2004). Efficient cloning of plant genomes into bacterial artificial chromosome (BAC) libraries with larger and more uniform insert size. Plant Biotechnol. J. **2:** 181–188.

Chalhoub, B.A., Thibault, S., Laucou, V., Rameau, C., Höfte, H., and Cousin, R. (1997). Silver staining and recovery of AFLP amplification products on large denaturing polyacrylamide gels. Biotechniques **22:** 216–218, 220.

Chantret, N., et al. (2005). Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*). Plant Cell **17:** 1033–1045.

Chaudhary, B., Flagel, L., Stupar, R.M., Udall, J.A., Verma, N., Springer, N.M., and Wendel, J.F. (2009). Reciprocal silencing, transcriptional bias and functional divergence of homeologs in polyploid cotton (*Gossypium*). Genetics **182:** 503–517.

Cheng, F., Wu, J., Fang, L., Sun, S., Liu, B., Lin, K., Bonnema, G., and Wang, X. (2012). Biased gene fractionation and dominant gene expression among the subgenomes of *Brassica rapa*. PLoS ONE **7:** e36442.

Chhabra, A.K., and Sethi, S.K. (1991). Inheritance of cleistogamic flowering in durum wheat (*Triticum aestivum* L.). Euphytica **55:** 147–150.

Clough, S.J., and Bent, A.F. (1998). Floral dip: A simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J. **16:** 735–743.

Culley, T.M., and Klooster, M.R. (2007). The cleistogamous breeding system: A review of its frequency, evolution, and ecology in angiosperms. Bot. Rev. **73:** 1–30.

Daniell, H. (2002). Molecular strategies for gene containment in transgenic crops. Nat. Biotechnol. **20:** 581–586.

Darwin, C. (1877). The Different Forms of Flowers on Plants of the Same Species. (London: John Murray).

Fargue, A.N., Colbach, P.J., Picault, H., Renard M., and Meynard, J.M. (2006). Predictive study of the advantages of cleistogamy in oilseed rape in limiting unwanted gene flow. Euphytica **151:** 1–13.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. Evolution **39:** 783–791.

Ford, C.S., Allainguillaume, J., Grilli-Chantler, P., Cuccato, G., Allender, C.J., and Wilkinson, M.J. (2006). Spontaneous gene flow from rapeseed (*Brassica napus*) to wild *Brassica oleracea*. Proc. Biol. Sci. **273:** 3111–3115.

Halfhill, M.D., Millwood, R.J., Raymer, P.L., and Stewart, C.N. Jr. (2002). *Bt*-transgenic oilseed rape hybridization with its weedy relative, *Brassica rapa*. Environ. Biosafety Res. **1:** 19–28.

Hampton, R.Y. (2002). ER-associated degradation in protein quality control and cellular regulation. Curr. Opin. Cell Biol. **14:** 476–482.

Hassink, G., Kikkert, M., van Voorden, S., Lee, S.J., Spaapen, R., van Laar, T., Coleman, C.S., Bartee, E., Früh, K., Chau, V., and Wiertz, E. (2005). TEB4 is a C4HC3 RING finger-containing ubiquitin ligase of the endoplasmic reticulum. Biochem. J. **388:** 647–655.

Hovav, R., Udall, J.A., Chaudhary, B., Rapp, R., Flagel, L., and Wendel, J.F. (2008). Partitioned expression of duplicated genes during development and evolution of a single cell in a polyploid plant. Proc. Natl. Acad. Sci. USA **105:** 6191–6195.

Hwang, C.S., Shemorry, A., and Varshavsky, A. (2010). N-terminal acetylation of cellular proteins creates specific degradation signals. Science **327:** 973–977.

Isidore, E., Scherrer, B., Bellec, A., Budin, K., Faivre-Rampant, P., Waugh, R., Keller, B., Caboche, M., Feuillet, C., and Chalhoub, B. (2005). Direct targeting and rapid isolation of BAC clones spanning a defined chromosome region. Funct. Integr. Genomics **5:** 97–103.

Jefferson, R.A., Kavanagh, T.A., and Bevan, M.W. (1987). GUS fusions: β-Glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. **6:** 3901–3907.

Jørgensen, R.B., and Andersen, B. (1994). Spontaneous hybridization between oilseed rape (*Brassica napus*) and weedy *B. campestris* (Brassicaceae): A risk of genetically modified oilseed rape. Am. J. Bot. **81:** 1520–1526.

Koornneef, M., Hanhart, C.J., and Thiel, F. (1989). A genetic and phenotypic description of *Eceriferum* (*cer*) mutants in *Arabidopsis thaliana*. J. Hered. **80:** 118–122.

Kraft, E., Stone, S.L., Ma, L., Su, N., Gao, Y., Lau, O.S., Deng, X.W., and Callis, J. (2005). Genome analysis and functional characterization of the E2 and RING-type E3 ligase ubiquitination enzymes of *Arabidopsis*. Plant Physiol. **139:** 1597–1611.

Kreft, S.G., Wang, L., and Hochstrasser, M. (2006). Membrane topology of the yeast endoplasmic reticulum-localized ubiquitin ligase *Doa10* and comparison with its human ortholog *TEB4* (*MARCH-VI*). J. Biol. Chem. **281:** 4646–4653.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J. Mol. Biol. **305:** 567–580.

Kumar, N., Srivastava, G.C., and Dixit, K. (2008). Flower bud opening and senescence in roses (*Rosa hybrida* L.). Plant Growth Regul. **55:** 81–99.

Kurauchi, N., Makino, T., and Hirose, S. (1993). Inheritance of cleistogamy-chasmogamy in barley. Barley Genet. Newsl. **23:** 19.

Kwon, Y.W., Kim, D.S., and Yim, K.O. (2001). Herbicide-resistant genetically modified crop: assessment and management of gene flow. Weed Biol. Manage. **1:** 42–51.

Li-Beisson, Y., Pollard, M., Sauveplane, V., Pinot, F., Ohlrogge, J., and Beisson, F. (2009). Nanoridges that characterize the surfacemorphology of flowers require the synthesis of cutin polyester. Proc. Natl. Acad. Sci. USA **106:** 22008–22013.

Lord, E.M. (1981). Cleistogamy: A tool for the study of floral morphogenesis, function and evolution. Bot. Rev. **47:** 421–449.

Lorick, K.L., Jensen, J.P., Fang, S., Ong, A.M., Hatakeyama, S., and Weissman, A.M. (1999). RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. Proc. Natl. Acad. Sci. USA **96:** 11364–11369.

Lovering, R., Hanson, I.M., Borden, K.L., Martin, S., O'Reilly, N.J., Evan, G.I., Rahman, D., Pappin, D.J., Trowsdale, J., and Freemont, P.S. (1993). Identification and preliminary characterization of a protein motif related to the zinc finger. Proc. Natl. Acad. Sci. USA **90:** 2112–2116.

Lu, B.R. (2003). Transgene containment by molecular means—Is it possible and cost effective? Environ. Biosafety Res. **2:** 3–8.

Lü, S., Zhao, H., Des Marais, D.L., Parsons, E.P., Wen, X., Xu, X., Bangarusamy, D.K., Wang, G., Rowland, O., Juenger, T., Bressan, R.A., and Jenks, M.A. (2012). *Arabidopsis ECERIFERUM9* involvement in cuticle formation and maintenance of plant water status. Plant Physiol. **159:** 930–944.

Lynch, M., and Force, A. (2000). The probability of duplicate gene preservation by subfunctionalization. Genetics **154:** 459–473.

Lysak, M.A., Cheung, K., Kitschke, M., and Bures, P. (2007). Ancestral chromosomal blocks are triplicated in *Brassiceae* species with varying chromosome number and genome size. Plant Physiol. **145:** 402–410.

Maeng, J.Y., Won, Y.J., Piao, R., Cho, Y.I., Jiang, W., Chin, J.H., and Koh, H.J. (2006). Molecular mapping of a gene '*Id(t)*' controlling cleistogamy in rice. Theor. Appl. Genet. **112:** 1429–1433.

Merwin, N.C., Glurey, L.M., and Blackwell, K.H. (1981). Inheritance of papery glume and cleistogamy in sorghum. Crop Sci. **21:** 953–956.

Moloney, M.M., Walker, J.M., and Sharma, K.K. (1989). High efficiency transformation of *Brassica napus* using *Agrobacterium* vectors. Plant Cell Rep. **8:** 238–242.

Nair, S.K., et al. (2010). Cleistogamous flowering in barley arises from the suppression of microRNA-guided *HvAP2* mRNA cleavage. Proc. Natl. Acad. Sci. USA **107:** 490–495.

Panikashvili, D., Shi, J.X., Schreiber, L., and Aharoni, A. (2009). The *Arabidopsis DCR* encoding a soluble BAHD acyltransferase is required for cutin polyester formation and seed hydration properties. Plant Physiol. **151:** 1773–1789.

Panikashvili, D., Shi, J.X., Schreiber, L., and Aharoni, A. (2011). The *Arabidopsis* ABCG13 transporter is required for flower cuticle secretion and patterning of the petal epidermis. New Phytol. **190:** 113–124.

Parkin, I.A., Gulden, S.M., Sharpe, A.G., Lukens, L., Trick, M., Osborn, T.C., and Lydiate, D.J. (2005). Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. Genetics **171:** 765–781.

Polsoni, L., Kott, S., and Beversdorf, W.D. (1988). Large-scale microspore culture technique for mutation selection studies in *Brassica napus*. Can. J. Bot. **66:** 1681–1685.

Prince, V.E., and Pickett, F.B. (2002). Splitting pairs: The diverging fates of duplicated genes. Nat. Rev. Genet. **3:** 827–837.

Pyke, K.A., and Page, A.M. (1998). Plastid ontogeny during petal development in *Arabidopsis*. Plant Physiol. **116:** 797–803.

Renard, M., and Tanguy, X. (1997). Obtention de mutants cléistogames de crucifères. Brevet FR 9715768.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:** 406–425.

Saxena, K.B., Ariyanayagam, R.P., and Reddy, L.J. (1992). Genetics of a high-selfing trait in pigeonpea. Euphytica **59:** 2–3.

Saxena, K.B., Singh, L., and Ariyanayagam, R.P. (1993). Role of partial cleistogamy in maintaining genetic purity of pigeonpea. Euphytica **66:** 225–229.

Schemske, D.W. (1978). Evolution of reproductive characteristics in *Impatiens* (*Balsaminaceae*): The significance of cleistogamy and chasmogamy. Ecology **59:** 596–613.

Shi, J.X., Malitsky, S., De Oliveira, S., Branigan, C., Franke, R.B., Schreiber, L., and Aharoni, A. (2011). SHINE transcription factors act redundantly to pattern the archetypal surface of *Arabidopsis* flower organs. PLoS Genet. **7:** e1001388.

Schranz, M.E., and Mitchell-Olds, T. (2006). Independent ancient polyploidy events in the sister families *Brassicaceae* and *Cleomaceae*. Plant Cell **18:** 1152–1165.

Silva, N.F., and Goring, D.R. (2001). Mechanisms of self-incompatibility in flowering plants. Cell. Mol. Life Sci. **58:** 1988–2007.

Smalle, J., and Vierstra, R.D. (2004). The ubiquitin 26S proteasome proteolytic pathway. Annu. Rev. Plant Biol. **55:** 555–590.

Song, X., Wang, Z., Zuo, J., Huangfu, C., and Qiang, S. (2010). Potential gene flow of two herbicide-tolerant transgenes from oilseed rape to wild *B. juncea* var. gracilis. Theor. Appl. Genet. **120:** 1501–1510.

Sparrow, P.A.C., Townsend, T.M., Morgan, C.L., Dale, P.J., Arthur, A.E., and Irwin, J.A. (2004). Genetic analysis of in vitro shoot regeneration from cotyledonary petioles of *Brassica oleracea*. Theor. Appl. Genet. **108:** 1249–1255.

Stone, S.L., Hauksdóttir, H., Troy, A., Herschleb, J., Kraft, E., and Callis, J. (2005). Functional analysis of the RING-type ubiquitin ligase family of *Arabidopsis*. Plant Physiol. **137:** 13–30.

Swanson, R., Locher, M., and Hochstrasser, M. (2001). A conserved ubiquitin ligase of the nuclear envelope/endoplasmic reticulum that functions in both ER-associated and *Matalpha2* repressor degradation. Genes Dev. **15:** 2660–2674.

Takahashi, R., Kurosaki, H., Yumoto, S., Han, O.K., and Abe, J. (2001). Genetic and linkage analysis of cleistogamy in soybean. J. Hered. **92:** 89–92.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. **24:** 1596–1599.

Tang, H., Woodhouse, M.R., Cheng, F., Schnable, J.C., Pedersen, B.S., Conant, G., Wang, X., Freeling, M., and Pires, J.C. (2012). Altered patterns of fractionation and exon deletions in *Brassica rapa* support a two-step model of paleohexaploidy. Genetics **190:** 1563–1574.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997). The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25:** 4876–4882.

Turuspekov, Y., Mano, Y., Honda, I., Kawada, N., Watanabe, Y., and Komatsuda, T. (2004). Identification and mapping of cleistogamy genes in barley. Theor. Appl. Genet. **109:** 480–487.

Uphof, J.C.T. (1938). Cleistogamic flowers. Bot. Rev. **4:** 21–49.

van Doorn, W.G., and Van Meeteren, U. (2003). Flower opening and closure: A review. J. Exp. Bot. **54:** 1801–1812.

van Engelen, F.A., Molthoff, J.W., Conner, A.J., Nap, J.P., Pereira, A., and Stiekema, W.J. (1995). pBINPLUS: An improved plant transformation vector based on pBIN19. Transgenic Res. **4:** 288–290.

Vierstra, R.D. (2009). The ubiquitin-26S proteasome system at the nexus of plant biology. Nat. Rev. Mol. Cell Biol. **10:** 385–397.

von Heijne, G., Steppuhn, J., and Herrmann, R.G. (1989). Domain structure of mitochondrial and chloroplast targeting peptides. Eur. J. Biochem. **180:** 535–545.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., and Zabeau, M. (1995). AFLP: A new technique for DNA fingerprinting. Nucleic Acids Res. **23:** 4407–4414.

Waller, D.M. (1984). Differences in fitness between seedlings derived from cleistogamous and chasmogamous flowers in *Impatiens capensis*. Evolution **38:** 427–440.

Wang, X., et al; *Brassica rapa* Genome Sequencing Project Consortium (2011). The genome of the mesopolyploid crop species *Brassica rapa*. Nat. Genet. **43:** 1035–1039.

Warwick, S.I., Simard, M.J., Légère, A., Beckie, H.J., Braun, L., Zhu, B., Mason, P., Séguin-Swartz, G., and Stewart, C.N. Jr.. (2003). Hybridization between transgenic *Brassica napus* L. and its wild relatives: *Brassica rapa* L., *Raphanus raphanistrum* L., *Sinapis arvensis* L., and *Erucastrum gallicum* (Willd.) O.E. Schulz. Theor. Appl. Genet. **107:** 528–539.

Yoshida, H., Itoh, J.I., Ohmori, S., Miyoshi, K., Horigome, A., Uchida, E., Kimizu, M., Matsumura, Y., Kusaba, M., Satoh, H., and Nagato, Y. (2007). *superwoman1*-cleistogamy, a hopeful allele for gene containment in GM rice. Plant Biotechnol. J. **5:** 835–846.

Zhang, Z., et al. (2011). Duplication and partitioning in evolution and function of homoeologous *Q* loci governing domestication characters in polyploid wheat. Proc. Natl. Acad. Sci. USA **108:** 18737–18742.

Zuckerkandl, E., and Pauling, L. (1965). Evolutionary divergence and convergence in proteins. In Evolving Genes and Proteins, V. Bryson and H.J. Vogel, eds (New York: Academic Press), pp. 97–166.

Zybailov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q., and van Wijk, K.J. (2008). Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. PLoS One **3:** e1994.

Supplemental Results

Detailed Description of the Positional Cloning of Bn-*CLG1A*.

Phenotyping and Primary Molecular Mapping

Primary molecular mapping of the cleistogamy trait was initiated using a segregating population of 255 doubled haploid (DH) lines of a cross between the French cleistogamous line 'B001-*Clg*' and a genetically-distant Korean cultivar called 'Yudal' (Figure 1).

We observed that the cleistogamy trait can be scored into five levels (noted 1 to 5) that can be grouped into three main categories (Supplemental Figure 1). Non-cleistogamous plants (note 1) show completely-open flowers similar to those of the Yudal parental line and represent 53% of segregating DH lines (Supplemental Figure 1). Cleistogamous plants and show all flowers completely-closed similar to those of the 'B001-*Clg*' parental line (note 5) or all flowers completely-closed with the sporadic exception of one or two incompletely-closed flowers (note 4) and represent together 20% of segregating DH lines (Supplemental Figure 1). Finally, intermediate-cleistogamous plants show incompletely-closed flowers with two levels (notes 2 and 3) and represent together the remaining 27% of segregating DH lines (Supplemental Figure 1). The intermediate-cleistogamous phenotype was also observed in other homozygous cleistogamous oilseed rape lines in which the Bn-*clg1A-1D* gene was introgressed, having thus a Bn-*clg1A-1D*Bn-*clg1A-1D* genotype (Fargue et al., 2006, Renard and Tanguy, 1997). This observation, added to the fact that intermediate-cleistogamous DH lines are also homozygous, suggest genetic background effects on the cleistogamy phenotype. Therefore for molecular mapping, phenotypes were converted into a Mendelian scoring where non-cleistogamous plants were considered as having the Bn-*CLG1A*Bn-*CLG1A* genotype whereas together intermediate-cleistogamous and cleistogamous ones were considered together as having the Bn-*clg1A-1D*Bn-*clg1A-1D* genotype. Genotypes of the different phenotyped cleistogamous categories and levels were checked *a posteriori* after the gene sequence was identified and it was possible to develop a molecular marker distinguishing the Bn-*clg1A-1D* allele from the Bn-*CLG1A* allele (*see below*). The observed segregation of 53% of non cleistogamous DH lines to 47% of intermediate-cleistogamous and cleistogamous ones (Supplemental Figure 1) fits a monogenic model, with Bn-*clg1A-1D* allele leading to cleistogamy or partial cleistogamy in this particular cross. Noteworthy, monogenic control of cleistogamy phenotype was previously suggested using a

different genetic F2 segregating population where about 25% of F2 plants were cleistogamous and 50% exhibiting partially closed flowers (Renard and Tanguy, 1997).

We initially used AFLP molecular markers (Supplemental Table 1) and bulk segregant analysis (BSA), using bulks of non-cleistogamous intermediate- and complete-cleistogamous DH lines and located Bn-*clg1A-1D* to a chromosomal region associated with five linked markers (Figure 2A) on the linkage group DY1b-N1 (Lombard and Delourme, 2001; Delourme et al. 2006). A QTL mapping approach (by MapMaker/QTL v1.1, Lander and Bostein, 1989) was applied *a posteriori* on the linkage groups and the results showed that the Bn-*clg1A-1D* locus explained 72.6% of the total phenotypic variance in the studied segregation population.

Fine Genetic and Physical Mapping of the Bn-*clg1A-1D* Gene

For fine genetic and physical mapping of Bn-*clg1A-1D*, we used a combination of genetic and comparative genomic approaches with Arabidopsis. All AFLP markers identified as linked to Bn-*clg1A-1D* were sequenced. Sequence of the AFLP marker T2A02 that mapped at 0 cM from Bn-*clg1A-1D* (Figure 2A) had 90% sequence identity with the *At4g33760* gene, located on the long-arm of *Arabidopsis* chromosome 4.

Thirty-two *Arabidopsis* genes located in an interval of 2,030 kb around *At4g33760* gene (from *At4g31810* to *At4g36930* genes), have then been used for generating PCR-based marker for genetic mapping in *B. napus* (Figure 2B). The use of different primer combinations generated from each single Arabidopsis gene allowed the detection of different homologous gene copies in oilseed rape, because of its highly duplicated genome. Nevertheless, among these, seven PCR-based markers (in addition to T2S005 and T2S008, converted from the AFLP marker T2A02 and T2A01, respectively) were polymorphic and mapped to the Bn-*clg1A-1D* gene region of oilseed rape across a 14.9 cM genetic interval, confirming the orthologous relationship with *Arabidopsis* (Figures 2B and 2C; Supplemental Table 1).

A better resolution of the genetic map was then realized by genotyping five of these markers (T2G108, T2G115, T2S005, T2S008 and T2G107) on a larger population of 2,158 DH lines derived from the cross between 'B001-*Clg*' and 'Yudal' (Figure 2D). The Bn-*clg1A-1D* locus was fine-positioned across a 1.6-cM interval, flanked by T2S005 and T2S008 markers and where 31 recombinant lines were revealed (Figure 2D). All the 31 recombinant lines were

checked several times with the molecular markers and were re-phenotyped in field conditions in order to confirm the patterns of recombination.

To identify BAC clones spanning the Bn-*clg1A-1D* gene, we screened an existing BAC library of the oilseed rape cultivar 'Darmor-*bzh*' (Chalhoub et al., 2004), which does not carry the cleistogamous mutation, with the closest flanking markers T2S005 and T2S008. Four BAC clones (CZ7N2, CZ19L14, CZ49B4 and CZ55N3) were shown to carry both flanking genetic markers and thus the non-cleistogamous allele (Bn-*CLG1A*) (Figure 2E).

Sequencing of Bn-*CLG1A* Genomic Region and Further Fine Mapping

The longest BAC clone CZ7N2 (~250 kb) containing the expected Bn-*CLG1A* allele (and the closest flanking markers) was sequenced. This BAC clone contains 72 predicted genes and three sequences of repeats and is well conserved in comparison to the orthologous regions in Arabidopsis (Supplemental Table 2, Supplemental Figure 2). The Bn-*clg1A-1D* flanking markers T2S005 and T2S008 were shown to delimit a ~150 kb interval containing 35 predicted genes. This suggests that the genetic distance of 1.6 cM should correspond to ~100 kb/cM around the Bn-*clg1A-1D* locus.

We developed additional PCR-based markers from genomic sequence spanning the Bn-*clg1A-1D* gene. Five new PCR-based markers, T2M004, T2M009, T2S027, T2043F/T2S051R, T2S064 were polymorphic and mapped on the 31 recombinant lines, allowing us to reduce the *Bn-clg1A-1D* interval from 150 kb to 8 kb and to reduce the number of candidate genes from 35 to only one (Figure 2F). This gene, delimited by T2S064 and T2S027 genetic markers, is orthologous to *At4g34100* gene of Arabidopsis and encodes a putative RING variant (RINGv) domain containing protein belonging to E3 ubiquitin ligase family.

Supplemental References

Delourme, R., Falentin, C., Huteau, V., Clouet, V., Horvais, R, Gandon, B., Specel, S., Hanneton, L., Dheu, J.E., Deschamps, M., Margale, E., Vincourt, P., and Renard, M. (2006). Genetic control of oil content in oilseed rape (*Brassica napus* L.). Theor. Appl. Genet. **113:** 1331-1345.

Lombard, V., and De lourme, R. (2001). A consensus linkage m ap for rapeseed (*Brassica napus* L.): Construction and integration of th ree individual m aps from DH populations. Theor. Appl. Genet. **103:** 491–507.

Supplemental Method.

Phenotyping, Molecular Marker Development, DNA Extraction and Fine Genetic Mapping of Bn-*CLG1A*.

Phenotyping of the Cleistogamy Trait

In both greenhouse and field conditions, several individual plants of each DH line were phenotyped visually every 2 to 3 days during the flowering stage (from appearance of the first flower bud). The cleistogamy phenotypes were scored into five levels (notes 1 to 5) according to their degree of cleistogamy, with level 1 being non-cleistogamous with completely open flowers and levels 4 and 5 being cleistogamous with either plants having all flowers completely closed (note 5) or plants having all flowers completely closed with sporadically, the exception of one or two incompletely-closed flowers (note 4). Finally, levels 2 and 3 were intermediate-cleistogamous (Supplemental Figure 1 online). Scoring was further converted into a Mendelian for genetic analysis where Cleistogamous and intermediate-cleistogamous DH lines were considered as homozygous for the mutated allele (Bn-*clg1A-1D*Bn-*clg1A-1D*) and non-cleistogamous DH lines were considered as homozygous for the native allele (Bn-*CLG1A*Bn-*CLG1A*) (Supplemental Figure 1 online).

DNA Extraction and Construction of DNA Pools

For initial mapping, genomic DNA from each of the 255 DH lines and the two parental genotypes (Yudal and B001-*Clg*) was extracted from 2-5 g of young leave tissues using a cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle, 1990). The DNA concentration in each sample was quantified and adjusted to 20 ng/µL. Two sets of three DNA pools (A1, B1 and C1; A2, B2 and C2) were constructed for bulked segregant analysis (BSA, Michelmore et al., 1991) using AFLP. The pools were constructed by mixing equal quantities of DNA from 12 DH lines with either normally opened flower type (note 1, pools A1 and A2), intermediate closed flower type (notes 2 and 3, pools B1 and B2) or completely closed flower type (notes 4 and 5, pools C1 and C2).

For fine mapping use, genomic DNA from 2,158 DH lines was extracted by an optimized 96-well plate, high-throughput method. Briefly, about 50 mg of young leaf tissue of each DH line was harvested and lyophilized in 96 well micro titre plates containing a 3 mm stainless steel

bead in each well. The samples were grounded into fine powder by using a Fluid Management paint shaker. 600 µL of preheated lysis buffer (500 mM NaCl, 100 mM Tris, 50 mM EDTA and 3.8g/L sodium metabisulfite) were added. The samples were homogenized by pipettings and then incubated for 1 hour at 95°C. The plates were then centrifuged for 20 minutes at 3400*g*. 300 µL cold isopropanol and 75 µL 8 M ammonium acetate was added to 400 µL of supernatant. Plates were incubated at 4°C for 30-60 minutes and centrifuged at 3400 *g* for 30 minutes. The DNA pellets were washed with 70% cold, dried and resuspended in 50 µL TE (Tris-HCl 10 mM, 1 mM EDTA, pH=8.0). The samples were further diluted 30 times to give a ready-to-use solution for PCR reactions.

AFLP Markers and Conversion to Sequence-Specific PCR Markers

AFLP analysis was performed essentially as described by Vos et al. (1995) using the AFLP Analysis System I and AFLP starter Primer Kits (Gibco BRL Life Technologies) following the manufacturer's instructions. Briefly, about 350 ng of genomic DNA was firstly digested with 20 units of restriction enzyme *Eco*RI and *Mse*I (Roche) in a final volume of 20 µL. The digested DNA was ligated to the double-stranded *Eco*RI and *Mse*I adapters using T4 ligase, pre-amplified using AFLP primers with one selective nucleotide and selectively amplified using combinations of *Eco*RI+3 and *Mse*I+3 primers. The amplified products were separated on 30-cm -long 6% denaturing polyacrylamide gels in 1 X TBE buffer. The amplified fragments were visualized by silver nitrate staining as described by Chalhoub et al. (1997).

The polymorphic AFLP fragments were cut out from air-dried denaturing polyacrylamide gels and eluted overnight in 100 µL sterile water at 4°C. Two aliquots (2.5 µL and 5 µL) were used as template DNA and re-amplified in a volume of 100 µL using the same selective primer combinations and conditions that generated the AFLP fragments. The quality and quantity of the re-amplified fragments were evaluated by running an aliquot (10 µL) of resulting PCR products onto a 2 % agarose gel. Fragments accompanied with co-amplification of non-desired fragments were re-purified by another round of "cut-out-of-gel" and PCR re-amplification. The purified resulting PCR products were sequenced by Genomexpress (Grenoble, France). PCR-specific primers were designed manually in combination with the use of Oligo 6.0 software (Rychlik and Rhoads, 1989) from internal sequences of corresponding AFLP primers. These primer pairs were tested by PCR (20 µL) on the two sets of DNA pools plus two parental lines 'Yudal' and 'B001-

Clg' for polymorphism checking. An aliquot (10µL) of monomorphic PCR products was digested by *Hae*III and electrophoresed on a 2% agarose gel.

ACGM, PFM and CAPS Markers

Three types of PCR-based markers were developed from Arabidopsis gene sequences and used to detect polymorphism and mapping in *B. napus.*

ACGMs (Amplified consensus gene markers), with prefix 'T2G-', were developed from Arabidopsis gene sequences as described by Brunel et al. (1999). Individual genomic gene sequences were downloaded from the public database of The Arabidopsis Information Resource (TAIR, http://www.arabidopsis.org/). For each selected gene, two forward (_F1 and _F2) and two reverse (_R1 and _R2) primers were designed on exons using Oligo v 6.0 software (Rychlik and Rhoads 1989) to give four possible primer combinations (_F1/_R1, _F1/_R2, _F2/_R1 and F2/_R2) spanning one or more introns of the gene. These primer combinations were used to screen the sets of DNA pools along with the two parent DNAs ('Yudal' and 'B0101_ *Clg*') in a volume of 25 µL. An aliquot (10µL) of the PCR products was firstly analyzed on a 1% agarose gel for identifying combinations giving polymorphism and/or allowing specific amplifications with well visible bands of expected size on the ethidium bromide stained gel.

STS (Sequence tagged site, with prefix T2S -) markers were generated from annotated *B. napus* BAC sequences.

For all three types of PCR-based markers, primer combinations that give specific amplifications on the oilseed rape DNAs, **CAPS** (Cleaved Amplified Polymorphic Sequences) were systematic developed by digestion of PCR products of sequence-specific primer pairs (ACGM or STS) by the restriction enzymes *Mse*I or *Hae*III and resolution on 30 cm -long 6% denaturing polyacrylamide gels in 1 x TBE buffer.

A **CAPS** marker involving the use of *Bgl II* enzyme was developed for the detection of the causal SNP of Bn-*clg1A-D* mutant as described in Supplemental Figure 3 online.

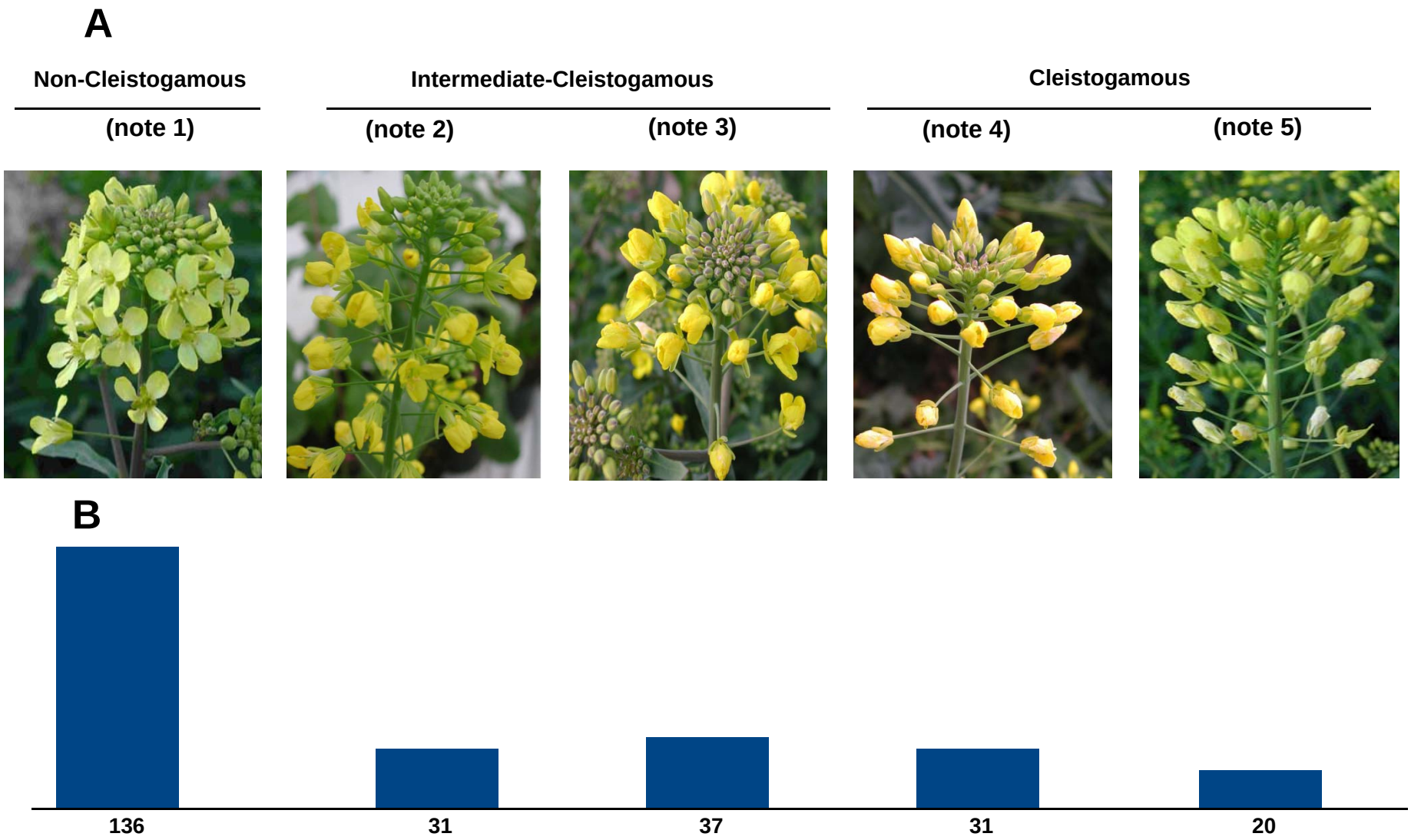
PCR analyses were carried out in 10 µL volumes of reaction mixture containing 0.4 µM of each primers, 50 mM KCl, 10 mM Tris-HCl pH 8.3, 1.5 mM $MgCl_2$, 20% of loading buffer (60% w/v sucrose, 5 mM cresol red in water), 0.25 U AmpliTaq Gold® DNA Polymerase (PE

Applied Biosystems), and 3 µL of tem plate DNA dilution (5~10 ng/µL). The reactions were prepared by using a Beckm an Biomeck 2000 robot system in 384-well plates coverrd with an adhesive foil (ABgene) on a GeneAm p PCR system 9700 thermocycler (P E Applied Biosystems). The therm al cycling program used consisted of one cycle of 4 m in at 94°C, followed by 35 cycles of 30 sec at 94°C, 30 sec at 56°C, and 1 min at 72°C, and a final cycle of 7 min at 72°C.

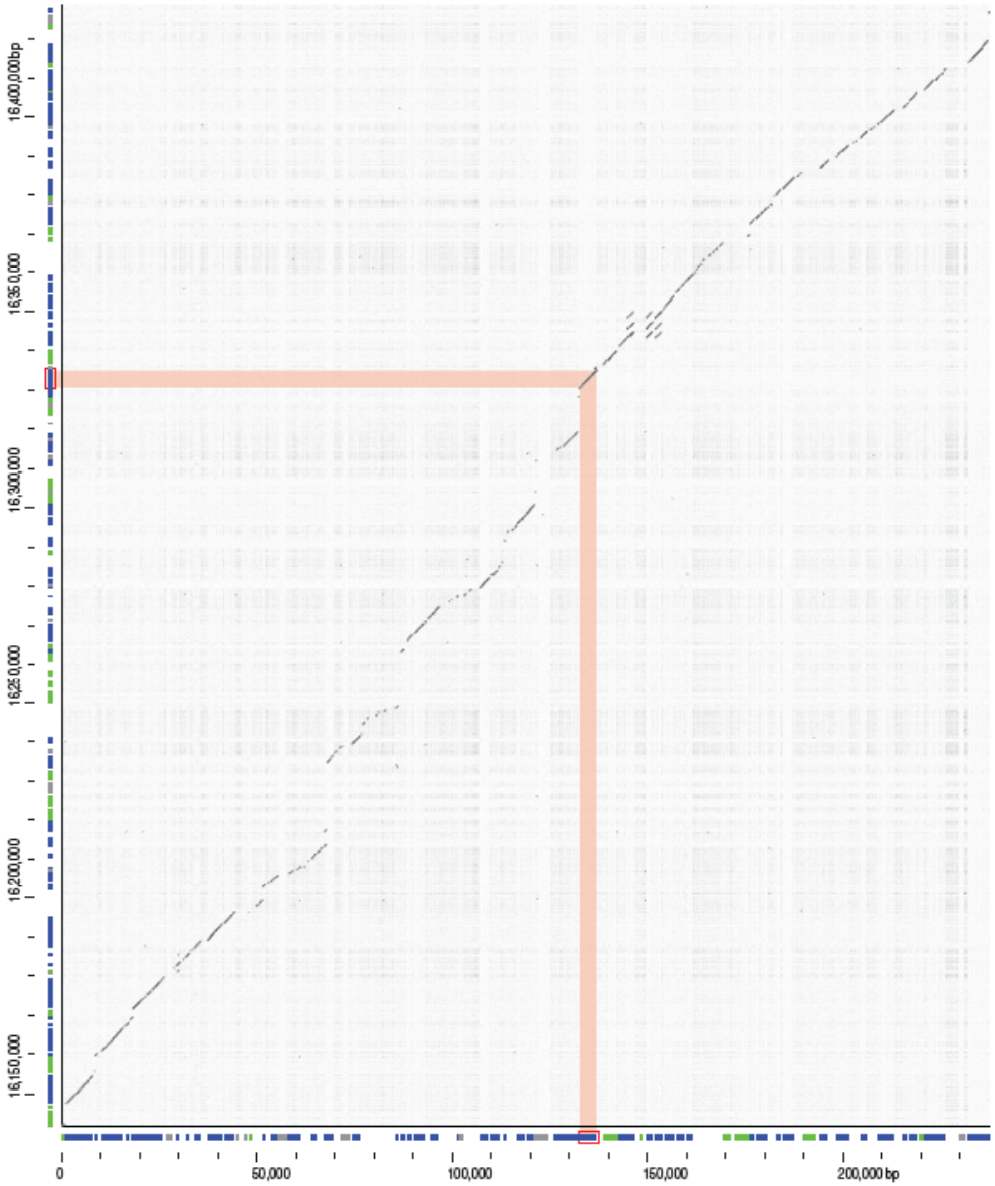
Linkage Analysis and Mapping

The multi-scored phenotypic data o f each DH line were averaged and converted into Mendelian data. The non-cleistog amous lines with an average sc ore value of 1 were genotype d as Bn-*CLG1A*Bn-*CLG1A* and the intermediate cleistogamous or cleistogamous lines with a value of 2-to 5 were genotyped as Bn- *clg1A-1D*Bn-*clg1A-1D* and analyzed together with m olecular markers.

The genotypic segregation data were analy zed with Mapmaker/Exp v3 .0 (Lincoln et al., 1992). Map distances were estimated with the Kosambi function (Kosambi 1944). Linkage maps were drawn using MapDisto software (Lorieux et al., 2000).



Supplemental Figure 1. Scoring and segregation of the cleistogamy phenotype in a population of 255 doubled haploid lines of *B. napus*. **(A)** and **(B)**: respectively scoring and segregation of the cleistogamy phenotype into five levels (1 to 5) (representing three main categories). Non-cleistogamous (note 1) ,with completely open flowers; Intermediately-cleistogamous (notes 2 and 3), with incompletely-closed flowers and; Cleistogamous (notes 4 and 5) with completely-closed flowers. A cross between the French cleistogamous line 'B001-*Clg*' and a genetically-distant Korean cultivar called 'Yudal'.

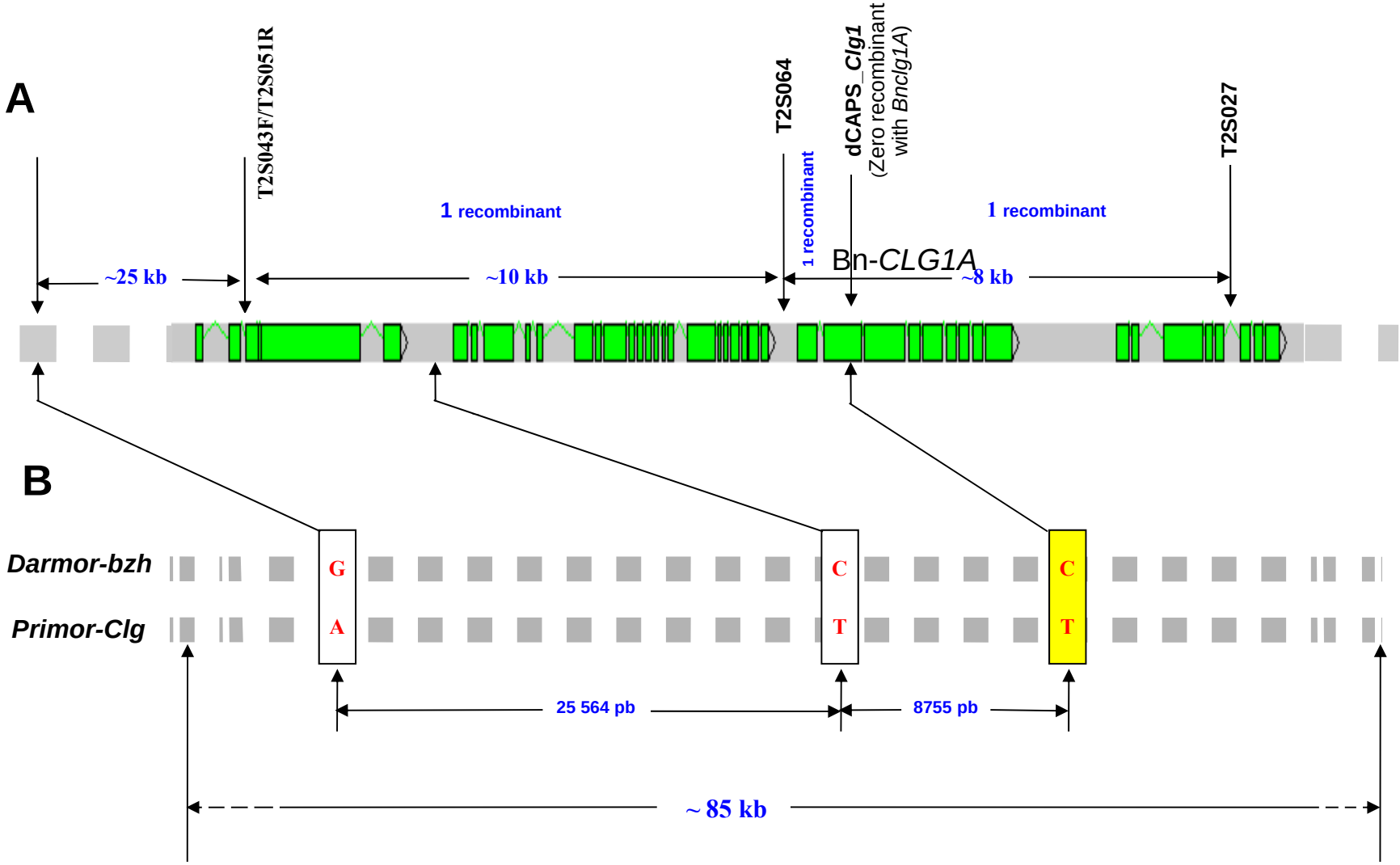


Supplemental Figure 2. Dotplot comparison between the *Brassica napus* region spanning the Bn-*CLG1A* wild type gene and the orthologous genomic sequence from *Arabidopsis thaliana*

Brassica napus region (BAC clone CZ7N2, horizontal, X axis.

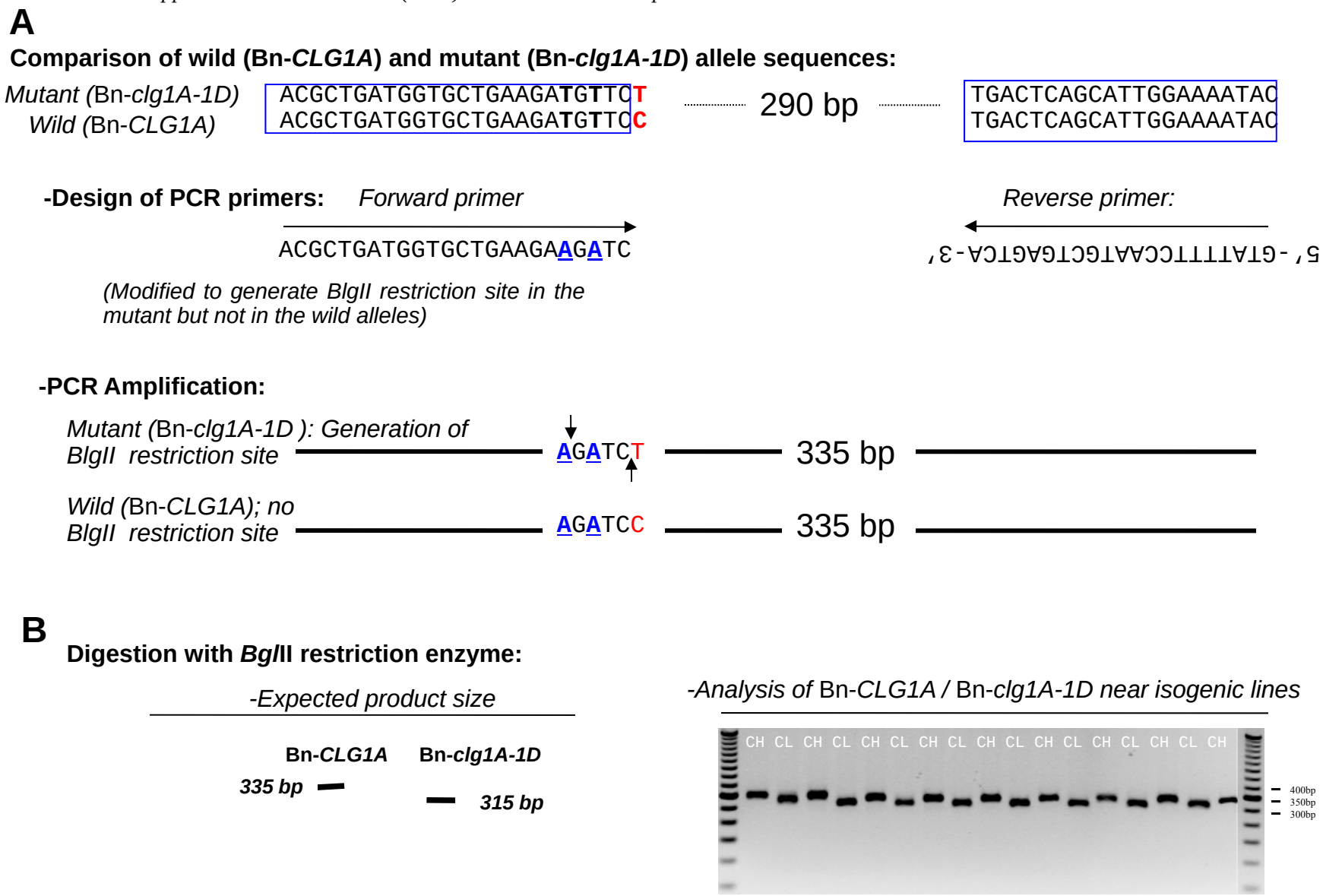
Orthologous genomic sequence from *Arabidopsis thaliana* (vertical, Y axis). Blue boxes: genes conserved between both species, Green boxes: genes present in one species and missing in the other species. *B. napus* Bn-*CLG1A* wild type gene and its corresponding orthologous At4g34100 in *A. thaliana* are marked in red.

Dotplot : Sonnhammer and Durbin, 1995). Annotation was done as in Chantret et al. (2005).

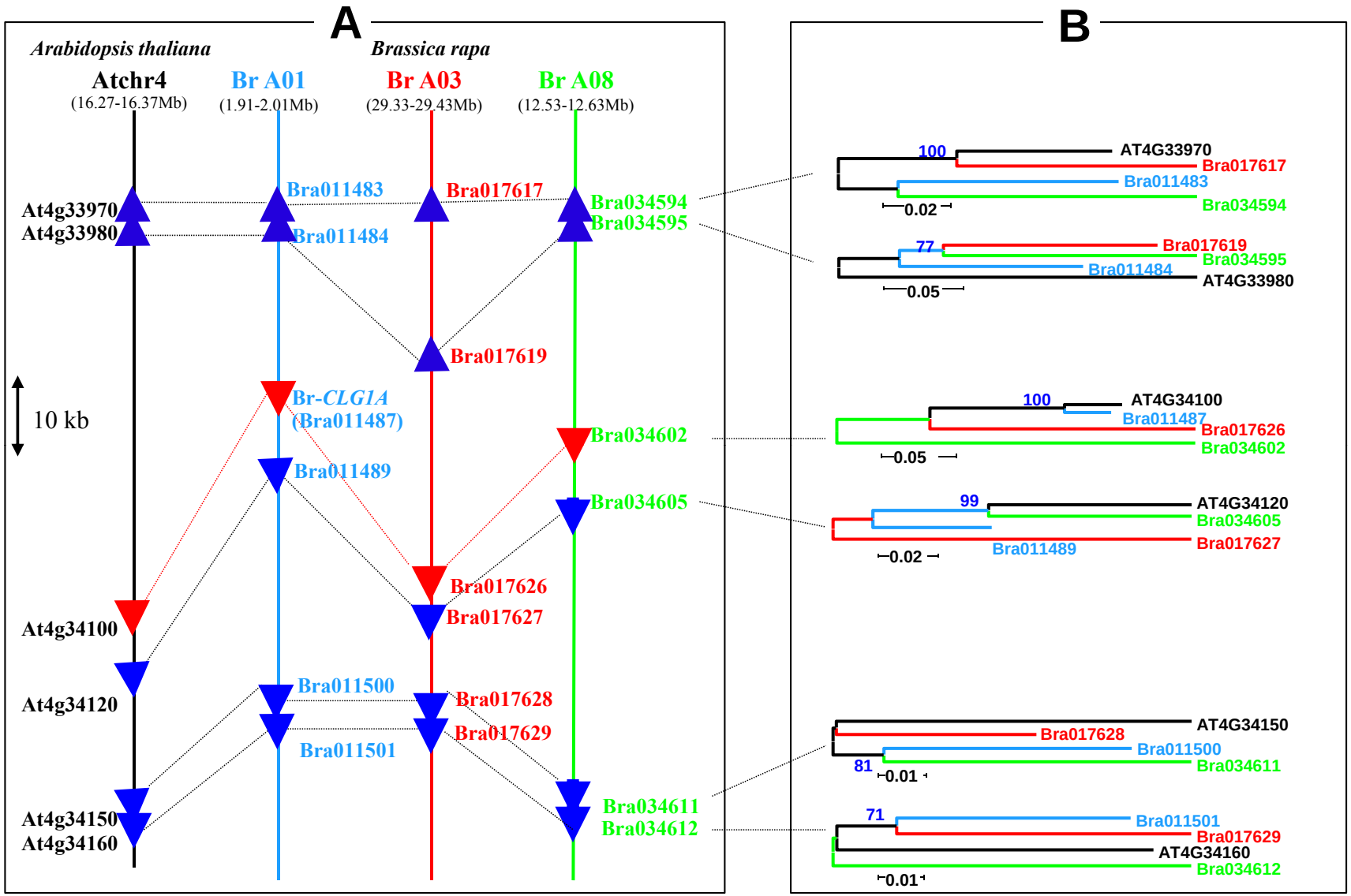


Supplemental Figure 3. Physical positions of the genetic markers closely-linked to Bn-*CLG1A* and sequence comparison between Bn-*CLG1A* and Bn-*clg1A-1D*.

(A) Bn-*CLG1A* gene (from *cv* Darmor-*bzh*, BAC clone CZ7N2) and the mutated Bn-*clg1A-1D* allele region (from Primor_*Clg*, BAC clone Clg1_H1B_P13a_3_J1). **(B)** Only three SNPs were identified between Darmor-*bzh* and Primor_*Clg* within a ~85-kb sequence interval of Bn-*clg1A-1D*/ Bn-*CLG1A* locus regions. The causal (C-to-T) is located inside Bn-*CLG1A* candidate gene (yellow box).

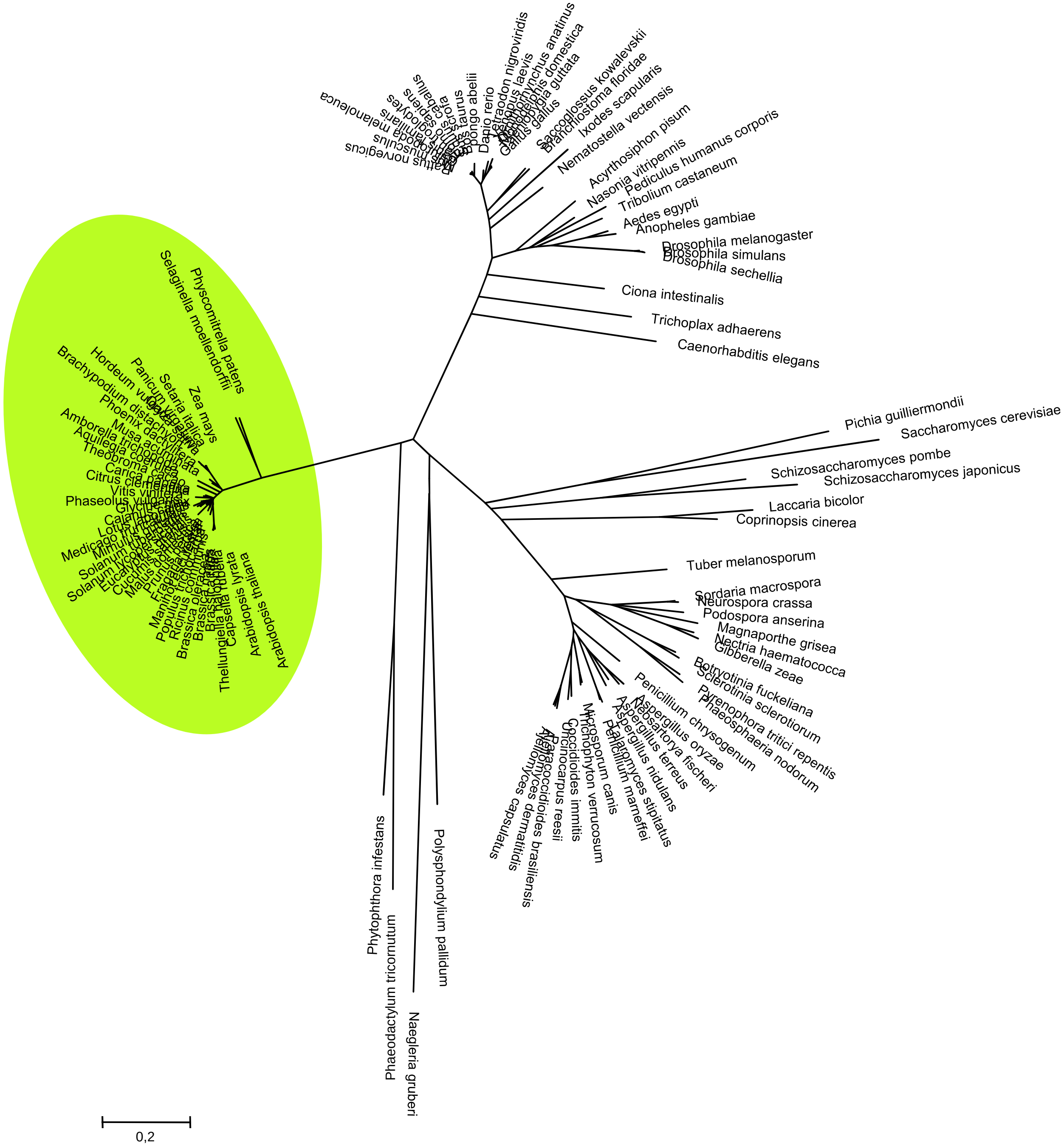


Supplemental Figure 4. Design of a dCAPS marker for the detection of the T/C SNP differentiating Bn-*clg1A-1D* and Bn-*CLG1A*.
(A) Design of modified forward PCR primer allowing generation of *Blg*II restriction site in the mutated Bn-*clg1A-1D* allele and not in the wild Bn-*CLG1A* allele. **(B)** Results obtained on cleistogamous / non-cleistogamous *B. napus* near isogenic lines showing perfect association with the Bn-*clg1A-1D* and Bn-*CLG1A* alleles respectively.



Supplemental Figure 5. Chromosomal colinearity and phylogenetic relationships between genes from the Arabidopsis chromosome 4 genomic region spanning the *At4g34100* gene and their triplicated orthologs found in *Brassica rapa*.

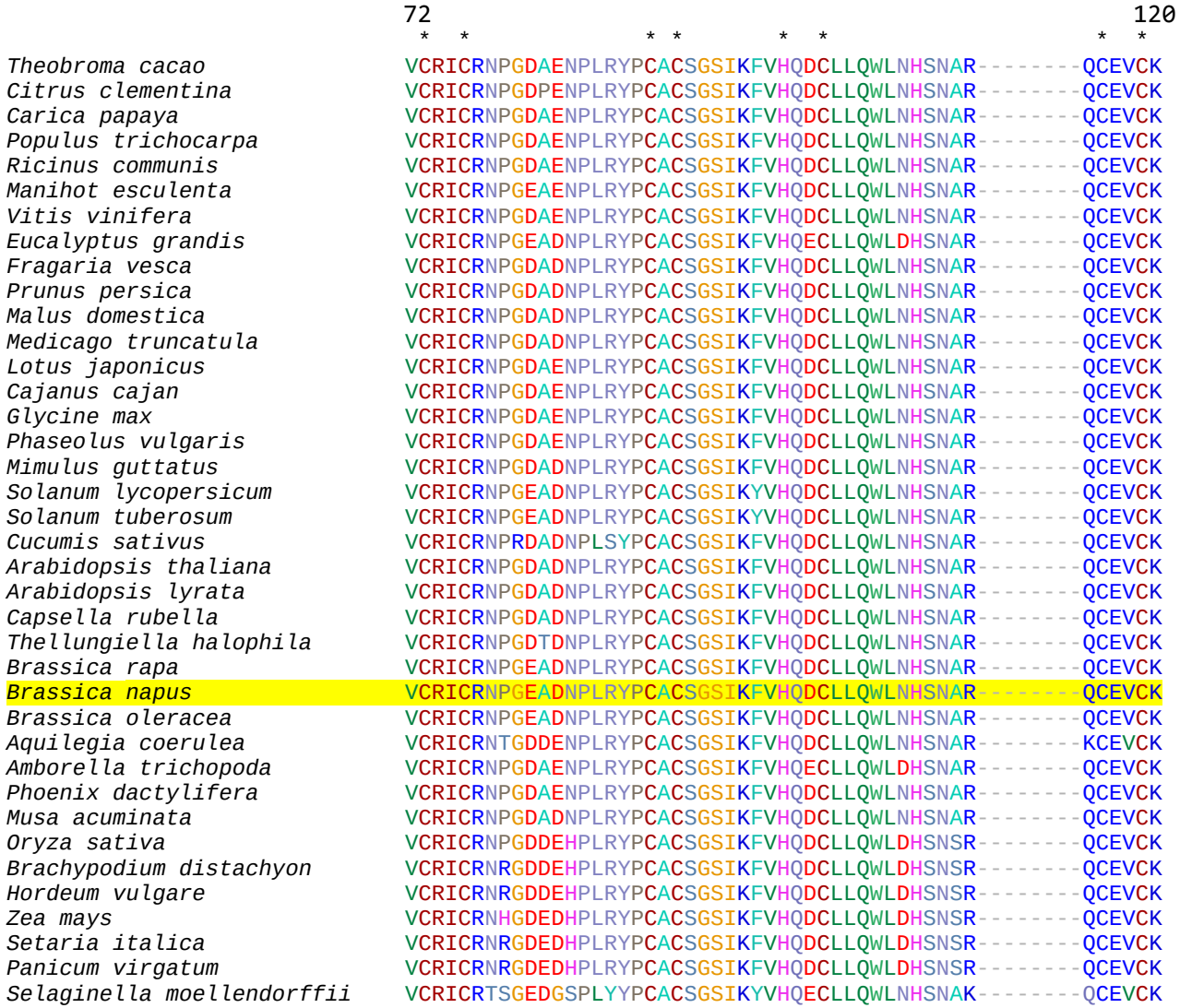
(A) Chromosomal colinearity. *B. rapa* subgenomes (BrA01, BrA03 and BrA08) (from Wang et al., 2011). In addition to *At4g34100* (red arrow), a set of five other genes are shown (indicated by blue Arrows). **(B)** Phylogeny relationships. Amino acid sequences were aligned using Clustal X (Thompson et al., 1997) . The phylogeny analysis was performed with MEGA version 4.1 using the Neighbor-Joining method (Tamura et al., 2007). The percent of replicates trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. . Blue line: *B. rapa* genomic region BrA01, red line: *B. rapa* genomic region BrA03, green line: *B. rapa* genomic region BrA08. The close relationship between Arabidopsis and one of the *Brassica* subgenome alternate (switch) between the three subgenomes, depending on the considered gene.



Supplemental Figure 6. Phylogenetic analysis of homologs of Bn-*CLG1A*.
Homologs of *Brassica napus* cleistogamy (Bn-*CLG1A)* gene from various eucaryotes showing the separation between plant (green) and non-plant Bn-*CLG1A* homologs were used.
A total of 107 entire protein sequences (see Supplemental Table 4 online) were considered. Amino acid sequences were aligned using Clustal X (Thompson et al., 1997) . The phylogeny analysis was performed with MEGA version 4.1 using the Neighbor-Joining method (Tamura et al., 2007).

Supplemental Figure 7. Alignment of RINGv domains of Bn-CLG1A homologous protein sequences identified in all eukaryote species.

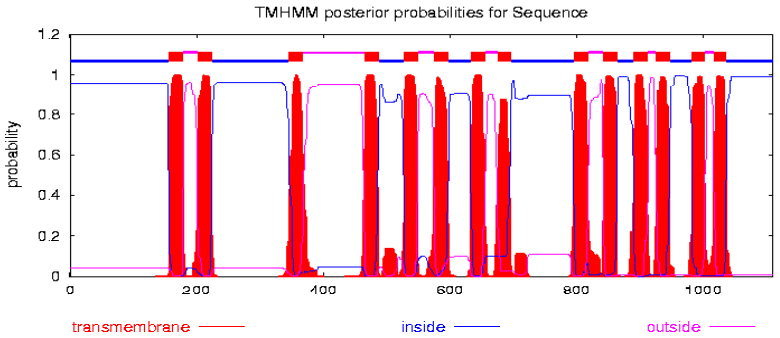
The position of eight conserved metal ligands is indicated by an asterisk. *B. napus* sequence is colored in yellow.



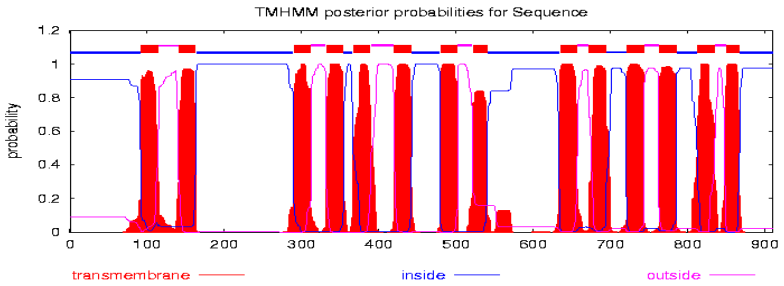
```
Physcomitrella patens          VCRICRTPGDEESSLYHPCACSGSIKYVHQECLLQWLNHSNAR--------QCEVCK
                                 *  *              * *       *  *                 *  *
Fungi
Ajellomyces capsulatus          TCRICRGEGTEEEQLYYPCKCSGSIKFVHQDCLMQWLSHSQKK--------YCELCK
Ajellomyces dermatitidis        TCRICRGEGTEEEQLFYPCKCSGSIKFVHQECLMQWLSHSQKK--------YCELCK
Aspergillus nidulans            TCRICRGEGTEQDELYYPCKCSGSIRFVHQPCLVQWLAHSQKK--------HCELCK
Aspergillus oryzae              TCRICRGEGTEEEQLFYPCKCSGSIKFVHQACLVEWLSHSQKK--------HCELCK
Aspergillus terreus             TCRICRGEGTPEEQLFYPCKCSGSIKFVHQACLVEWLSHSQKK--------HCELCK
Botryotinia fuckeliana          TCRICRAEATETEPLFYPCKCSGSIKFVHQDCLMEWLSHSQKK--------HCELCK
Coccidioides immitis            TCRICRGEGSEEEELFYPCKCSGSIKFVHQSCLMEWLSHSQKK--------YCELCK
Coprinopsis cinerea             TCRICSAPAEPDQPLFHPCKCSGTIRYIHQDCLTTWLAHSKKK--------SCDVCK
Gibberella zeae                 ICRICRGEGTPEEPLFYPCKCSGSIKYVHQDCLMEWLSHSQKK--------YCELCK
Laccaria bicolor                TCRICSAPAEPDQPLFHPCKCSGTIRYIHQDCLTTWLAHSKKK--------TCDVCK
Magnaporthe grisea              TCRICRGEATADEPLFYPCKCSGSIKYVHQDCLMEWLSHSQKK--------HCELFK
Microsporum canis               TCRICRGEGTEEEQLFYPCKCSGSIKFVHQNCLMEWLSHSQKK--------HCELCK
Nectria haematococca            ICRICRGEATPEEPLFYPCKCSGSIKYVHQDCLMEWLSHSQKK--------YCELCK
Neosartorya fischeri            TCRICRGEGTLEEPLFYPCKCSGSIKFVHQPCLVEWLSHSQKK--------HCELCK
Neurospora crassa               QCRICRGDASPDDPLYHPCKCSGSIKWVHQECLMQWLAQTQRK--------HCELCK
Paracoccidioides brasiliensis   TCRICRGEGTEEEQLFYPCKCSGSIKFVHQDCLMQWLSHSQKK--------YCELCK
Penicillium chrysogenum         TCRICHGEATEDEPLFYPCKCSGSIKFVHQVCLVEWLSHSQKK--------HCELCK
Penicillium marneffei           TCRICRGEGSEEEQLFYPCKCSGSIKFVHQSCLMEWLSHSQKK--------YCELCK
Phaeosphaeria nodorum           SCRICRGEGSVDEPLFYPCKCSGSIKYVHQECLMEWLSHTQKK--------HCELCK
Pichia guilliermondii           TCRICRGEATTSQPLLHPCKCRGSIRYIHQDCLLEWLKHSNKTTK------KCDICN
Podospora anserina              TCRICRGEGTPEEPLFYPCRCSGSIKHVHQDCLMEWLSHSQKK--------HCELCK
Pyrenophora tritici repentis    TCRICRGEGTTEEPLFFPCKCSGSIKYVHQECLMEWLSHTQKK--------HCELCK
Saccharomyces cerevisiae        TCRICRGEATEDNPLFHPCKCRGSIKYMHESCLLEWVASKNIDISKPGADVKCDICH
Schizosaccharomyces japonicus   FCRVCRCEGTPESPLYHPCKCSGSIRYVHQECLVEWLKHSRKK--------YCELCY
Schizosaccharomyces pombe       ICRVCRCEGAPDSPLFHPCKCTGSIRYVHQECLVEWLGHSKKT--------HCELCKA
Sclerotinia sclerotiorum        TCRICRAEATEREPLFYPCKCSGSIKFVHQDCLMEWLSHSQKK--------HCELCK
Sordaria macrospora             QCRICRGEATPDDPLYHPCKCSGSIKWVHQECLMQWLAQTQRK--------HCELCK
Talaromyces stipitatus          TCRICRGEGSEEEQLFYPCKCSGSIKFVHQSCLMEWLSHSQKK--------YCELCK
Trichophyton verrucosum         TCRICRGEGTDEEQLFYPCKCSGSIKFVHQNCLMEWLSHSQKK--------HCELCK
Tuber melanosporum              HCRICRSEGSREEPLFHPCKCSGSIKFVHQDCLLEWLQHSQKK--------HCELCK
Uncinocarpus reesii             TCRICRGEGSEEEELFYPCKCSGSIKFVHQSCLMEWLSHSQKK--------YCELCK
Animals
Acyrthosiphon pisum             ICRVCRSEGMPERPLFHPCICTGSIKYIHQECLVQWMRYSRKE--------YCELCS
Aedes egypti                    ICRVCRCEAQSDRPLFHPCICTGSIKWIHQDCLMQWMRYSRKE--------YCELCG
Anopheles gambiae               ICRVCRCEAQSDRPLFHPCICTGSIKWIHQDCLMQWMRYSRKE--------YCELCG
Bos taurus                      ICRVCRSEGTPEKPLYHPCVCTGSIKFIHQECLVQWLKHSRKE--------YCELCK
Caenorhabditis elegans          MCRVCRGN---EGSLYYPCLCTGSIKYVHQECLVEWLKYSKKE--------VCELCN
Canis familiaris                ICRVCRSEGTPEKPLYHPCVCTGSIKFIHQECLVQWLKHSRKE--------YCELCK
Ciona intestinalis              ICRVCRSTAFPDRPLFHPCICTGSIRHIHQDCLLQWLRHSRKE--------YCELCK
Danio rerio                     ICRVCRSEGTQDKPLYHPCVCTGSIKFIHQECLVQWLKHSRKE--------YCELCK
Drosophila melanogaster         ICRVCRCEAQPDRPLFYPCICTGSIKYIHQDCLMLWMRYSHKE--------YCELCS
Drosophila sechellia            ICRVCRCEAQPDRPLFYPCICTGSIKYIHQDCLMQWMRYSHKE--------YCELCS
Drosophila simulans             ICRVCRCEAQPDRPLFYPCICTGSIKYIHQDCLMQWMRYSHKE--------YCELCS
```

Equus caballus	ICRVCRSEGTPEKPLYHPCVCTGSIKFIHQECLVQWLKHSRKE--------YCELCK
Homo sapiens	ICRVCRSEGTPEKPLYHPCVCTGSIKFIHQECLVQWLKHSRKE--------YCELCK
Monodelphis domestica	ICRVCRSEGTPEKPLYHPCVCTGSIKFIHQECLVQWLKHSRKE--------YCELCK
Mus musculus	ICRVCRSEGTPEKPLYHPCVCTGSIKFIHQECLVQWLKHSRKE--------YCELCK
Nasonia vitripennis	ICRVCRSEGLPDRPLFHPCICTGSIKWIHQECLVQWMRYSRKE--------YCELCG
Nematostella vectensis	ICRVCRAEGTPDKPLYFPCICTGSIKYIHQECLLQWLKHSKKE--------YCELCN
Ornithorhynchus anatinus	ICRVCRSEGTPEKPLYHPCVCTGSIKFIHQECLVQWLKHSRKE--------YCELCK
Pan troglodytes	ICRVCRSEGTPEKPLYHPCVCTGSIKFIHQECLVQWLKHSRKE--------YCELCK
Pediculus humanus corporis	ICRVCRSEGLPDRPLFHPCICTGSIKWIHQECLMQWMRYSRKE--------YCELCS
pongo abelii	ICRVCRSEGTPEKPLYHPCVCTGSIKFIHQECLVQWLKHSRKE--------YCELCK
Rattus norvegicus	ICRVCRSEGTPEKPLYHPCVCTGSIKFIHQECLVQWLKHSRKE--------YCELCK
Saccoglossus kowalevskii	ICRVCRSEGAHDRPLFHPCICTGSIKFIHQDCLLQWLKHSKKE--------YCELCK
Sus scrofa	ICRVCRSEGTPEKPLYHPCVCTGSIKFIHQECLVQWLKHSRKE--------YCELCK
Taeniopygia guttata	ICRVCRSEGTPEKPLYHPCVCTGSIKFIHQECLVQWLKHSRKE--------YCELCK
Tetraodon nigroviridis	ICRVCRSEGTPDKPLYHPCVCTGSIKFIHQECLVQWLKHSRKE--------YCELCK
Tribolium castaneum	ICRVCRSEGLPDRPLFHPCICTGSIKWIHQECLMQWMRYSRKE--------YCELCS
Trichoplax adhaerens	ICRVCRLEGSTDKPLYFPCHCTGSIKYIHEACLLQWLKHSGKD--------FCELCR
Xenopus laevis	ICRVCRSEGTPEKPLYHPCVCTGSIKFIHQECLVLWLKHSRKE--------YCELCK
Protists	
Naegleria gruberi	LCRICKQPAADDDPLFHPCKCSGSIKYIHESCLNEWMKHSNKGK-------YCEICK
Phaeodactylum tricornutum	ECRVCRGPEEEGRPLFKPCKCSGSIGLTHQDCLQSWLEVQRGDG-------RCELCH
Phytophthora infestans	ECRVCRGEAEPDRRLFAPCKCSGSIRFTHSDCLEQWLEHSGKS--------FCELCG
Polysphondylium pallidum	VCRVCRNGPTTNNPLSYPCKCNGSIKFIHQNCLLDWIKFSKSS--------ACELCG

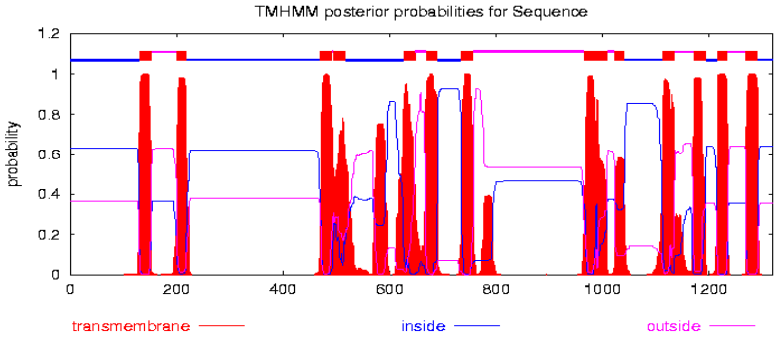
A- Bn-CLG1A



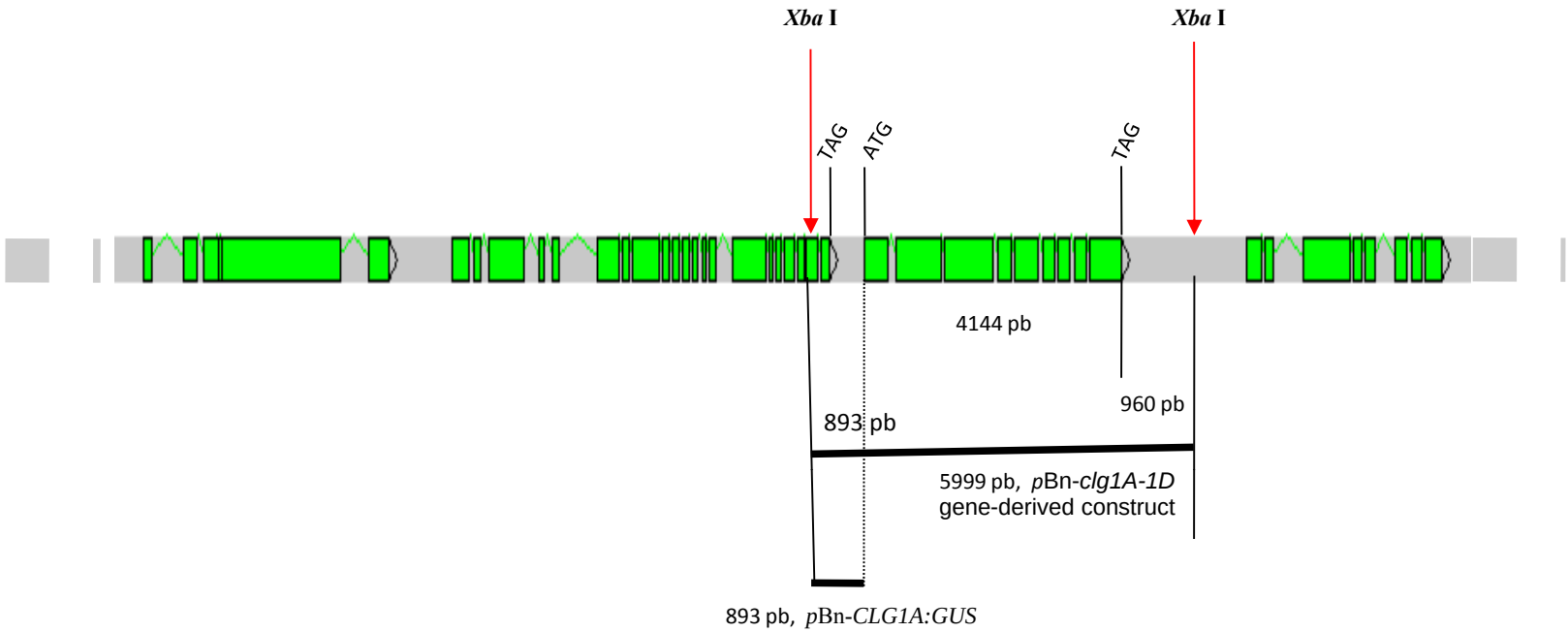
B- TEB4p



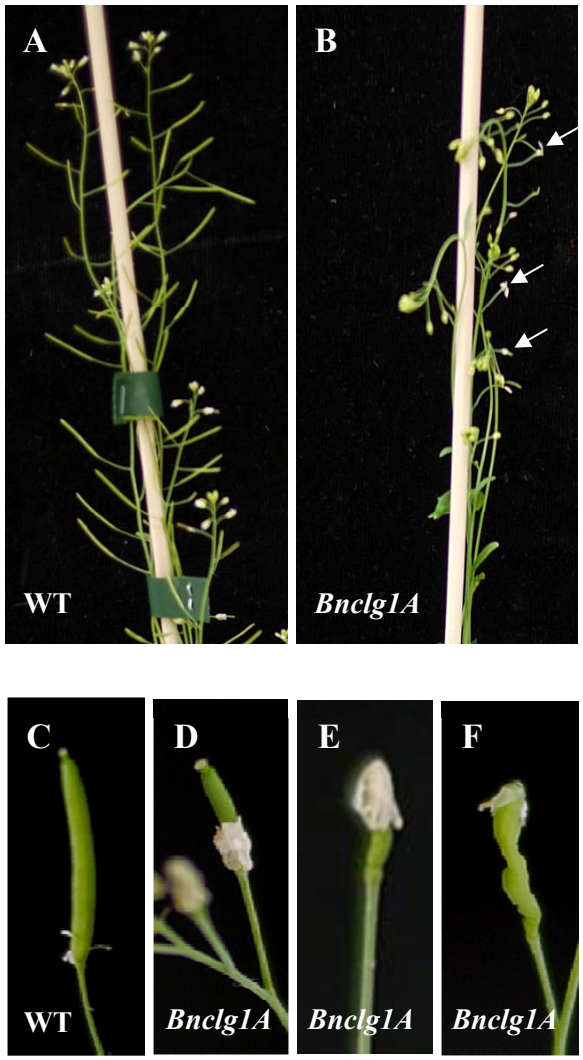
C- DOA10p



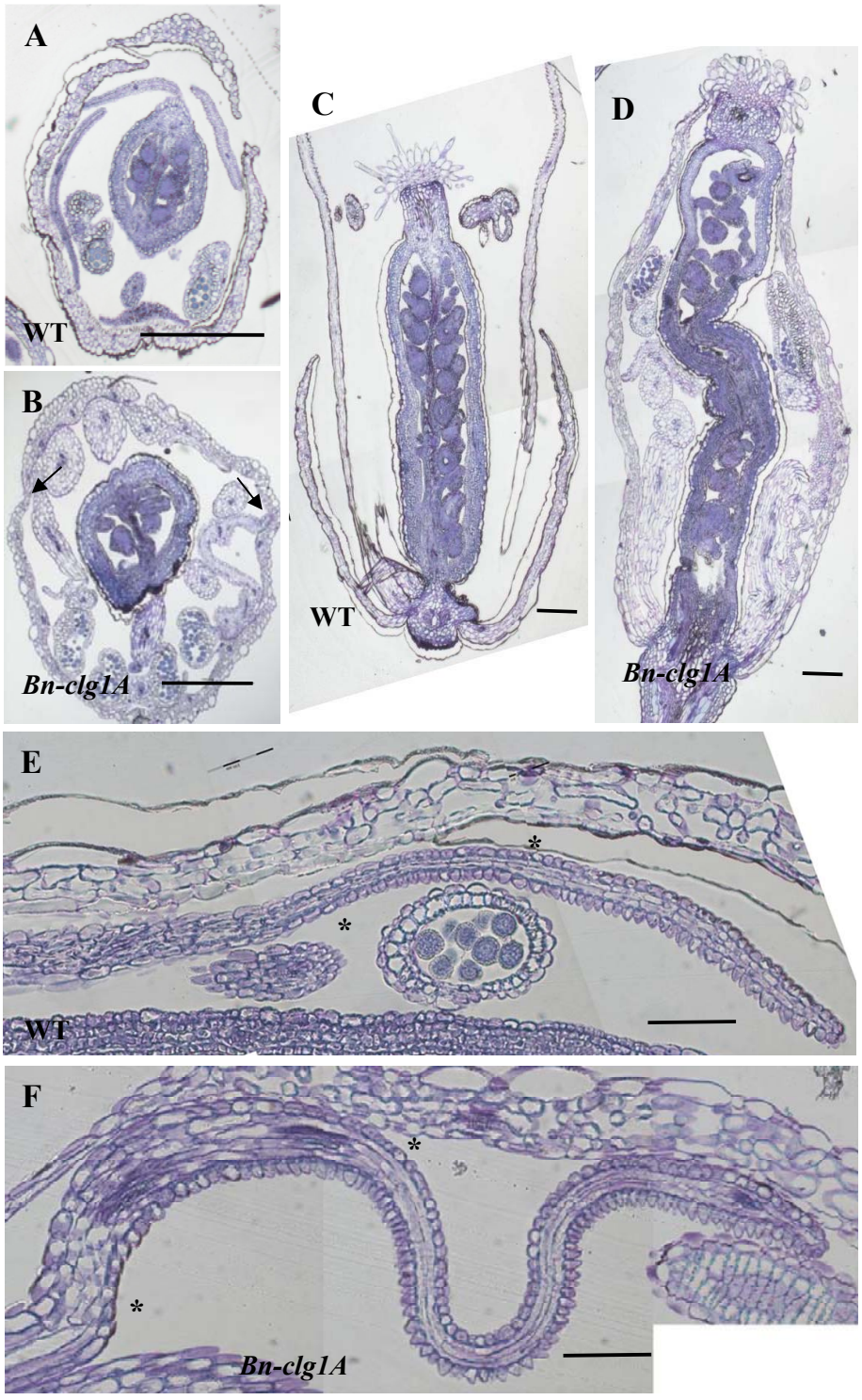
Supplemental Figure 8. Comparison of Membrane topology between Bn-CLG1A TEB4p and DOA10p proteins.
Membrane topology (predicted by TMHMM2.0, http://www.cbs.dtu.dk/services/TMHMM/ ; Krogh et al., 2001) is shown here.



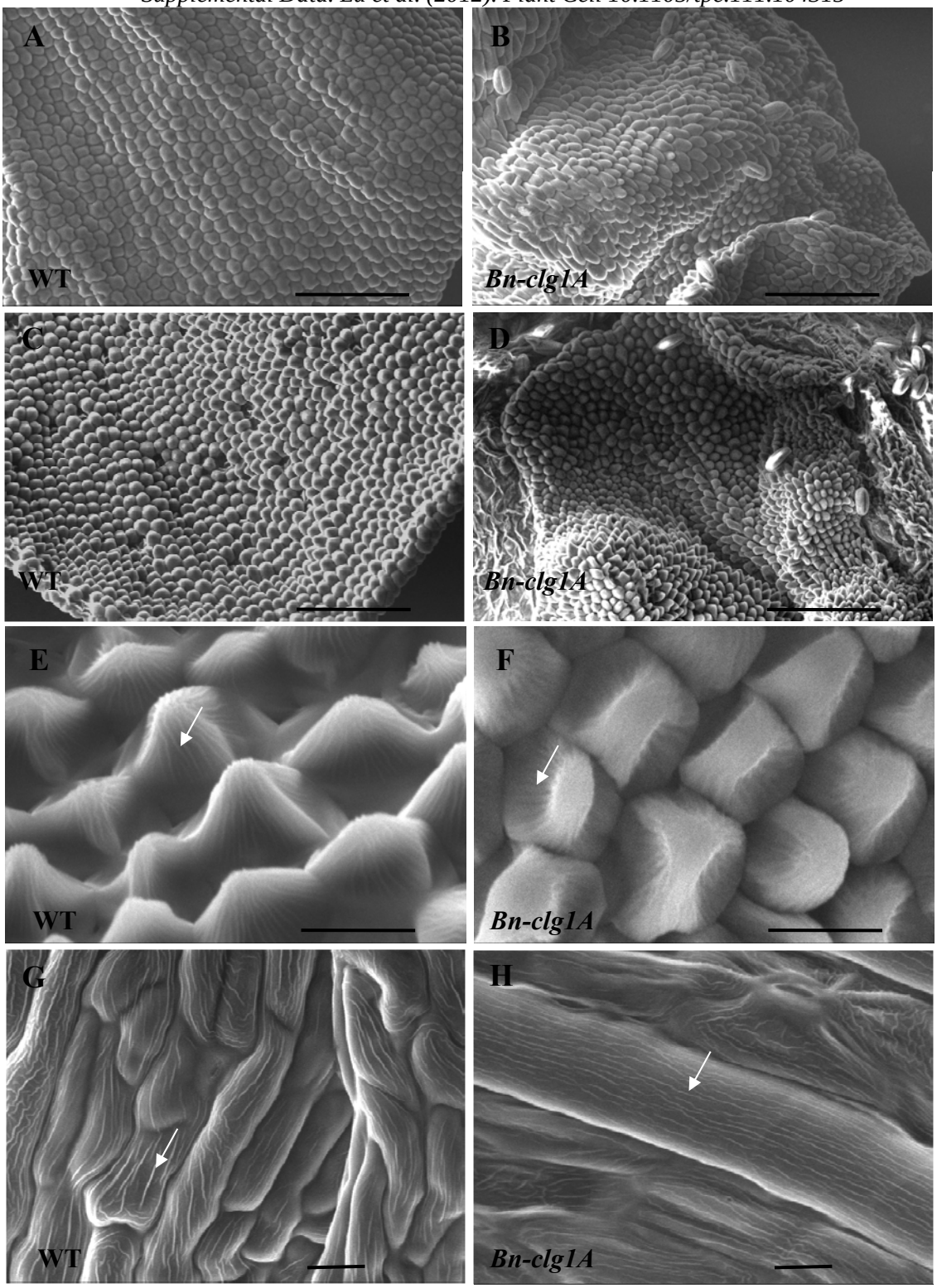
Supplemental Figure 9. Schematic presentation of genomic DNA promoter fragments used for the *p*Bn-*CLG1A:GUS* construct, as well as the entire gene sequence fragment used for Bn-*clg1A-1D* gene-derived construct (*p*Bn-*clg1-1D*).
The 893-bp DNA promoter fragment, used for the *p*Bn-*CLG1A:GUS* construct, was generated by PCR amplifications (forward primer 5-'GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTA GAGAGTTATGTAATGGGTTGG-3', reverse primer 5'-GGGG AC CAC TTT GTA CAA GAA AGC TGG GTT GGAGAAATCCGAGCTTGGGTTGAC-3'), and then inserted into a pBI101-GUS vector by the Gateway (Invitrogene) method.
For the the *p*Bn-*clg1-1D* gene-derived construct, an *Xba*I generated 5,999-bp genomic fragment of the mutant allele (895-bp native promoter region + 4144-pb predicted coding region + 960-bp downstream region of the *Bnclg1* gene) was inserted into the binary vector pBINPLUS.



Supplemental Figure 10. Reproductive organ morphology of Arabidopsis plants transformed with Bn-*clg1A-1D*.
(A) and (**B**) Flowering stem phenotype of 7 week-old Col-0 (WT) and transgenic Arabidopsis T1 Col-0 plants transformed with Bn-*clg1A-1D* : Arrows indicate representative semi-sterile siliques in the transgenic plant (line #1).
(**C**) to (**F**) Siliques of WT and Bn-*clg1A* -1D transgenic Arabidopsis T1 plants. Petals of *Bn-clg1A-1D* transgenic plants remain attached to the top of semi-sterile siliques that are also shorter or twisted.



Supplemental Figure 11. Light microscopy of Toluidine Blue stained sections of WT and Bn-*clg1A-1D* flowers.
(A) and **(B)** cross section of WT **(A)** and Bn-*clg1A-1D* **(B)** flowers at developmental stage 12. Arrows indicate sepal to sepal and sepal to petal fusions.
(C) and **(D)** assembled images of transverse section of WT **(C)** and Bn-*clg1A-1D* **(D)** flowers at developmental stage 14.
(E) and **(F)** assembled images of transverse section of WT **(E)** and Bn-*clg1A-1D* **(F)** petals at developmental stage 13. Asterisks indicate for both the abaxial and adaxial surface of petals the boundary between thin and elongated epidermal cell at the base of petals and the conical cells at the top (blade) of petals.



Supplemental Figure 12. Scanning electron microscopy of ArabidopsisBn-*clg1A-1D* transgenic plants.

(**A**) and (**B**) Abaxial petal blade of Col-0 WT (**A**) and Bn-*clg1A-1D* transgenic T1 flowers (**B**).
(**C**) and (**D**) Abaxial petal surface epidermis of Col-0 WT (**C**) and Bn-*clg1A-1D* flowers (**D**).
(**E**) and (**F**) Adaxial petal surface epidermis of Col-0 WT (**E**) and Bn-*clg1A-1D* flowers (**F**).
(**G**) and (**H**) Abaxial sepal surface epidermis of Col-0 WT (**G**) and Bn-*clg1A-1D* floral buds (**H**).
Note that Bn-*clg1A-1D* plant exhibit folded petals (**B**) and (**D**), and misshaped petal and sepal epidermal cells with absence or reduced nanoridges formation (**B**), (**D**), (**F**) and (**H**). Arrows indicate nanoridges. Scale bars: 100 µm (**A**) to (**D**), 10 µm (**E**) and (**F**), 20 µm (**G**) and (**H**).

Supplemental Table 1. Sequence of PCR primers and conditions used for genetic mapping of the Brassica napus cleistogamy gene (Bn-*clg1A-1D*) as well as for gene expression analysis using quantitative-PCR..

Primersna	Typestypes[a]	Primer sequence	condition
T2A01	AFLP	E-AGC / M-GTA	6% denaturing polyacrylamide gel
T2A02	AFLP	E-AGG / M-GAG	6% denaturing polyacrylamide gel
T2A04	AFLP	E-CTC / M-ACC	6% denaturing polyacrylamide gel
T2A10	AFLP	E-CGT / M-GAA	6% denaturing polyacrylamide gel
T2A16	AFLP	E-CAG / M-AAG	6% denaturing polyacrylamide gel
T2G101	ACGM (*Mse*I)[a]	Forward (5'->3'): GCT GCT ATT AAA TTC CGN GG Reverse (5'->3'): TCA AAG TTG GTC ACG GCR TC	6% denaturing polyacrylamide gel
T2G107	ACGM	Forward (5'->3'): ATC GTA AAG TGG ATA TGG GAA RCA Reverse (5'->3'): TCT CTC TGT GTT CGG ATG GAA AYT T	1.5% agarose gel
T2G108	ACGM	Forward (5'->3'): TAC CAT CAG ATC AGA TCC TCT NCA Reverse (5'->3'): GGG TGT GTA ACT GTG GAT CTA RCT	1.5% agarose gel
T2G115	ACGM	Forward (5'->3'): GGA GTT GAC TCA TGA CGA AAC NAC Reverse (5'->3'): GTC ATA TAA AGG CAA GAT TTC NGG	1.5% agarose gel
T2G116	ACGM	Forward (5'->3'): TGG TCC CAA TCT AAA TCA ACT NAC Reverse (5'->3'): ACC AGG AAA TGG TTC TTT GAA NGT	1.5% agarose gel
T2G119	ACGM (*Mse*I)[a]	Forward (5'->3'): TTC TAA CTC CGG TAC TAG GTG GGC Reverse (5'->3'): AGA TGT CTA TCA AGG AGT TCC NAA	1.5% agarose gel
T2G124	ACGM (*Mse*I)[a]	Forward (5'->3'): GAG ACG CCT CCA GCT CCT CGA YAC Reverse (5'->3'): TGA TCC AGT CCC AAA TCT ATA NCC	6% denaturing polyacrylamide gel
T2S004	microsatellite	Forward (5'->3'): GCG TCG CGA GAA ACA TTA TTC ACC Reverse (5'->3'): AAT TTG TGT TTT GTT CTT TTG CGT	3 % agarose gel
T2S005	STS	Forward (5'->3'): TTA CCC TTC AGA GAG TAT AAG C Reverse (5'->3'): CAA GCT GAG TAA ATT CAG GYT G	1.5% agarose gel
T2S008	STS	Forward (5'->3'): TTA CCC TTC AGA GAG TAT AAG C Reverse (5'->3'): CAA GCT GAG TAA ATT CAG GYT G	2% agarose gel
T2S009	microsatellite	Forward (5'->3'): GAG ACG CAG ATA CTT GAG ATA GGG Reverse (5'->3'): AAT AGT GGT TCG CTT TGG TTT ACT	3 % agarose gel
T2S043F/051R	CAPS (*Hae* III)[a]	Forward (5'->3'): CCA ATG GAT AAG GCG TCT GAC Reverse (5'->3'): GAG TAA CGA AAG ACA TAA GGG	3 % agarose gel
T2S064	CAPS (*Hae* III)[a]	Forward (5'->3'): AAC TCC AGA AGG ACT CGT CCG Reverse (5'->3'): GGG AGA AGA TAC CGA CGA TTC	3 % agarose gel
T2S027	STS	Forward (5'->3'): TGA GAA ACT CAA CGG GAT GCT GCT Reverse (5'->3'): CAT GAA GTA ATT GGG TAC AGG	1.5% agarose gel
dCAPS_Clg1	dCAPS (*Bgl* II)[a]	Forward (5'->3'): ACG CTG ATG GTG CTG AAG AAG ATC Reverse (5'->3'): GTA TTT TCC AAT GCT GAG TCA	3 % agarose gel
			-
BnEF1-F	Q-PCR	Forward (5'->3'):GAAGCAGGTATCTCCAAGGATG	-
BnEF1-R	Q-PCR	Reverse (5'->3') GGTAGTGGCATCCATCTTGTTAC	-
BnUBQ2-F	Q-PCR	Forward (5'->3'):CAAGATCCAGGACAAAGAAGGGA	-
BnUBQ2-R	Q-PCR	Reverse (5'->3') CTAAGCCTCAGGACAAGATGCA	-
BnClg1A-F	Q-PCR	Forward (5'->3'):AGAGAGATTGTATTCCCGATTGTA	-

[a] Restriction enzyme used to cut the amplified fragments is shown between brackets

Supplemental Table 2. Comparison of gene content between the *Brassica napus* BAC clone CZ7N2 and the corresponding orthologous region of Arabidopsis chromosome 4.

Brassica napus		Arabidopsis
Coordinate on CZ7N2	Predicted Gene	Gene Model
471 - 43	Unnamed1	-
-	AT	4G33585
-	AT	4G33590
-	AT	4G33600
-	AT	4G33610
1209 - 6526	GSBNAPT00096217001	AT 4G33620
8007 - 6778	GSBNAPT00096218001	AT 4G33625
-	AT	4G33630
9259 - 8768	GSBNAPT00096219001	AT 4G33640
10540 - 15579	GSBNAPT00096220001	AT4G33650
17019 - 16607	GSBNAPT00096221001	AT4G33660
17971 - 18222	GSBNAPT00096222001	AT4G33666
-	AT	4G33670
21476 - 18478	GSBNAPT00096223001	AT4G33680
21995 - 22982	GSBNAPT00096224001	AT4G33690
25892 - 23335	GSBNAPT00096225001	AT4G33700
28113 - 27193	Unnamed2	-
-	AT	4G33710
29661 - 30185	GSBNAPT00096226001	AT4G33720
32141 - 32659	GSBNAPT00096227001_7N2	AT4G33730
34250 - 35563	GSBNAPT00096227010_7N2	AT4G33740
41112 - 37717	GSBNAPT00096228001	AT4G33760
44217 - 41764	GSBNAPT00096228010_7N2	AT4G33770
44887 - 45121	Unnamed3	-
46870 - 47270	Unnamed4	-
48881 - 48199	GSBNAPT00096228020_7N2	-
52206 - 51533	GSBNAPT00096229001	AT4G33780
61132 - 53698	GSBNAPT00096230001	AT4G33790
55728 - 57590	Unnamed6	-
65143 - 64178	GSBNAPT00096231001	AT4G33800
67724 - 67450	GSBNAPT00096231010_7N2	pseudo AT4G33810
69553 - 68062	GSBNAPT00096232001	AT4G33820
72486 - 71681	GSBNAPT00096232010_7N2	-
72513 - 73739	GSBNAPT00096232020_7N2	-
-	AT	4G33830
-	AT	4G33840
-	AT	4G33850
-	AT	4G33860
76179 - 74592	GSBNAPT00096233001	AT4G33880
85639 - 86223	GSBNAPT00096234001	AT4G33865
-	AT	4G33890
-	AT	4G33900
-	AT	4G33905
-	AT	4G33910
-	AT	4G33920
86795 - 87680	GSBNAPT00096235001	AT4G33925
-	AT	4G33930
88514 - 89652	GSBNAPT00096236001	AT4G33940
90445 - 92795	GSBNAPT00096237001	AT4G33945
94433 - 96527	GSBNAPT00096238001	AT4G33950
101268 - 101465	GSBNAPT00096239001	AT4G33960
101868 - 101968	Unnamed7	-
102343 - 102696	Unnamed8	-
109210 - 107210	GSBNAPT00096240001	AT4G33970
112079 - 110013	GSBNAPT00096241001_7N2	AT4G33980
-	AT	4G33985
113636 - 113463	Unnamed9	AT 4G33990
116691 - 118613	GSBNAPT00096241010_7N2	AT4G34000
121063 - 119119	GSBNAPT00096241020_7N2	AT4G34020
-	AT	4G34030
-	AT	4G34040
121071 - 124368	Unnamed10	AT 4G34050
126162 - 132217	GSBNAPT00096242001	AT4G34060
-	AT	4G34070

132579 - 132367	Unnamed11	AT	4G34090
-	AT		4G34080
132808 - 136951	GSBNAPT00096243001		AT4G34100
139001 - 142121	GSBNAPT00096244001		-
-	AT		4G34110
142858 - 144436	GSBNAPT00096245001		AT4G34120
146464 - 144744	GSBNAPT00096246001		AT4G34131
148604 - 148434	Unnamed12	-	
151199 - 149739	GSBNAPT00096247001		AT4G34135
153627 - 152006	GSBNAPT00096248001		AT4G34138
154449 - 156772	GSBNAPT00096249001		AT4G34140
157491 - 159278	GSBNAPT00096250001		AT4G34150
160162 - 161530	GSBNAPT00096251001		AT4G34160
171177 - 169664	GSBNAPT00096252001		-
172571 - 175799	GSBNAPT00096253001		-
-	AT		4G34170
-	AT		4G34180
177128 - 176361	GSBNAPT00096254001		AT4G34190
180485 - 177923	GSBNAPT00096255001		AT4G34200
-	AT		4G34210
183983 - 183119	GSBNAPT00096256001		AT4G34215
187166 - 184783	GSBNAPT00096257001		AT4G34220
192886 - 190111	GSBNAPT00096258001		-
196091 - 194281	GSBNAPT00096259001		AT4G34230
198448 - 201320	GSBNAPT00096260001		AT4G34240
204721 - 206151	GSBNAPT00096261001		AT4G34250
208973 - 212977	GSBNAPT00096262001		AT4G34260
215463 - 216183	GSBNAPT00096263001		AT4G34265
218838 - 217130	GSBNAPT00096264001		AT4G34270
220566 - 219504	GSBNAPT00096265001		-
-	AT		4G34272
220993 - 224926	GSBNAPT00096266001		AT4G34280
225387 - 226213	GSBNAPT00096267001		AT4G34290
-	AT		4G34300
232217 - 237044	GSBNAPT00096268001		AT4G34310
-	AT		4G34320
-	AT		4G34330
237490 - 237660	Unnamed14	AT	4G34340

-: No orthologous gene was found at this position

Supplemental Table 3. List of 19 Bn-*clg1A-1D* / Bn-*CLG1A* near-isogenic lines and 19 representative oilseed rape cultivars analyzed by the dCAPS_Clg1 marker.

Bn-*clg1A-1D* / Bn-*CLG1A* near-isogenic lines :

AKAMAR/AKAMAR-*Clg*, ASCONA/ASCONA-*Clg*, BRISTOL/BRISTOL-*Clg*, CAPITOL/CAPITOL-*Clg*, CCH 1083/CCH 1083-*Clg*, EXPRESS/EXPRESS-*Clg*, FALCON/FALCON-*Clg*, GIN 3683/GIN 3683-*Clg*, H133 (zeruca)/H133 (zeruca)-*Clg*, H5 7741/H5 7741-*Clg*, HR 51/HR 51-*Clg*, ISN 1770/ISN 1770-*Clg*, JAZZ/JAZZ-*Clg*, MOHICAN/MOHICAN-*Clg*, NAVAJO/NAVAJO-*Clg*, NPZ 01 (jetton)/NPZ 01 (jetton)-*Clg*, NPZ 12 (zorro)/NPZ 12 (zorro)-*Clg*, VIVOL/VIVOL-*Clg*, YUDAL/YUDAL-*Clg*

Representative oilseed rape cultivars:

B001_*Clg*, Darmor, Darmor_Nain, Samouraï, Stella, Norin 9, Drakkar, Cheyenne, NSL, Rapid, Euro A, Euro B, Euro C, Euro D, Euro E, Euro F, Z821, Primor, Primor-*Clg*

Supplemental Table 4. List of Bn-*clg1A-1D* orthologs used for phylogenic tree construction in Figure 5 and Supplemental Figure 6 online..

		species names	sequence reference	database*
Protists				
	1	*Naegleria gruberi*	XP_002670283.1	NCBI
	2	*Phaeodactylum tricornutum*	XP_002176733.1	NCBI
	3	*Phytophthora infestans*	EEY59020.1	NCBI
	4	*Polysphondylium pallidum*	EFA82525.1	NCBI
Fungi				
	1	*Ajellomyces capsulatus*	EEH07677.1	NCBI
	2	*Ajellomyces dermatitidis*	EEQ83336.1	NCBI
	3	*Aspergillus nidulans*	XP_663740.1	NCBI
	4	*Aspergillus oryzae*	XP_001826413.1	NCBI
	5	*Aspergillus terreus*	XP_001218322.1	NCBI
	6	*Botryotinia fuckeliana*	XP_001549835.1	NCBI
	7	*Coccidioides immitis*	XP_001247058.1	NCBI
	8	*Coprinopsis cinerea*	XP_001831723.1	NCBI
	9	*Gibberella zeae*	XP_386758.1	NCBI
	10	*Laccaria bicolor*	XP_001883668.1	NCBI
	11	*Magnaporthe grisea*	XP_364239.2	NCBI
	12	*Microsporum canis*	EEQ31659.1	NCBI
	13	*Nectria haematococca*	EEU39888.1	NCBI
	14	*Neosartorya fischeri*	XP_001260371.1	NCBI
	15	*Neurospora crassa*	XP_958364.1	NCBI
	16	*Paracoccidioides brasiliensis*	EEH49250.1	NCBI
	17	*Penicillium chrysogenum*	XP_002558835.1	NCBI
	18	*Penicillium marneffei*	XP_002146559.1	NCBI
	19	*Phaeosphaeria nodorum*	XP_001799855.1	NCBI
	20	*Pichia guilliermondii*	XP_001484747.1	NCBI
	21	*Podospora anserina*	XP_001907566.1	NCBI
	22	*Pyrenophora tritici repentis*	XP_001930705.1	NCBI
	23	*Saccharomyces cerevisiae*	NP_012234.1	NCBI
	24	*Schizosaccharomyces japonicus*	XP_002174794.1	NCBI
	25	*Schizosaccharomyces pombe*	NP_596733.1	NCBI
	26	*Sclerotinia sclerotiorum*	XP_001592095.1	NCBI
	27	*Sordaria macrospora*	CBI55668.1	NCBI
	28	*Talaromyces stipitatus*	XP_002478858.1	NCBI
	29	*Trichophyton verrucosum*	EFE38201.1	NCBI
	30	*Tuber melanosporum*	CAZ85005.1	NCBI
	31	*Uncinocarpus reesii*	XP_002541277.1	NCBI
Plants				
	1	*Amborella trichopoda*	evm_15.model.AmTr_v1.0_scaffold 00035.33:5558183-5591397(-)	AGD
	2	*Aquilegia coerulea*	Aquca_015_00003	
	3	*Arabidopsis lyrata*	XP_002869176.1	NCBI
	4	*Arabidopsis thaliana*	NP_195136.3	NCBI
	5	*Brachypodium distachyon*	XP_003563290.1	NCBI
	6	*Brassica napus*	BnClg1.1An	INRA
	7	*Brassica oleracea*	C01#2011-08-02#BGI	BRAD
	8	*Brassica rapa*	Bra011487	BRAD
	9	*Cajanus cajan*	C.cajan_31076	PGDD
	10	*Capsella rubella*	Carubv10004028m	Phytozome
	11	*Carica papaya*	evm.model.supercontig_200.25	Phytozome
	12	*Citrus clementina*	clementine0.9_000892m	Phytozome
	13	*Cucumis sativus*	Cucsa.219750.1	Phytozome
	14	*Eucalyptus grandis*	Eucgr.I01130.	Phytozome
	15	*Fragaria vesca*	scf0512966	GDR
	16	*Glycine max*	XP_003518705.1	NCBI
	17	*Hordeum vulgare*	BAJ93994.1	NCBI
	18	*Lotus japonicus*	chr4.CM0227.540.nc	PlantGDB
	19	*Malus domestica*	MDP0000939108	Phytozome
	20	*Manihot esculenta*	cassava4.1_000587m	Phytozome
	21	*Medicago truncatula*	XP_003607923.1	NCBI
	22	*Mimulus guttatus*	mgv1a000808m	Phytozome
	23	*Musa acuminata*	fusion_GSMUA_Achr1T26090_001 _GSMUA_Achr1T26100_001	Genoscope
	24	*Oryza sativa*	EEE66081.1	NCBI

25	*Panicum virgatum*	Pavirv00041158m	Phytozome
26	*Phaseolus vulgaris*	Phvulv091002017m	Phytozome
27	*Phoenix dactylifera*	PDK_30s749921g004	WCMC
28	*Physcomitrella patens*	XP_001775262.1	NCBI
29	*Populus trichocarpa*	XP_002313128.1	NCBI
30	*Prunus persica*	ppa000536m	Phytozome
31	*Ricinus communis*	XP_002518729.1	NCBI
32	*Selaginella moellendorffii*	XP_002966406.1	NCBI
33	*Setaria italica*	Si005707m	Phytozome
34	*Solanum lycopersicum*	Solyc01g107880.2	PGDD
35	*Solanum tuberosum*	PGSC0003DMB000000073	SGN
36	*Thellungiella halophila*	Thhalv10024280m	Phytozome
37	*Theobroma cacao*	Tc02_g009070	PGDD
38	*Vitis vinifera*	XP_002282447.1	NCBI
39	*Zea mays*	GRMZM2G051792_T01	PlantGDB
Animals			
1	*Acyrthosiphon pisum*	XP_001942926.1	NCBI
2	*Aedes egypti*	XP_001659978.1	NCBI
3	*Ailuropoda melanoleuca*	EFB15044.1	NCBI
4	*Anopheles gambiae*	XP_308581.3	NCBI
5	*Bos taurus*	XP_593602.3	NCBI
6	*Branchiostoma floridae*	XP_002604587.1	NCBI
7	*Caenorhabditis elegans*	NP_492823.2	NCBI
8	*Canis familiaris*	XP_535791.2	NCBI
9	*Ciona intestinalis*	XP_002119395.1	NCBI
10	*Danio rerio*	XP_002660875.1	NCBI
11	*Drosophila_melanogaster*	NP_647715.2	NCBI
12	*Drosophila sechellia*	XP_002035057.1	NCBI
13	*Drosophila simulans*	XP_002083352.1	NCBI
14	*Equus caballus*	XP_001500611.2	NCBI
15	*Gallus gallus*	XP_419012.2	NCBI
16	*Homo sapiens*	NP_005876.2	NCBI
17	*Ixodes scapularis*	XP_002401333.1	NCBI
18	*Monodelphis domestica*	XP_001372049.1	NCBI
19	*Mus musculus*	NP_766194.2	NCBI
20	*Nasonia vitripennis*	XP_001605983.1	NCBI
21	*Nematostella vectensis*	XP_001639316.1	NCBI
22	*Ornithorhynchus anatinus*	XP_001519170.1	NCBI
23	*Pan troglodytes*	XP_517630.2	NCBI
24	*Pediculus humanus corporis*	XP_002431316.1	NCBI
25	*Pongo abelii*	Q5R9W1.1	NCBI
26	*Rattus norvegicus*	XP_001065952.2	NCBI
27	*Saccoglossus kowalevskii*	XP_002737454.1	NCBI
28	*Sus scrofa*	NP_001124007.1	NCBI
28	*Taeniopygia guttata*	XP_002189350.1	NCBI
30	*Tetraodon nigroviridis*	CAF97998.1	NCBI
31	*Tribolium castaneum*	XP_966509.1	NCBI
32	*Trichoplax adhaerens*	XP_002107876.1	NCBI
33	*Xenopus laevis*	NP_001091317.1	NCBI

* *AGD:* http://www.amborella.org/;
BRAD: http://brassicadb.org/brad/
GDR: http://www.rosaceae.org/
LIS: http://cajca.comparative-legumes.org/
NCBI: http://www.ncbi.nlm.nih.gov/;
Phytozome: http://www.phytozome.com/
PlantGDB: http://www.plantgdb.org/;
SGN: http://solgenomics.net/
WCMC: http://qatar-weill.cornell.edu/research/datepalmGenome/
PGDD: http://chibba.pgml.uga.edu/duplication/index/home