

# One-step creation of CMS lines using a *BoCENH3*-based haploid induction system in *Brassica* crop

Received: 13 September 2023

Accepted: 4 February 2024

Published online: 18 March 2024

 Check for updates

Fengqing Han<sup>1,5</sup>, Xiaoli Zhang<sup>2,5</sup>, Yuxiang Liu<sup>1,3,5</sup>, Yumei Liu<sup>1</sup>, Hong Zhao<sup>4</sup> & Zhansheng Li<sup>1</sup>✉

Heterosis utilization in a large proportion of crops depends on the use of cytoplasmic male sterility (CMS) tools, requiring the development of homozygous fertile lines and CMS lines<sup>1</sup>. Although doubled haploid (DH) technology has been developed for several crops to rapidly generate fertile lines<sup>2,3</sup>, CMS lines are generally created by multiple rounds of backcrossing, which is time consuming and expensive<sup>4</sup>. Here we describe a method for generating both homozygous fertile and CMS lines through in vivo paternal haploid induction (HI). We generated in-frame deletion and restored frameshift mutants of *BoCENH3* in *Brassica oleracea* using the CRISPR/Cas9 system. The mutants induced paternal haploids by outcrossing. We subsequently generated HI lines with CMS cytoplasm, which enabled the generation of homozygous CMS lines in one step. The *BoCENH3*-based HI system provides a new DH technology to accelerate breeding in *Brassica* and other crops.

Heterosis utilization has been extensively applied in plant breeding and contributes greatly to crop yield<sup>5</sup>. One of the most important processes for heterosis utilization is to develop homozygous (true-breeding) lines, which requires 6–8 generations of selfing<sup>2,6</sup>. In addition, heterosis utilization in many crops including most *Brassica* crops, as well as millet, carrot and some rice, maize, soybean and sorghum crops, exploits cytoplasmic male sterility (CMS) as a pollination control tool for hybrid seed production, requiring another 6–8 rounds of backcrossing to generate homozygous CMS lines<sup>4</sup>. These selfing and backcrossing processes are time consuming and costly for most crops<sup>7</sup>.

Doubled haploid (DH) technology has broad application in generating true-breeding lines within 1–2 generations<sup>8</sup>. In vitro haploid induction (HI) methods such as anther culture and microspore embryogenesis are expensive and limited by genotype recalcitrance<sup>9</sup>. In vivo

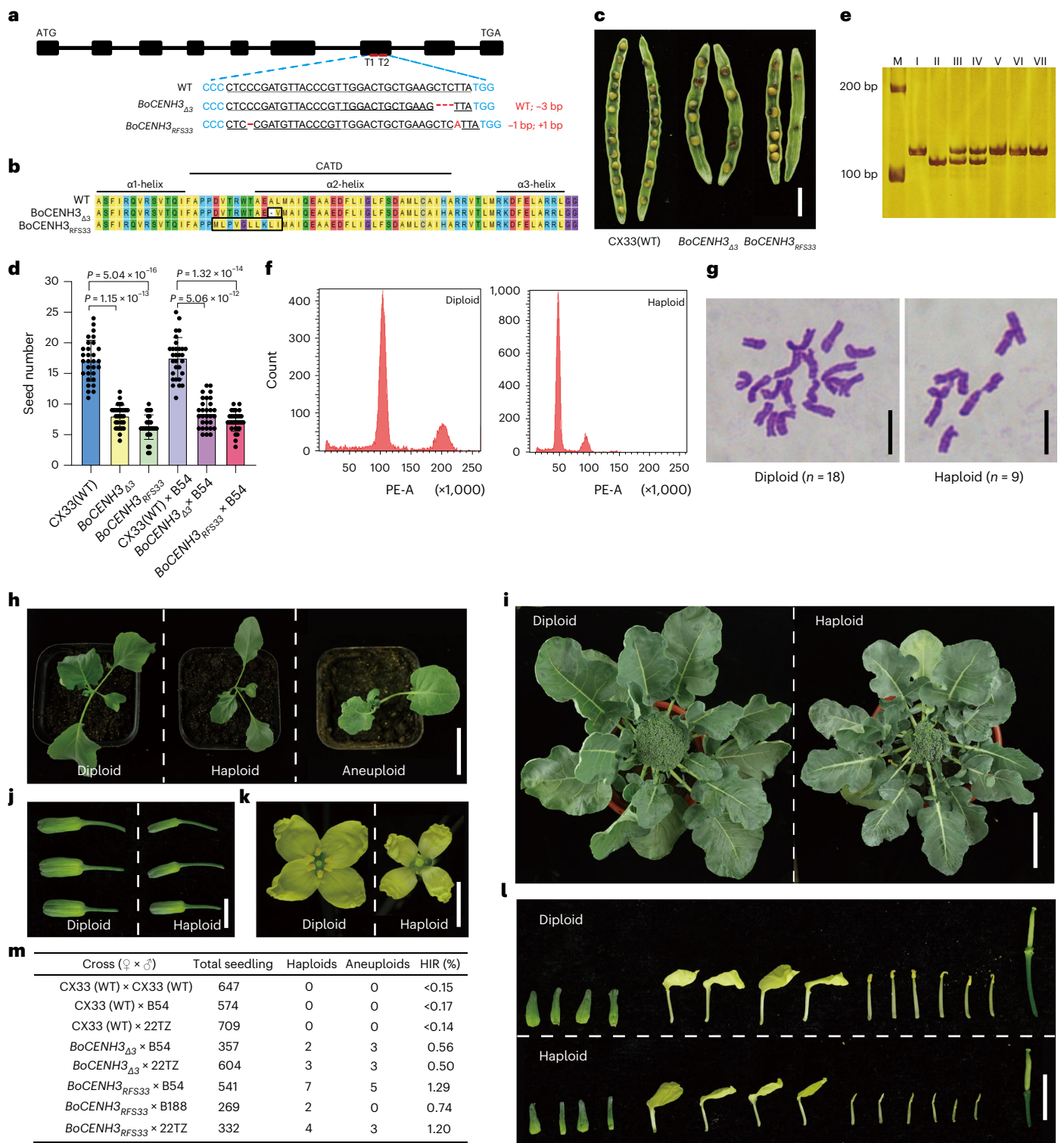
HI technology has commonly been used in maize breeding for several decades<sup>10</sup>. In recent years, due to the cloning of genes including *MTL/NLD/ZmPLA1*, *DMP*, *ZmPOD65*, *CENH3* and *ECS1/2* that are responsible for HI<sup>3,11–13</sup>, in vivo HI has been extended to several plant species including *Arabidopsis*, rice, wheat, tomato, rapeseed, cabbage and *Medicago truncatula*<sup>2,7–9,14–16</sup>.

The *Arabidopsis* *CENTROMERIC HISTONE3* (*CENH3*) gene encodes a centromere-identifying protein histone H3 variant<sup>17</sup>. Modification of this gene has been shown to induce mainly paternal haploid progeny upon outcrossing to the wild type (WT)<sup>12,18</sup>. However, the establishment of a *CENH3*-based system is relatively difficult because it requires modifications such as the use of GFP–tailswap chimaeric proteins, the expression of non-native *CENH3*, in-frame deletions, single amino acid substitutions and heterozygous *cenh3* null mutation (+/*cenh3*), while

<sup>1</sup>State Key Laboratory of Vegetable Biobreeding, Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing, China.

<sup>2</sup>State Key Laboratory of Vegetable Biobreeding, Tianjin Academy of Agricultural Sciences, Tianjin, China. <sup>3</sup>Key Laboratory for Vegetable Biology of Hunan Province, Engineering Research Center for Horticultural Crop Germplasm Creation and New Variety Breeding, Ministry of Education, Hunan Agricultural University, Changsha, China. <sup>4</sup>Beijing Vegetable Research Center (BVRC), Beijing Academy of Agricultural and Forestry Sciences, National Engineering Research Center for Vegetables, Beijing, China. <sup>5</sup>These authors contributed equally: Fengqing Han, Xiaoli Zhang, Yuxiang Liu.

✉e-mail: [lizhansheng@caas.cn](mailto:lizhansheng@caas.cn)



**Fig. 1 | In-frame deletion and restored frameshift mutations of *BoCENH3* induce paternal haploids. a**, Schematic diagram of *BoCENH3* gene structure and CRISPR/Cas9-based genome editing. Black blocks, gene coding region; red lines, sgRNA-targeted regions (T1, T2); underlining indicates the target sequence; protospacer adjacent motif sequences are highlighted in blue; insertions/deletions are highlighted in red. **b**, Amino acid alignment of BoCENH3, BoCENH3<sub>Δ3</sub> and BoCENH3<sub>RFS33</sub>. Putative CATD and α-helix are indicated by black lines. Amino acid changes in BoCENH3<sub>Δ3</sub> and BoCENH3<sub>RFS33</sub> are outlined in black boxes. **c**, Representative siliques from selfed CX33 (WT) and BoCENH3 mutants. **d**, Per silique seed setting performance of selfed WT, BoCENH3<sub>Δ3</sub> and BoCENH3<sub>RFS33</sub>.

Means ± s.d. (n = 30) (unpaired, two-tailed *t*-test). **e**, Haploid genotyping with DNA marker. M, DNA size marker; I–II, PCR bands of the B54 inbred line and the BoCENH3<sub>Δ3</sub> mutant; III–IV, F<sub>1</sub> hybrids from BoCENH3<sub>Δ3</sub> × B54; V–VII, three haploids from B54. **f**, Flow cytometry analysis of diploid B54 and haploid B54 generated by BoCENH3<sub>Δ3</sub> inducer. **g**, Chromosome numbers in diploid B54 (n = 18) and haploid B54 (n = 9) plants. **h**, Phenotypes of diploid, haploid and aneuploid B54 seedlings. **i–l**, Phenotypes of diploid and haploid B54 adult plants (**i**), buds (**j**), flowers (**k**) and dissected flower tissues (**l**). **m**, Haploid induction rates of BoCENH3<sub>Δ3</sub> and BoCENH3<sub>RFS33</sub>. Scale bars, 1 cm (**c**), 10 μm (**g**), 5 cm (**h**), 10 cm (**i**), 0.5 cm (**j, k**) and 1 cm (**l**).

homozygous *cenh3* null mutation is lethal<sup>8,19–22</sup>. At present, although *CENH3* is conserved in both monocotyledonous and dicotyledonous plants, *CENH3*-based paternal HI is available only for *Arabidopsis*, maize and wheat<sup>8,21,22</sup>.

The paternal haploid inducer line has the advantage of transferring cytoplasm between different genotypes, as the cytoplasm is maternally inherited during crossing. The maize *indeterminate gametophyte1 (ig1)* mutation induces the formation of androgenetic haploids with an HI rate of 1–3% (refs. 23,24). This *ig1* haploid inducer is routinely used by breeders to transfer CMS cytoplasm, such as C, S, SD, Vg, ME and CA, to inbred lines<sup>23</sup>. The *CENH3*-based haploid inducer also enables the transfer of cytoplasm among different ecotypes in *Arabidopsis*<sup>8,25</sup>, but it has not been employed to produce CMS lines.

We tested whether modification of the *CENH3* homologue can be used for paternal HI and whether a *CENH3*-based paternal inducer line with a CMS cytoplasm enables the creation of CMS lines through the exchange of cytoplasmic and nuclear genomes in broccoli (*Brassica oleracea* var. *italica*), a globally important vegetable crop.

Using the *Arabidopsis* *CENH3* as query, we searched the broccoli HDEM genome with the Basic Local Alignment Search Tool for protein (BLASTP) and found only one *CENH3* homologue BoC8t52879H (BoCENH3). Phylogenetic analysis revealed that BoCENH3 is highly similar to *Arabidopsis* *CENH3* with 66% sequence identity, and they are assigned to the same subclade (Supplementary Fig. 1 and Supplementary Table 1). We cloned *BoCENH3* from the broccoli inbred line CX33 and detected its expression in various broccoli tissues by reverse transcription quantitative PCR (RT-qPCR), which indicated that *BoCENH3* is highly expressed in pistils and young buds, especially in pistils at 24–72 h after pollination (Supplementary Fig. 2), consistent with what has been found for authentic *CENH3* genes in *Arabidopsis*, barley and wheat<sup>19,21</sup>.

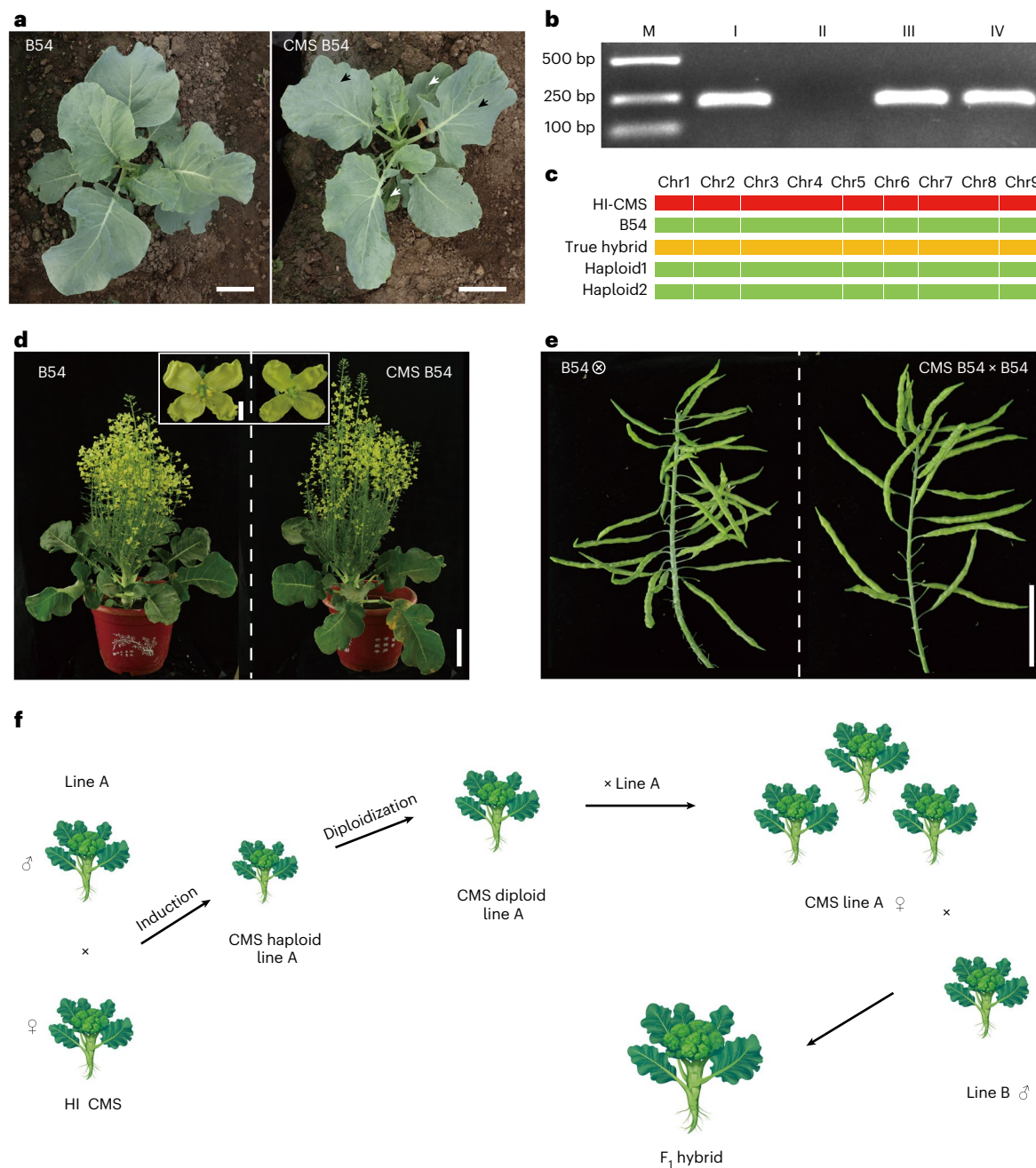
We employed a CRISPR/Cas9 system to create hypomorphic alleles of *BoCENH3*. A CRISPR/Cas9 construct was generated with two specific guide RNA sequences (sgRNA1 and sgRNA2) targeting the 7th exon of *BoCENH3*, located within the sequence encoding the putative  $\alpha$ -helix region<sup>26</sup> (Fig. 1a,b). The construct was subsequently introduced into the inbred line CX33, after which 21 independent transgenic lines were obtained. Among these transgenic lines, 11 lines harboured mutations in either of the targeted regions (3 for both sgRNA1 and sgRNA2, and 8 for only sgRNA2), corresponding to editing efficiencies of 14.29% for sgRNA1 and 52.38% for sgRNA2 (Supplementary Table 2). Eight lines were heterozygous or chimaeric mutants with deletions/insertions that led to frameshift and premature termination. Although not homozygous mutants, they displayed wrinkled leaves, a typical feature of defects in cell division (Supplementary Fig. 3). No homozygous mutants were produced from the T0 genome editing lines or selfing of T0 lines, consistent with the previous reports that *CENH3* is essential for plant development and reproduction<sup>12,26</sup>. Two representative heterozygous (*BoCENH3/bocenh3*) mutants were crossed by pollen donors, the broccoli lines B54 and 22TZ, and 2,512 progenies from four crosses were assessed (Supplementary Table 2). Although the heterozygous *CENH3* mutant (+/*cenh3*) with a null allele triggered haploid production in maize<sup>22</sup>, the broccoli heterozygous (*BoCENH3/bocenh3*) mutants did not induce haploids at the scale tested here (zero haploid in a screen of ~600 plants, for an average HI frequency of <0.16). The HI rate in lines homozygous for WT *CENH3* is probably considerably lower than this upper limit, as in a test for spontaneous paternal haploid in the related species *A. thaliana*, paternal haploids were not observed in a population of over 800,000 F<sub>1</sub> progenies (<1 × 10<sup>-6</sup>)<sup>27</sup>. The remaining two lines, #3 and #8, harboured in-frame deletions in targeted region 2 and a restored frameshift (RFS) between the two targeted regions, respectively (Fig. 1a,b and Supplementary Table 2). From the T1 generation, we generated non-transgenic homozygous mutants *BoCENH3<sub>Δ3</sub>* (*BoCENH3<sub>Δ3</sub>/BoCENH3<sub>Δ3</sub>*) and *BoCENH3<sub>RF533</sub>* (*BoCENH3<sub>RF533</sub>/BoCENH3<sub>RF533</sub>*), which exhibited a normal appearance compared with that of WT plants. When

selfing or crossing as a female parent, the *BoCENH3<sub>Δ3</sub>* and *BoCENH3<sub>RF533</sub>* mutants showed significantly reduced seed setting rates ( $P < 0.01$ ; unpaired, two-tailed *t*-test) (Fig. 1c,d).

To test whether the *BoCENH3<sub>Δ3</sub>* and *BoCENH3<sub>RF533</sub>* mutants could induce haploids, these mutants were crossed by 3 pollen donors: B54, B188 and 22TZ. These pollen donors were broccoli inbred lines with distinct genetic backgrounds and phenotypes (Supplementary Fig. 4). We developed 9 molecular markers (one for each chromosome) showing insertion–deletion (indel) polymorphisms between the mutants and the male parents to screen all the progenies. Potential haploids that exhibited genotypes identical to those of the male parents were identified (Fig. 1e and Supplementary Fig. 5). Flow cytometry analyses and chromosome spreads confirmed that all the plants were true haploids (Fig. 1f,g). These haploids were morphologically similar to the corresponding male parents but had smaller plant sizes, thinner leaves and male sterility (Fig. 1h–l). The HI rate was 0.52% on average for *BoCENH3<sub>Δ3</sub>* and 1.14% for *BoCENH3<sub>RF533</sub>*, substantially higher than that observed in selfed WT plants (0 in 647 progenies, HI rate < 0.15%), WT × B54 crosses (0 in 574 progenies, HI rate < 0.17%) and WT × 22TZ crosses (0 in 709 progenies, HI rate < 0.14%) (Fig. 1m). Nevertheless, this HI rate is relatively lower than the previously reported 1–44% in *Arabidopsis* (varied in different *CENH3* mutants)<sup>26</sup>, ~7% in wheat (in-frame modification of the N-terminal domain)<sup>21</sup> and 5% in maize (heterozygous mutant with a *cenh3* null allele)<sup>22</sup>. The generation of additional types of *BoCENH3* mutants may be helpful in increasing the HI rate in broccoli. In addition, we identified aneuploids from the outcrossing progenies (Fig. 1h,m and Supplementary Fig. 6). Given that some *CENH3* mutants could induce both paternal maternal and maternal haploids<sup>26</sup>, we tested whether the *BoCENH3* mutants could trigger maternal haploids in reciprocal crosses. Although the seed setting rate was also slightly reduced when these two mutants were crossed as males, no haploids were identified at the scale (2,699 individuals) tested here (Supplementary Table 4).

Almost all commercial *B. oleracea* hybrids are produced on the basis of the Ogura CMS system, which requires the creation of CMS lines<sup>28</sup>. Currently, CMS lines in *Brassica* crops are exclusively developed through the traditional backcrossing procedure<sup>28,29</sup>. Although swapping of CMS via a paternal HI line (*ig1*) is already used in maize<sup>23</sup>, this strategy has not been extended to other species. We tested whether the CMS cytoplasm can be exchanged to inbred lines via *BoCENH3*-based HI. To create HI lines in the Ogura CMS cytoplasm, the CMS line CMS219 was crossed with *BoCENH3<sub>RF533</sub>* pollen and then backcrossed to *BoCENH3<sub>RF533</sub>*. From the BC<sub>1</sub> generation, we identified homozygous *BoCENH3<sub>RF533</sub>* individuals with Ogura CMS cytoplasm (harbouring the *orf138* gene)<sup>28</sup>. The HI-CMS line was crossed with a B54 pollen donor. We identified 5 haploids from 492 progenies (1.02%) using molecular markers, flow cytometry and plant phenotype analyses (Fig. 2a and Supplementary Fig. 7). Using the *orf138* specific gene marker and nuclear genomic markers, we confirmed that these haploids had a B54 genetic background and Ogura CMS cytoplasm (Fig. 2b,c). These CMS B54 plants were treated with 200 mg l<sup>-1</sup> colchicine solution for chromosome doubling. We successfully obtained two diploid CMS-B54 lines, which were morphologically similar to its maintainer B54 and showed typical Ogura male sterile traits (Fig. 2a,d and Supplementary Fig. 8). When crossed to its maintainer line, CMS-B54 exhibited normal seed setting performance (Fig. 2e and Supplementary Fig. 9). To further confirm the genomic background of the haploids and diploids, the parental lines, two diploid CMS B54 plants and six B54 haploids were subjected to whole-genome resequencing. More than 318,000 single-nucleotide polymorphisms (SNPs) were identified from the samples. Most SNPs (>99.5%) from B54 CMS lines and B54 haploids showed genotypes identical to the original B54 line and no credible maternally derived SNPs were found (Supplementary Table 5), indicating that *BoCENH3<sub>RF533</sub>* could induce clean paternal haploids.

In summary, we successfully established a *CENH3*-based paternal HI system in broccoli, providing a novel and cost-effective DH technology



**Fig. 2 | One-step creation of CMS lines using the *BoCENH3*<sub>RFS33</sub> inducer line with an Ogura CMS cytoplasm. **a**, Phenotypes of WT B54 and doubled haploid CMS B54 seedlings. White arrows indicate haploid leaves before diploidization with colchicine and black arrows indicate diploid leaves after treatment with colchicine. **b**, PCR confirmation of the cytoplasm type using an *orf138* specific marker. M, DNA size marker; I, Ogura CMS line; II, inbred line B54; III–IV, B54**

haploids with Ogura CMS cytoplasm. Three independent experiments were performed with similar results. **c**, Genetic background of the HI-CMS line, B54, F<sub>1</sub> hybrid and haploid revealed by DNA markers on 9 chromosomes. **d, e**, Phenotype (**d**) and seed setting performance (**e**) of the CMS-B54 line. **f**, A proposed model of an effective CMS-based heterosis utilization system. Scale bars, 5 cm (**a**), 10 cm (whole plant in **d**), 5 cm (single flower in **d**) and 5 cm (**e**).

without genotype recalcitrance. Importantly, by creating paternal haploids with maternal cytoplasm, homozygous CMS lines can be created in one step, which is a breakthrough for crops using CMS systems for hybrid seed production (a proposed system is shown in Fig. 2f). The success of this HI system could accelerate the improvement of *B. oleracea* cultivars and pave the way for extending *CENH3*-based HI to other crops.

## Methods

### Plant materials

The broccoli inbred line CX33 was employed for the purpose of genetic transformation and *BoCENH3* knockout. The inbred lines B54, B188

and the 22TZ were used as test lines for HI of *BoCENH3* inducers. The broccoli CMS line CMS219 was used as a donor of Ogura cytoplasm. All plants and the derived transgenic plants were grown under local conditions in Beijing, China. All selfing and cross-pollinations were performed in a greenhouse.

### RNA isolation and RT–qPCR analyses

Fresh broccoli tissues, including leaves, stems, roots, floral buds, petals, sepals, mature pollens, pistils and siliques were sampled and submerged in liquid nitrogen. An RNA extraction kit (Tiangen Biotech, 4992239) was used for total RNA isolation from these broccoli tissues.

A complementary (c)DNA Synthesis kit (Tiangen, 4992910) was used for synthesizing cDNA following standard protocol. For qPCR assays, we prepared 25  $\mu$ l PCR mixture using SYBR Green master mix (Takara, RR82WR) and performed PCR experiments on an RT–qPCR detection platform (CFX96 Touch, Bio-Rad). As an internal control, RT–qPCR for the *Actin* gene of *B. oleracea* was also performed alongside the experimental group. Each assay was biologically repeated at least three times. The data analysis was performed using the  $2^{-\Delta\Delta CT}$  method<sup>30</sup>.

### Phylogenetic analyses

The protein sequence of BoCENH3 was downloaded from the broccoli draft genome HDEM ([https://www.genoscope.cns.fr/projet\\_BKL/cgi-bin/gbrowse/boleracea/](https://www.genoscope.cns.fr/projet_BKL/cgi-bin/gbrowse/boleracea/))<sup>31</sup>. To search for proteins homologous to this sequence, we utilized the BLASTP tool in Ensembl Plants (<http://plants.ensembl.org/index.html>) with BoCENH3 protein sequence as a query. A total of 12 BoCENH3 homologue proteins from 10 species were used for phylogenetic analyses. The protein sequences were subjected to alignment through ClustalW in MEGA software (v.7) and then a neighbour-joining phylogenetic tree was constructed using MEGA (Poisson model, 1,000 bootstrap replicates). Sequences of the homologous proteins used in this study are shown in Supplementary Table 1.

### Genetic transformation and CRISPR/Cas9-mediated gene editing in broccoli

The sequence of *BoCENH3* from broccoli CX33 was used for sgRNA design. We searched for the conserved motifs of  $\alpha$ -helix and centromere-targeting domain (CATD) in *BoCENH3* by comparing it to the known *Arabidopsis CENH3* (ref. 26). Two sgRNA sequences targeting the putative  $\alpha$ 2-helix region were chosen and inserted into a modified vector downstream of the *Arabidopsis* U6 promoter. We further cloned and inserted this cassette to a vector with a CaMV 35S promoter-derived *Bar* selection marker and a CaMV 35S promoter-derived *Cas9* gene as previously described<sup>32</sup>. The protocol for *Agrobacterium*-mediated broccoli transformation using the material CX33 has been described previously<sup>32,33</sup>. Positive transgenic plants were screened by Basta resistance and *Bar* gene specific marker.

The genomic fragment encompassing the sgRNA-targeted regions was amplified from positive transgenic plants and subjected to Sanger sequencing. To determine the mutations in T0 plants, PCR amplicons from these lines with expected mutations in *BoCENH3* were inserted into a cloning vector, introduced into *Escherichia coli* strain *EH5 $\alpha$*  and then sequenced (at least 20 independent clones for each line)<sup>32</sup>. Homozygous in-frame deletion and restored frameshift mutants of *BoCENH3* in the T1 or BC<sub>1</sub> generations were confirmed by PCR and Sanger sequencing.

### Whole-genome resequencing and genotyping

Genomic DNA was extracted from leaves of B54, *BoCENH3*<sub>RF533</sub> haploid inducer, two B54 CMS lines and six B54 haploids. DNA libraries were constructed and then subjected to whole-genome sequencing using the DNBSEQ-T7 platform. Low-quality reads were filtered using Trimmomatic software and the retained clean reads were processed with the BWA software by aligning the reads to the broccoli HDEM draft genome. Uniquely mapped reads were employed in the process of whole-genome SNP calling via GATK4 following the pipeline for GATK best practices. High-quality SNPs were selected and filtered following the reported parameters<sup>34</sup>. Only sites where genotyping was available for all accessions were retained.

### Haploid screening

The progenies of the *BoCENH3* mutants  $\times$  broccoli lines were analysed by genetic background, flow cytometry, chromosome number and plant phenotype analyses. For genetic background analyses, markers were developed on the basis of indel variations between the *BoCENH3* HI inducer and broccoli test lines, and used to genotype all the progenies. For potential haploids identified by molecular markers, flow cytometry

and chromosome number analyses were performed following previously described protocols<sup>35</sup>.

To determine the cytoplasmic background of the haploids or DH individuals, PCR was performed using an *orf138* (the causal gene of Ogura CMS) specific marker. Information on all primers used in the study can be found in Supplementary Table 6.

### Chromosome doubling

To generate DH lines, the identified CMS B54 haploid seedlings at the stage of 3–4 true leaves were treated with 200 mg l<sup>-1</sup> colchicine solution by root dip treatment for 20 h. The ploidy levels of plants were detected after treatment.

### Phenotyping and statistical analyses

Images of plant tissues were acquired using a digital camera (EOS M6, Canon). Subsequent processing of these images was conducted using Adobe Photoshop CS6 (v.13.0) and Adobe Illustrator 2022 (v.26.2). Statistical analyses were conducted using GraphPad Prism (v.9) and Microsoft Excel (2019).

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

The whole-genome resequencing data used in this study are accessible in NCBI through the accession code PRJNA1050660. The broccoli HDEM reference genome is publicly available ([https://www.genoscope.cns.fr/projet\\_BKL/cgi-bin/gbrowse/boleracea/](https://www.genoscope.cns.fr/projet_BKL/cgi-bin/gbrowse/boleracea/)). CENH3 homologue proteins can be obtained from the database Ensembl Plants (<http://plants.ensembl.org/index.html>). Source data are provided with this paper.

### References

- Chen, L. & Liu, Y. G. Male sterility and fertility restoration in crops. *Annu. Rev. Plant Biol.* **65**, 579–606 (2014).
- Liu, C. et al. Extension of the in vivo haploid induction system from diploid maize to hexaploid wheat. *Plant Biotechnol. J.* **18**, 316 (2020).
- Kelliher, T. et al. MATRILINEAL, a sperm-specific phospholipase, triggers maize haploid induction. *Nature* **542**, 105–109 (2017).
- Gautam, R. et al. Male sterility in plants: an overview of advancements from natural CMS to genetically manipulated systems for hybrid seed production. *Theor. Appl. Genet.* **136**, 195 (2023).
- Wan, X. et al. Breeding with dominant genic male-sterility genes to boost crop grain yield in the post-heterosis utilization era. *Mol. Plant* **14**, 531–534 (2021).
- Zhong, Y. et al. Establishment of a *dmp* based maternal haploid induction system for polyploid *Brassica napus* and *Nicotiana tabacum*. *J. Integr. Plant Biol.* **64**, 1281–1294 (2022).
- Li, Y. et al. An in planta haploid induction system in *Brassica napus*. *J. Integr. Plant Biol.* **64**, 1140–1144 (2022).
- Ravi, M. et al. A haploid genetics toolbox for *Arabidopsis thaliana*. *Nat. Commun.* **5**, 5334 (2014).
- Zhao, X. et al. In vivo maternal haploid induction based on genome editing of *DMP* in *Brassica oleracea*. *Plant Biotechnol. J.* **20**, 2242 (2022).
- Prigge, V. et al. New insights into the genetics of in vivo induction of maternal haploids, the backbone of doubled haploid technology in maize. *Genetics* **190**, 781–793 (2012).
- Zhong, Y. et al. Mutation of *ZmDMP* enhances haploid induction in maize. *Nat. Plants* **5**, 575–580 (2019).
- Ravi, M. & Chan, S. W. L. Haploid plants produced by centromere-mediated genome elimination. *Nature* **464**, 615–618 (2010).

13. Zhang, X. et al. A female in vivo haploid-induction system via mutagenesis of egg cell-specific peptidases. *Mol. Plant* **16**, 471–480 (2023).
14. Yao, L. et al. *OsMATL* mutation induces haploid seed formation in indica rice. *Nat. Plants* **4**, 530–533 (2018).
15. Zhong, Y. et al. In vivo maternal haploid induction in tomato. *Plant Biotechnol. J.* **20**, 250–252 (2021).
16. Wang, N. et al. In planta haploid induction by genome editing of *DMP* in the model legume *Medicago truncatula*. *Plant Biotechnol. J.* **20**, 22 (2022).
17. Talbert, P. B. et al. Centromeric localization and adaptive evolution of an *Arabidopsis* histone H3 variant. *Plant Cell* **14**, 1053–1066 (2002).
18. Sanei, M. et al. Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. *Proc. Natl Acad. Sci. USA* **108**, E498–E505 (2011).
19. Karimi-Ashtiyani, R. et al. Point mutation impairs centromeric *CENH3* loading and induces haploid plants. *Proc. Natl Acad. Sci. USA* **112**, 11211–11216 (2015).
20. Wang, Z. et al. A simple and highly efficient strategy to induce both paternal and maternal haploids through temperature manipulation. *Nat. Plants* **9**, 699–705 (2023).
21. Lv, J. et al. Generation of paternal haploids in wheat by genome editing of the centromeric histone *CENH3*. *Nat. Biotechnol.* **38**, 1397–1401 (2020).
22. Wang, N. et al. Haploid induction by a maize *cenh3* null mutant. *Sci. Adv.* **7**, eabe2299 (2021).
23. Kindiger, B. & Hamann, S. Generation of haploids in maize: a modification of the indeterminate gametophyte (ig) system. *Crop Sci.* **33**, 342–344 (1993).
24. Evans, M. M. S. The *indeterminate gametophyte1* gene of maize encodes a LOB domain protein required for embryo sac and leaf development. *Plant Cell* **19**, 46–62 (2007).
25. Flood, P. J. et al. Reciprocal cybrids reveal how organellar genomes affect plant phenotypes. *Nat. Plants* **6**, 13–21 (2020).
26. Kuppu, S. et al. A variety of changes, including CRISPR/Cas9-mediated deletions, in *CENH3* lead to haploid induction on outcrossing. *Plant Biotechnol. J.* **18**, 2068–2080 (2020).
27. Portemer, V. et al. Large genetic screens for gynogenesis and androgenesis haploid inducers in *Arabidopsis thaliana* failed to identify mutants. *Front. Plant Sci.* **6**, 147 (2015).
28. Yu, H. et al. Creation of fertility-restored materials for Ogura CMS in *Brassica oleracea* by introducing *Rfo* gene from *Brassica napus* via an allotriploid strategy. *Theor. Appl. Genet.* **133**, 2825–2837 (2020).
29. Shu, J. et al. Evaluation and selection of sources of cytoplasmic male sterility in broccoli. *Euphytica* **215**, 125 (2019).
30. Livak, K. & Schmittgen, T. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods* **25**, 402–408 (2001).
31. Genete, M. et al. Chromosome-scale assemblies of plant genomes using nanopore long reads and optical maps. *Nat. Plants* **4**, 879–887 (2018).
32. Liu, Y. et al. Functional characterization of *BoGL5* by an efficient CRISPR/Cas9 genome editing system in broccoli. *Sci. Hortic.* **319**, 112136 (2023).
33. Han, F. et al. A natural mutation in the promoter of *Ms-cd1* causes dominant male sterility in *Brassica oleracea*. *Nat. Commun.* **14**, 6212 (2023).
34. Zhong, Y. et al. A *DMP*-triggered in vivo maternal haploid induction system in the dicotyledonous *Arabidopsis*. *Nat. Plants* **6**, 466–472 (2020).
35. Zhu, M. et al. Introgression of clubroot resistant gene into *Brassica oleracea* L. from *Brassica rapa* based on homoeologous exchange. *Hortic. Res.* **9**, uhac195 (2022).

## Acknowledgements

This work was supported by the National Key Research and Development Program of China (Grant No. 2022YFF1003000 to F.H.), the National Natural Science Foundation of China (Grant No. 32302544 to F.H.; 32002042 to X.Z.; 32172580 to Z.L.), the China Agriculture Research System (Grant No. CARS-23-05A to Z.L.), and the Agricultural Science and Technology Innovation Program (Grant No. ASTIP to Z.L.).

## Author contributions

F.H. designed and conceived the overall research. F.H., Yuxiang Liu and X.Z. performed the experiments, analysed the data and wrote the manuscript. Yumei Liu, H.Z. and Z.L. contributed to the revision of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41477-024-01643-w>.

**Correspondence and requests for materials** should be addressed to Zhansheng Li.

**Peer review information** *Nature Plants* thanks Anne Britt, Jing Li and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a | Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection	In this study, no custom-made software or codes/mathematical algorithms were generated by the authors. Whole-genome resequencing data were generated using the high-throughput sequencing platform DNBSEQ-T7. Plant tissues were captured with a Canon EOS M6 digital camera (Canon, Tokyo, Japan).
Data analysis	Statistical analyses were conducted using GraphPad Prism (Version 9) and Microsoft Excel (Version 2019). Protein sequences were aligned by ClustalW in MEGA software (version 7). Phylogenetic tree was constructed with MEGA software (version 7). Whole-genome resequencing data were processed with GATK (Version 4.1.4.0), Trimmomatic (Version 0.39) and BWA (Version 0.7.17). Images were processed with Adobe Illustrator 2022 (version 26.2) and Adobe Photoshop CS6 (version 13.0).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The genomic resequencing data generated in this study have been deposited at NCBI under accession PRJNA1050660. Broccoli HDEM reference genome is public available ([https://www.genoscope.cns.fr/projet\\_BKL/cgi-bin/gbrowse/boleracea/](https://www.genoscope.cns.fr/projet_BKL/cgi-bin/gbrowse/boleracea/)). CENH3 homolog proteins can be obtained from the database Ensembl Plants (<http://plants.ensembl.org/index.html>). Source data are provided with this paper.

## Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	<input type="text" value="n/a"/>
Reporting on race, ethnicity, or other socially relevant groupings	<input type="text" value="n/a"/>
Population characteristics	<input type="text" value="n/a"/>
Recruitment	<input type="text" value="n/a"/>
Ethics oversight	<input type="text" value="n/a"/>

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We follow related studies and instructions for standard protocols to determine the sample size. For qRT-PCR, at least three independent samples were analyzed. For HI testing, at least 300 plants derived from each cross/selfing were screened. The sample size is sufficient as determined in related studies. We have cited the references in the manuscript. 18. Sanei, M et al. Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. <i>Proc. Natl. Acad. Sci. USA</i> 108, E498–E505 (2011). 19. Karimi-Ashtiyani, R. et al. Point mutation impairs centromeric CENH3 loading and induces haploid plants. <i>Proc. Natl. Acad. Sci. USA</i> 112, 11211–11216 (2015). 21. Lv, J. et al. Generation of paternal haploids in wheat by genome editing of the centromeric histone CENH3. <i>Nat. Biotechnol.</i> 38, 1397–1401 (2020). 22. Wang, N. et al. Haploid induction by a maize cenh3 null mutant. <i>Sci. Adv.</i> 7, eabe2299 (2021).
Data exclusions	No data was excluded from the analyses.
Replication	We conducted at least three biological replicates for each experiment. These information has been described in the methods or relevant figure legends.
Randomization	Plants used in this study were randomly grown in the same open field, greenhouse or growth chamber. Plants samples were collected randomly.
Blinding	Researchers were blinded during ploidy level detection. Researchers were blinded during plant cultivation, qRT-PCR, phylogenetic analyses, CRISPR/Cas9-mediated gene editing, genome resequencing, genotyping and HI rate calculation as blinding is not a necessary experimental design in plant sciences.



# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

## Methods

- | n/a                                 | Involvement                                            |
|-------------------------------------|--------------------------------------------------------|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Involved in the study         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies                    |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                 |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Plants             |

- | n/a                                 | Involvement                                        |
|-------------------------------------|----------------------------------------------------|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Involved in the study     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq                  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging    |

## Dual use research of concern

Policy information about [dual use research of concern](#)

### Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

- | No                                  | Yes                                                 |
|-------------------------------------|-----------------------------------------------------|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Public health              |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> National security          |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Crops and/or livestock     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Ecosystems                 |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Any other significant area |

### Experiments of concern

Does the work involve any of these experiments of concern:

- | No                                  | Yes                                                                                                  |
|-------------------------------------|------------------------------------------------------------------------------------------------------|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Demonstrate how to render a vaccine ineffective                             |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Confer resistance to therapeutically useful antibiotics or antiviral agents |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Enhance the virulence of a pathogen or render a nonpathogen virulent        |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Increase transmissibility of a pathogen                                     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Alter the host range of a pathogen                                          |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Enable evasion of diagnostic/detection modalities                           |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Enable the weaponization of a biological agent or toxin                     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Any other potentially harmful combination of experiments and agents         |

## Plants

Seed stocks	n/a
Novel plant genotypes	We used CRISPR/Cas9 system to generate BoCENH3 mutants. Information of targets for editing, sgRNAs and application of the editor was provided within this paper.
Authentication	We used broccoli inbred line CX33 to generate BoCENH3 mutants. We assessed the phenotypes, seed setting, and cross pollinations of the BoCENH3 mutants and their wild type plants.

## Flow Cytometry

### Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

Sample preparation	Young leaves were sampled, and subjected to lysate treatment to release nuclears. The samples were subsequently centrifuged with 1000rpm and dyed for 30 min.
Instrument	BD LSRFortessa™ Flow Cytometer, BD Biosciences, San Jose, CA, USA
Software	LSRFortessa
Cell population abundance	By centrifuge
Gating strategy	Bulk cell population

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.