Demonstration of targeted crossovers in hybrid maize using CRISPR technology

Andrei Kouranov (andrei.kouranov@bayer.com)
Bayer Crop Sciences
Charles Armstrong
Bayer Crop Sciences https://orcid.org/0000-0002-8780-6239
Ashok Shrawat
Bayer Crop Sciences
Vladimir Sidorov
Bayer Crop Sciences
Scott Huesgen
Bayer Crop Sciences
Bryce Lemke
Bayer Crop Sciences
Timothy Boyle
Bayer Crop Sciences
Michelle Gasper
Bayer Crop Sciences
Richard Lawrence
Bayer Crop Sciences
Samuel Yang
Bayer Crop Sciences

Article

Keywords: genetic diversity, chromosomal crossovers (CO), RNA guided nuclease technology

Posted Date: May 19th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-476417/v1

Version of Record: A version of this preprint was published at Communications Biology on January 13th, 2022. See the published version at https://doi.org/10.1038/s42003-022-03004-9.

Demonstration of targeted crossovers in hybrid maize using CRISPR technology.

Abstract:

Naturally occurring chromosomal crossovers (CO) during meiosis are a key driver of genetic diversity. The ability to target CO at specific allelic loci in hybrid plants would provide a significant advantage to the plant breeding process by facilitating trait introgression, and potentially increasing the rate of genetic gain. We present the first demonstration of targeted CO in hybrid maize utilizing the CRISPR Cas12a system. Our experiments showed that stable and heritable targeted CO can be produced in F1 somatic cells using Cas12a at a significantly higher rate than the natural CO in the same interval. These results are a step towards the use of RNA guided nuclease technology to simplify the creation of targeted genome combinations in progeny and accelerate breeding.

Introduction:

Improvement of genetic gain to increase yield while maintaining genetic diversity are key fundamentals of plant breeding (1). Variation in genetic gain and diversity is introduced by homologous recombination (HR) resolved by CO during meiosis. Naturally occurring CO are rare events, introduced at a frequency of about one to three CO per chromosome pair in meiotic cells (2, 3). Targeted CO at specific genomic loci could facilitate crop improvement by breaking linkage drag and/or stacking haplotypes associated with high yield. Bernardo (4) predicted that targeted CO at prescribed loci in hybrid maize can significantly improve genetic gain and increase yield. The author's analysis showed that maize grain yield can be doubled if one targeted CO is introduced in each chromosome (4). A similar effect of targeted CO on average yield gain was predicted in other crop species including soybean, wheat, barley, and pea (5). Plant breeders rely on natural CO, creating and screening large populations of biparental crosses to identify and stabilize the desired genotypes for crop improvement. However, this approach is time-consuming, expensive, and limited due to low CO frequencies at some genomic locations. CRISPR-Cas

technology provides an alternative genome engineering approach to implement targeted CO and accelerate breeding. Several advancements in precision genome engineering have already been demonstrated using this technology in plants, including homologous chromosome recombination (6), reciprocal trans-fragment translocations (7), and large fragment inversions (8, 9). In addition, the guided nuclease technology can be employed to induce homologous recombination in somatic cells, avoiding competition with the naturally occurring process during meiosis (6).

Results & Discussion:

To evaluate whether targeted CO can be directed by guided nuclease activity, we developed an experimental strategy to induce allelic CO by introducing DSBs (Double Stranded Break) at the same position in both parental chromosomes in hybrid somatic cells. As a guided nuclease, we employed *LbCas12a* whose expression was driven by strong constitutive Zm-Ubiquitin1 promoter (10). Two gRNA target sites about 180 kb apart, located between 170.9 and 173.1 cM on chromosome 3 were selected for targeted recombination (Figure 2 A).

We performed two independent plant transformation experiments using T-DNA plasmids containing the guided nuclease and either gRNA1 or gRNA2 to evaluate efficiency of targeted CO at two distinct chromosomal positions. F1 hybrid plants were produced by crossing two elite maize inbred lines (Parent A or Parent B) (Figure 1-1). Embryo explants were isolated from mature F1 hybrid seeds and separated into editing and control treatments for *Agrobacterium* -mediated transformation (Figure 1-2). The editing gRNAs mediated double-stranded breaks at the desired recombination locations whereas the control gRNA was lacking the target sites in the maize genomes. The regenerated F1-T0 plants were reciprocally backcrossed to one of the parents for the identification of chromosomal recombination (Figure 1-3 & 1-4). A small amount of endosperm tissue was non-destructively isolated from the harvested BC1-F1 seeds for genotyping analysis (Figure 1-5) to identify seeds with reciprocal CO between parental chromosomes (Figure 1-6).

The targeted CO at the selected target sites were assayed using 18 and 20 TaqMan markers, respectively (M1-M17, and M33 for the gRNA1 target site; M3, M16-M34 for the gRNA2 target site; Figure 2A). The assayed SNPs (Single-Nucleotide Polymorphism) were distributed across 556 kilobases (Kb) (M1-M33) and 359 Kb (M3-M34) regions. The most distant markers were located about 290 Kb, 140 Kb, 70 Kb, and 6 Kb on each side from the gRNA target site (see Supplementary material). In order to increase confidence that the targeted recombination was introduced precisely at the gRNA target sites we designed TaqMan assays located within a 6 Kb region spanning the recombination target sites (M4-M15 spanning gRNA1 and M19-M31 spanning gRNA2 target sites). The closest TaqMan markers were located about 400 nt and 200nt on each side of the targeted DNA break introduced by gRNA1 and gRNA2, respectively (Figure 2A: Marker 8 and 9, and Marker 25 and 26). Both gRNAs showed strong editing activity in the F1-T0 plants (Supplementary material: Figure 1).

In the first guided CO experiment 42 edited and 15 control F1-T0 plants were selected for reciprocal backcrossing with inbred Parent B. Subsets of BC1-F1 seeds from each backcross event were sampled for genotyping analysis (Figure 2B). Seeds from both reciprocal backcrosses were included in the analysis (Supplementary Table 1). The DNA extracted from the seed tissue was genotyped using 18 TaqMan PCR assays (Figure 2A, M1-M17, M33). We observed a significant difference in the recombination frequency between the edited and control BC1-F1 seed populations in the 0.8 Kb (between M8 and M9) region spanning the guided nuclease target site (Figure 2B). Within this narrow interval, overall recombination frequency in the edited population was 0.71% (30 out of 4200 genotyped seeds) versus 0.04% (1 out of 2265 genotyped seeds) in the control, a significant increase of ~18-fold in the presence of the guided nuclease. Chi-square test of independence validated by permutation test confirmed that the observed difference in recombination frequency between the control and treatment is statistically significant. The probability that the frequency of recombination in the edited population is independent of guided nuclease activity is significantly low (p-value < 0.001). However, we did not observe an overall increase in recombination frequency between the edited and control populations within the larger 556 Kb (between

M1 & M33) interval (Figure 2B). The observed 3% recombination frequency within this interval is consistent with previously defined natural recombination rates (data not shown), suggesting a precise and specific increase in targeted recombination at the gRNA1 target site.

Recombinant chromosomes from both reciprocal CO were identified by genotyping (Figure 2D). About 50% recombinant BC1-F1 seeds (14 of 30) from edited plants were *LbCas12a* negative, suggesting germinal transmission of the recombinant chromosome from the F1-T0 plant to the next generation (Figure 2D).

Our analysis showed that BC1-F1 seeds with recombinant genotype derived from only three edited F1-T0 plants (Figure 2B: Event 1, Event 2, Event 3). We genotyped all seeds derived from these three edited and three control F1-T0 plants and compared the recombination frequency (Figure 2B). One of the control plants was selected for the comparison because it contained one seed with CO within the 0.8 Kb targeted region (Figure 2B: Event 4). Two other control seed sets were selected randomly (Event 5 and Event 6). The rate of targeted CO in two of the three edited plants ranged from 3.5 to 6.0 % versus 0.1% in the control (Figure 2C). These editing rates are consistent with previously reported genome editing rates in plants (18,19). The single recombinant seed identified in Event 3 may potentially represent a natural CO. All recombinant BC1-F1 seeds listed in Table 2C derived from backcrosses where F1-T0 plants were used as females suggesting that guided CO transmitted to the female gamete before pollination with the inbred pollen. We were not able to identify examples of the guided CO in the male gametes in this experiment for Even1, Event2, or Event3. However, this may be due to the lower number of the genotyped seeds derived from the reciprocal backcross (173 versus 887 seeds, see Supplementary Table1 and 2 for additional details).

In our second genome editing experiment we intended to confirm that CO in somatic cells could be targeted at a different gRNA site and transmitted to the male gametes. We produced and backcrossed 43 F1-T0 plants (Figure 2E) as males onto Parent A. The total number of edited and control BC1-F1 seeds

genotyped in this experiment are shown in Figure 2E. We used a subset of seeds from the same negative control population used in the analysis of the first genome editing experiment to evaluate and compare the efficiency of targeted and natural CO. Similar to the edited BC1-F1 seed population, all control seeds were produced by crossing the control F1-T0 plants as males onto inbred females. We employed 20 TaqMan PCR assays to evaluate targeted recombination at the gRNA2 target site (M3, M16-M34, Figure 2A). In this experiment, we observed a higher recombination rate in the edited versus control BC1-F1 seed populations within the 359 Kb and 0.4 Kb regions (Figure 2, Table E). The recombination rate within the 359 Kb region (between M3 and M34) was 4-fold higher in the edited than the control seed population. Remarkably, most of the observed targeted CO were within the narrow 400 nt region between M25 and M26, representing a rate of 3.6%. We did not identify any seeds in the control population with chromosomal recombination between the M25 and M26 markers. All seeds with a targeted CO (total number: 175 seeds) were derived from a single F1-T0 plant: ZM_S22440456. This result suggests that a guided CO was introduced into an early meristematic precursor cell and transmitted to developing male gametes. Both reciprocal chromosomal CO were identified in the ZM_S22440456 BC1-F1 seed population in about equal proportions (Figure 2F; 72 and 96 seeds respectively). Examples of both reciprocal CO are shown in Figure 2G.

In order to further characterize the precise nature of the targeted CO at the gRNA target site, we selected all LbCas12a-negative BC1-F1 seeds with targeted CO from the first genome editing experiment with gRNA1 (Total number 13 seeds: Event-1, 2, and 3. Figure 2C) for further analysis. The seeds were germinated, and leaf tissue was re-analyzed by genotyping which confirmed targeted CO and the absence of the LbCas12a cassette (data not shown). To analyze the editing patterns at the site of targeted CO the gRNA target sites were amplified and sequenced using Illumina technology. As non-edited controls, we included samples from Event-4 plants, one of which showed natural CO between M8 and M9 SNP markers. Our analysis showed that in 5 out of 13 analyzed edited plants DNA breaks were repaired without any mutations (Events 1.4, 1.8, 1.9, 2.10, 2.11, Figure 3B). Error prone NHEJ is the most

prevalent DNA-repair pathway in plant somatic cells. This mechanism may introduce deletions or insertions at the site of DSB repair, but it may also be repaired without any mutations (11,20). We also cannot exclude that the observed accurate DNA repair was driven by HDR pathway. Only one DNA editing pattern was identified in each BC1-F1 plant sample consistent with stability of the targeted CO transmission in nuclease-negative progeny (Figure 3B). We observed multiple editing patterns in BC1-F1 plants positive for LbCas12a transgene (data not shown). Interestingly, DNA editing patterns at the targeted CO site differed across BC1-F1 progeny derived from Event-1. In eight examined Event-1 siblings we observed four different DNA repair patterns (Figure 3B). This result is consistent with mosaic editing outcomes frequently observed in the presence of CRISPR-mediated nucleases in plants (12). The mosaicism can be explained by the editing machinery expression delay relative to the meristematic cell division. Corn ear contains 400-500 kernels that develop from different meristematic precursor cells (13). The germ line fate of the somatic precursors is defined late in plant development after several cell divisions, leading to variation in DNA repair patterns. Targeted CO with repair without any change at the target sequence could also be subjected to recurrent editing after additional cell divisions.

To confirm the stability of the targeted CO, we selected several LbCas12a-negative BC1-F1 seeds heterozygous for the chromosome recombination (Event-1.18, Event-2.10, and Event-2.11: Figure 3A). The plants were self-pollinated to produce BC1-F2 segregating population. The segregating populations of the one-week old BC1-F2 seedings were sampled and genotyped with 18 TaqMan assays. The genotyping analysis identified individual plants homozygous for the targeted CO (Figure 4B: labeled Light Green and Blue) at the expected segregation ratio 1:2:1 in all three tested segregating populations. Chromosome recombination in this study was stable and did not cause any negative impact on plant growth or development. The analysis of the guided CO sites in the homozygous recombinant BC1-F2 plants with Illumina sequencing confirmed previously observed DNA editing patterns (compare Figure 3B and Figure 4C). BC1-F2 siblings from Event-1.18 showed a deletion at the point of the guided CO

presumably driven by short nucleotide microhomology (Figure 4C, highlighted in gray), whereas in siblings from Event-2.10 & 11 DNA break was repaired without any mutation.

In summary, we demonstrated guided CO in F1 hybrid maize in two independent experiments using a guided LbCas12a nuclease. Both experiments confirmed that a transgenically introduced guided nuclease can effectively drive targeted CO in somatic maize cells transmittable to the next generation. Our study represents an initial step in genome engineering utilizing CRISPR-associated nuclease technology with a potential to improve genetic gain in commercial crops.

Material & Methods

Plant material & plant transformation

F1 seeds were produced by cross-pollinating two maize (*Zea mays*) inbred lines, a proprietary elite Stiff Stalk female line (Parent A) and LH244 (Parent B) by hand pollination in a field. Pollen was collected from Parent B tassels and crossed onto silks of Parent A. F1 embryo explants were transformed with *Agrobacterium tumefaciens (14)* using cp4 as the plant selectable marker (15).

Plasmids & agrobacterium strains

Agrobacterium tumefaciens AB32 strain (16) carrying pCpf1_gRNA1, pCpf1_gRNA2, or pCpf1_gRNA2 plasmids with a plant selectable marker (*cp4*) were used for transformation. All three plasmids contained *LbCpf1* and gRNA expression cassettes. Editing plasmids expressed gRNA1 (pCpf1_gRNA1) or gRNA2 (pCpf1_gRNA2), a control plasmid (pCpf1_gRNA_control) expressed a "dummy" gRNA lacking the target sites in the maize genomes. Plasmid structural components driving expression of the genome editing machinery and the selection marker are listed in Supplementary material**.**

Generation of BC1-F1 seeds for genotyping

The transgenic F1-T0 generation and the parental inbred line used for reciprocal backcrossing were grown in a greenhouse with a day temperature of 29°C and a night temperature of 21°C, with supplemental lighting added in order to provide a 16-hour daylength. The inbred line was planted prior to, at the same time as, and after the F1 generation, in order to ensure overlap of pollen shed of the male population and silking of the female population. The pollination procedure was performed by hand-pollination technique. The F1-T0 plants were backcrossed with Parent B (LH244) or Parent A in the first and the second genome editing experiment, respectively. The change of the inbred parent in the backcrossing scheme was due to ear/pollen synchronized inbred plant population available in the greenhouse at the time of the experiment.

BC1-F1 seed chipping and genotyping

Subsets of 80 to 120 BC1-F1 seeds produced from each backcrossing event were sampled for genotyping analysis. Nondestructive sampling of BC1-F1 seed populations was performed using an automated high throughput seed chipper (17). The small amount of endosperm tissue was collected from each seed into 96 well plates; DNA was isolated and genotyped using quantitative endpoint TaqMan PCR assay.

Qualitative endpoint TaqMan assays were performed using TaqPath ProAmp master mix obtained from ThermoFisher Scientific according to manufacture protocol. Thermal cycling was performed on an Applied Biosystems GeneAmp PCR system 9700 and fluorescence measurement by Tecan Spark microplate reader. TaqMan FAM- and VIC-labeled probes were obtained from ThermoFisher Scientific and primers were obtained from either ThermoScientific or Integrated DNA Technologies.

DNA isolation

DNA from leaf or seed endosperm tissue was extracted using a DNA-binding filter method. Plant tissue samples were collected into 96-deep well plates, frozen, and lyophilized prior to extraction. Samples were ground by paint shaker with 3/16 in. stainless steel balls in 440µl extraction buffer (0.1M Tris-HCl pH 8.0, 0.05M EDTA, 0.1M NaCl, 1%SDS), preheated to 65ºC. Following grinding, samples were

incubated at 65ºC for 45 minutes, followed by addition of 135 µl of 5M potassium acetate. After a brief centrifugation, 40µl of cleared lysate was added along with 40µl of isopropanol to a 384-well binding filter (PALL), and the plates were centrifuged to remove the liquid waste. The bound DNA was washed with 50µl of 70% ethanol and centrifuged. The bound DNA was eluted with 60µl of DNA elution buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA).

Evaluation of gRNA editing activity by Illumina sequencing

Samples were prepared following the Illumina DNA Prep protocol (formerly FLEX) provided by Illumina using in-house primers, adapted from the original Illumina primer design with unique oligonucleotides that incorporate the Illumina adaptor overhang, unique indexes and genome target specific primers. All samples (~5 ng of total DNA) were amplified in separate reactions using Phusion High-Fidelity DNA Polymerase (ThermoFisher). Amplicons of 424 nt or 283 nt representing gRNA target region#1 and target #2 respectively were purified using Agencourt AMPure XP beads and correctly sized products were verified using an Agilent 2100 Bioanalyzer DNA 1000 chip. Individually barcoded samples were pooled equimolarly and sequenced on an Illumina NextSeq using the NextSeq Reagent Kit 2 × 150 bp paired-end sequencing kit.

The reads from Illumina libraries were mapped to the genome reference sequences representing gRNA target regions to identify edited or wild type reads. The editing activity was measured as a percentage of edited reads in the total number of reads mapped to the reference sequence. Only reads with deletions but not substitution were counted as edited reads.

Statistical analysis

We used the Chi-square test of independence to confirm statistical significance of the observed difference in recombination frequency between the control and treatment in the first genome editing experiment with gRNA1. However, the precision of the Chi-square test result decreases when any the contingency table values are low.

Contingency table:

	Seeds with No Cross-over	Seeds with Cross-over	Total
Treatment	4170	30	4200
Control	2264	1	2265
Total	6434	31	6465

To validate the significances of the Chi-square test result, we implemented a permutation test that is not impacted by the low values in the contingency table. The following R functions were used to calculate the estimated p-value:

```
# Create the contingency table
m = matrix(c(30,1,4170,2264), nr=2, by=T)

# Perform Pearson's Chi-squared test
chisq.test(m)
data:  m
X-squared = 13.847, df = NA, p-value = 0.00024

# Perform Pearson's Chi-squared test with simulated p-value (based on 1e+05 replicates)

chisq.test(m, sim=T, B=1e5)
data:  m
X-squared = 13.847, df = NA, p-value = 0.00028
```

References:

1. Lenaertsa, B., Collardb, B.C.Y., Demontb, M. Improving global food security through accelerated plant breeding. Plant Science. **287**, 110207, (2019).
2. Mercier, R., Mezard, C., Jenczewski, E., Macaisne, N. and Grelon, M. The Molecular Biology of Meiosis in Plants. Annu. Rev. Plant Biol. **66**, 297-327, (2015)
3. Sidhua, G.K., Fanga, C., Olsona, M.A., Falqueb, M., Martinb, O.C. & Pawlowski, W.P. Recombination patterns in maize reveal limits to crossover homeostasis. PNAS **112**, 52, 15982–15987, (2015)
4. Bernardo, R. Prospective Targeted Recombination and Genetic Gains for Quantitative Traits in Maize. Plant Genome. **10**, 2, 1-9, (2017)
5. Ru, S., Bernardo, R. Targeted recombination to increase genetic gain in self-pollinated species. Theor. Appl. Genet. **132**, 2, 289-300, (2019)
6. Filler, H.S., Melamed, B.C., Levy, A.A. Targeted recombination between homologous chromosomes for precise breeding in tomato. Nature Communications **8**, 15605, (2017)
7. Beying, N., Schmidt, C., Pacher, M., Houben, A., Puchta H. CRISPRCas9-mediated induction of heritable chromosomal translocations in Arabidopsis. Nature Plants. **6**, 638–645, (2020)
8. Schmidt, C., Pacher, M., Puchta, H. Efficient induction of heritable inversions in plant genomes using the CRISPR/Cas system. The Plant Journal **98**, 577–589, (2019)
9. Schwartz, C., Lenderts, B., Feigenbutz, L., Barone, P., Llaca, V., Fengler, K. and Svitashev, S. CRISPR–Cas9-mediated 75.5-Mb inversion in maize. Nature Plants. **6**, 1427–1431, (2020)
10. Svitashev, S., Schwartz, C., Lenderts, B., Young, J.K., Cigan, A.M. Genome editing in maize directed by CRISPR-Cas9 ribonucleoprotein complexes. Nat. Commun. **16**,7, 13274, (2016)
11. Geisinger, J.M, Turan, S., Hernandez, S., Spector, L.P., Michele P. Calos, M.P. In vivo blunt-end cloning through CRISPR/Cas9-facilitated non-homologous end-joining. Nucleic Acids Research. **44**, 8, e76, (2016)

12. Liu, H.J., Jian, L., Xu, J., Zhang, Q., Zhang, M., Jin, M., Peng, Y., Yan, J., Han, B., Jie Liu, Gao, F., Liu, X., Huang, L., Wei, W., Ding, Y., Yang, X., Li, Z., Zhang, M., Sun, J., Bai, M., Song, W., Chen, H., Sun, X., Li, W., Lu, Y., Liu, Y., Zhao, J., Qian, Y., Jackson, D., Fernie, A.R., Yan, J. High-Throughput CRISPR/Cas9 Mutagenesis Streamlines Trait Gene Identification in Maize. The Plant Cell. **32**, 1397–1413, (2020)
13. Dumas, C., and Mogensen, H.L. Gametes and fertilization: maize as a model system for experimental embryogenesis in flowering plants. Plant Cell **5**, 1337–1348, (1993)
14. Sidorov, V., Duncan, D. Agrobacterium-Mediated Maize Transformation: Immature Embryos Versus Callus. Methods In Molecular Biology, Book: Transgenic Maize. 47-58, (2009).
15. Howe, A.R., Gasser, C.S, Brown, S.M., Stephen R. Padgette, S.R., Hart, J., Parker, G.B., Fromm, M.E. & Armstrong, C.L. Glyphosate as a selective agent for the production of fertile transgenic maize (Zea mays L.) plants. Molecular Breeding. **10**, 153–164, (2002)
16. Ye, X., Chen, Y., Wan, Y., Hong, Y.J., Ruebelt, M.C. and Gilbertson, L.A. Constitutive expression of the tzs gene from Agrobacterium tumefaciens virG mutant strains is responsible for improved transgenic plant regeneration in cotton meristem transformation. Plant Cell Reports. **35**, 601–611, (2016)
17. Deppermann, K.L., Yannakakis, L., Singleton, A.M., Koestel, A.R., Finley, D.W., Forinash, B.J., Ondes, A.N., Eathington, S.R., Forbes, H.M., Schnicker, B.J., Bull, J.K. Automated high-throughput seed sampler and methods of sampling, testing and bulking seeds. Patent Number US 20080317279A1, (2011)
18. Begemann, M.B., Gray, B.N., January, E., Gordon, G.C., He, Y., Liu, H., Wu, X., Brutnell, T.P., Mockler, T.C. & Oufattole, M.. Precise insertion and guided editing of higher plant genomes using Cpf1 CRISPR nucleases. Sci. Rep., **7,** 11606, (2017)
19. Wolter, F., Puchta, H.. In planta gene targeting can be enhanced by the use of CRISPR/Cas12a. The Plant Journal, **100**, 5, 1083-1094, (2019)

20. Be´termier, M., Bertrand, P., Bertrand, S.L. Is Non-homologous end-joining really an inherently errorprone process? PLoS Genet. 10, e1004086 (2014).

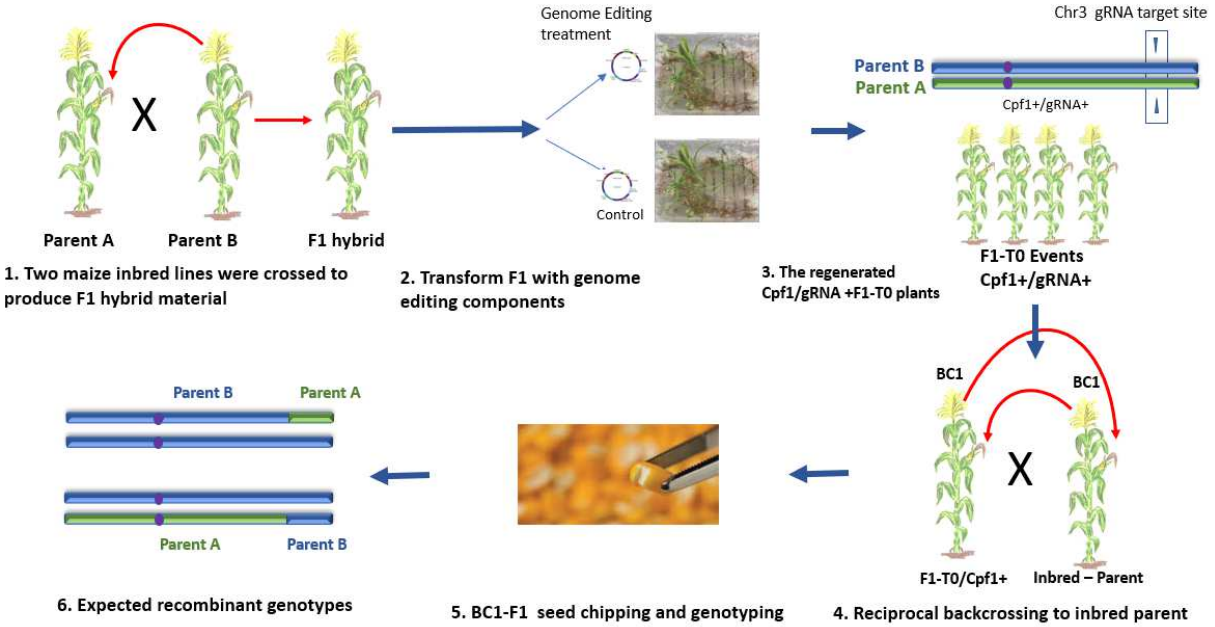
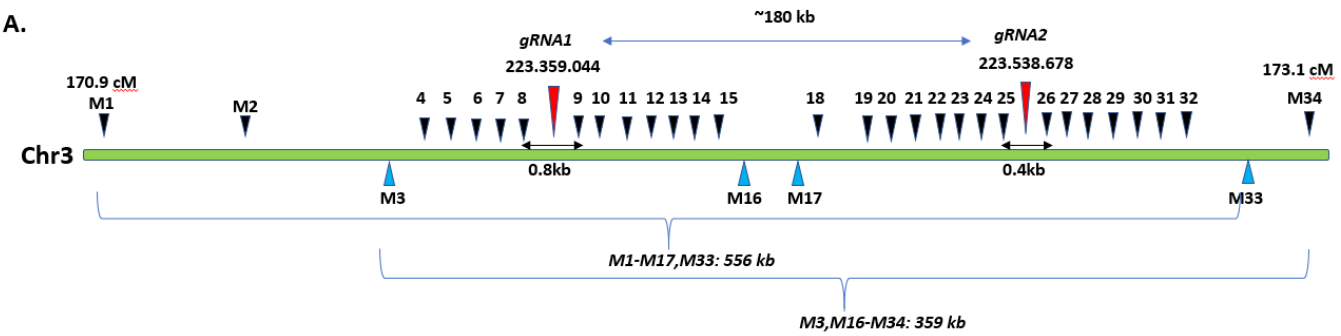


Figure 1: Experimental workflow to demonstrate guided homologous chromosome recombination in maize. 1) Two maize inbred lines were crossed to produce F1 hybrid material. 2) Isolated F1 embryo explants were separated into two groups for *Agrobacterium* -mediated transformation: a) editing treatment; b) control treatment. 3-4) The regenerated F1-T0 plants were reciprocally back-crossed to Parent B (the first experiment: gRNA1) or Parent A (the second experiment: gRNA2) for the identification of CO by genotyping. 5) The harvested BC1-F1 seeds were chipped to isolate a small amount of endosperm tissue for genotyping analysis using SNP TaqMan assays. 6) Isolation reciprocal chromosomal recombination between parental chromosomes.

A.



B.

Constructs	Transformation	Events#	Total seeds genotyped	Seeds with Recombinations between M1-M33 556 kb region (%)	Seeds with Recombinations between M8-M9 0.8 kb region (%)
pCpf1_gRNA1	Editing plasmid	42	4200	135 (3%)	30 (0.71%)
pCpf1_gRNA_control	Control plasmid	15	2265	69 (3%)	1 (0.04%)

C.

Transformation	Event #	Total seeds genotyped	Seeds with Recombinations between M1-M33 556 kb region (%)	Seeds with Recombinations between M8-M9 0.8 kb region (%)	Cpf1 negative seeds
pCpf1_gRNA1	Event 1	308	27 (9%)	*18 (6%)*	8
pCpf1_gRNA1	Event 2	311	20 (6.5%)	*11 (3.5%)*	4
pCpf1_gRNA1	Event 3	461	19 (4%)	*1 (0.2%)*	1
pCpf1_gRNA_control	Event 4	542	13 (2.4%)	*1 (0.1%)*	
pCpf1_gRNA_control	Event 5	543	15 (3%)	*0*	
pCpf1_gRNA_control	Event 6	196	1 (0.5%)	*0*	

D.

Recomb#	Event	Construct	M1	M2	M3	M4	M5	M6	M7	M8	gRNA1	M9	M10	M11	M12	M13	M14	M15	M16	M17	M33	Cpf1	Recomb.Genotype
1	Event 1.1	Edited event	TT	GG	GG	CC	TT	CC	AA	AA		AG	GC	CT	TC	TG	GA	TC	AG	GA	TC	POS	B/A
2	Event 1.2	Edited event	TT	GG	GG	CC	TT	CC	AA	AA		AG	GC	CT	TC	TG	GA	TC	AG	GA	TC	POS	B/A
3	Event 1.3	Edited event	TT	GG	GG	CC	TT	CC	AA	AA		AG	GC	CT	TC	TG	GA	TC	AG	GA	TC	POS	B/A
4	Event 1.4	Edited event	TT	GG	GG	CC	TT	CC	AA	AA		AG	GC	CT	TC	TG	GA	TC	AG	GA	TC	NEG	B/A
5	Event 1.5	Edited event	TT	GG	GG	CC	TT	CC	AA	AA		AG	GC	CT	TC	TG	GA	TC	AG	GA	TC	NEG	B/A
6	Event 1.6	Edited event	TT	GG	GG	CC	TT	CC	AA	AA		AG	GC	CT	TC	TG	GA	TC	AG	GA	TC	POS	B/A
7	Event 2.1	Edited event	TT	GG	GG	CC	TT	CC	AA	AA		AG	GC	CT	TC	TG	GA	TC	AG	GA	TC	POS	B/A
8	Event 2.2	Edited event	TT	GG	GG	CC	TT	CC	AA	AA		AG	GC	CT	TC	TG	GA	TC	AG	GA	TC	NEG	B/A
9	Event 2.3	Edited event	TT	GG	GG	CC	TT	CC	AA	AA		AG	GC	CT	TC	TG	GA	TC	AG	GA	TC	POS	B/A
10	Event 2.4	Edited event	TT	GG	GG	CC	TT	CC	AA	AA		AG	GC	CT	TC	TG	GA	TC	AG	GA	TC	POS	B/A
11	Event 4.1	Control event	TT	GG	GG	CC	TT	CC	AA	AA		AG	GC	CT	TC	TG	GA	TC	AG	GA	TC	POS	B/A
12	Event 1.7	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	NEG	A/B
13	Event 1.8	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	NEG	A/B
14	Event 1.9	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	NEG	A/B
15	Event 1.10	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	POS	A/B
16	Event 1.11	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	NEG	A/B
17	Event 1.12	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	POS	A/B
18	Event 1.13	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	POS	A/B
19	Event 1.14	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	NEG	A/B
20	Event 1.15	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	POS	A/B
21	Event 1.16	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	POS	A/B
22	Event 1.17	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	POS	A/B
23	Event 1.18	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	NEG	A/B
24	Event 2.5	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	POS	A/B
25	Event 2.6	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	NEG	A/B
26	Event 2.7	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	POS	A/B
27	Event 2.8	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	POS	A/B
28	Event 2.9	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	POS	A/B
29	Event 2.10	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	NEG	A/B
30	Event 2.11	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	NEG	A/B
31	Event 3.1	Edited event	CT	AG	AG	TC	AT	GC	TA	GA		GG	CC	TT	CC	GG	AA	CC	GG	AA	CC	NEG	A/B

E.

Constructs	Transformation	Event#	Total Seeds Genotyped	Seeds with Recombination between M3-M34 359 kb region	Seeds with Recombination between M25-M25 0.4 kb region
pCpf1_gRNA2	Editing plasmid	43	4798	197(4%)	175(3.6%)
pCpf1_gRNA_control	Control plasmid	15	3370	37(1%)	0

F.

Transformation	Event	Seeds with B/A recombination	Seeds with A/B recombination	Total Seeds genotyped	Seeds with B/A recombination Cpf1 negative	Seeds with A/B recombination Cpf1 negative	Total Seeds Cpf1 negative
pCpf1_gRNA2	ZM_S22440456	79	96	175	34	57	91

G.

Recomb#	Event	Construct	M3	M16	M17	M18	M19	M20	M21	M22	M23	M24	M25	gRNA2	M26	M27	M28	M29	M30	M31	M32	M33	M34	Cpf1	Recomb.genotype
1	ZM_S22440456	Editing	AG	AG	AG	AG	AG	AG	AG	AG	CT	AG	GT		CC	AA	AA	TT	AA	AA	GG	TT	GG	POS	B/A
2	ZM_S22440456	Editing	AG	AG	AG	AG	AG	AG	AG	AG	CT	AG	GT		CC	AA	AA	TT	AA	AA	GG	TT	GG	NEG	B/A
3	ZM_S22440456	Editing	AG	AG	AG	AG	AG	AG	AG	AG	CT	AG	GT		CC	AA	AA	TT	AA	AA	GG	TT	GG	NEG	B/A
4	ZM_S22440456	Editing	AG	AG	AG	AG	AG	AG	AG	AG	CT	AG	GT		CC	AA	AA	TT	AA	AA	GG	TT	GG	POS	B/A
5	ZM_S22440456	Editing	AG	AG	AG	AG	AG	AG	AG	AG	CT	AG	GT		CC	AA	AA	TT	AA	AA	GG	TT	GG	NEG	B/A
6	ZM_S22440456	Editing	AG	AG	AG	AG	AG	AG	AG	AG	CT	AG	GT		CC	AA	AA	TT	AA	AA	GG	TT	GG	NEG	B/A
7	ZM_S22440456	Editing	AG	AG	AG	AG	AG	AG	AG	AG	CT	AG	GT		CC	AA	AA	TT	AA	AA	GG	TT	GG	NEG	B/A
8	ZM_S22440456	Editing	AG	AG	AG	AG	AG	AG	AG	AG	CT	AG	GT		CC	AA	AA	TT	AA	AA	GG	TT	GG	NEG	B/A
9	ZM_S22440456	Editing	AG	AG	AG	AG	AG	AG	AG	AG	CT	AG	GT		CC	AA	AA	TT	AA	AA	GG	TT	GG	POS	B/A
10	ZM_S22440456	Editing	AG	AG	AG	AG	AG	AG	AG	AG	CT	AG	GT		CC	AA	AA	TT	AA	AA	GG	TT	GG	NEG	B/A
11	ZM_S22440456	Editing	AA	AA	GG	AA	GG	GG	GG	GG	CC	GG	TT		CG	AG	AT	GT	AG	AG	GT	CT	AG	NEG	A/B
12	ZM_S22440456	Editing	AA	AA	GG	AA	GG	GG	GG	GG	CC	GG	TT		CG	AG	AT	GT	AG	AG	GT	CT	AG	NEG	A/B
13	ZM_S22440456	Editing	AA	AA	GG	AA	GG	GG	GG	GG	CC	GG	TT		CG	AG	AT	GT	AG	AG	GT	CT	AG	POS	A/B
14	ZM_S22440456	Editing	AA	AA	GG	AA	GG	GG	GG	GG	CC	GG	TT		CG	AG	AT	GT	AG	AG	GT	CT	AG	NEG	A/B
15	ZM_S22440456	Editing	AA	AA	GG	AA	GG	GG	GG	GG	CC	GG	TT		CG	AG	AT	GT	AG	AG	GT	CT	AG	NEG	A/B
16	ZM_S22440456	Editing	AA	AA	GG	AA	GG	GG	GG	GG	CC	GG	TT		CG	AG	AT	GT	AG	AG	GT	CT	AG	POS	A/B
17	ZM_S22440456	Editing	AA	AA	GG	AA	GG	GG	GG	GG	CC	GG	TT		CG	AG	AT	GT	AG	AG	GT	CT	AG	POS	A/B
18	ZM_S22440456	Editing	AA	AA	GG	AA	GG	GG	GG	GG	CC	GG	TT		CG	AG	AT	GT	AG	AG	GT	CT	AG	NEG	A/B
19	ZM_S22440456	Editing	AA	AA	GG	AA	GG	GG	GG	GG	CC	GG	TT		CG	AG	AT	GT	AG	AG	GT	CT	AG	NEG	A/B
20	ZM_S22440456	Editing	AA	AA	GG	AA	GG	GG	GG	GG	CC	GG	TT		CG	AG	AT	GT	AG	AG	GT	CT	AG	NEG	A/B

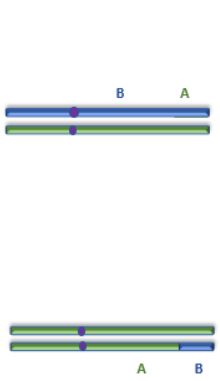
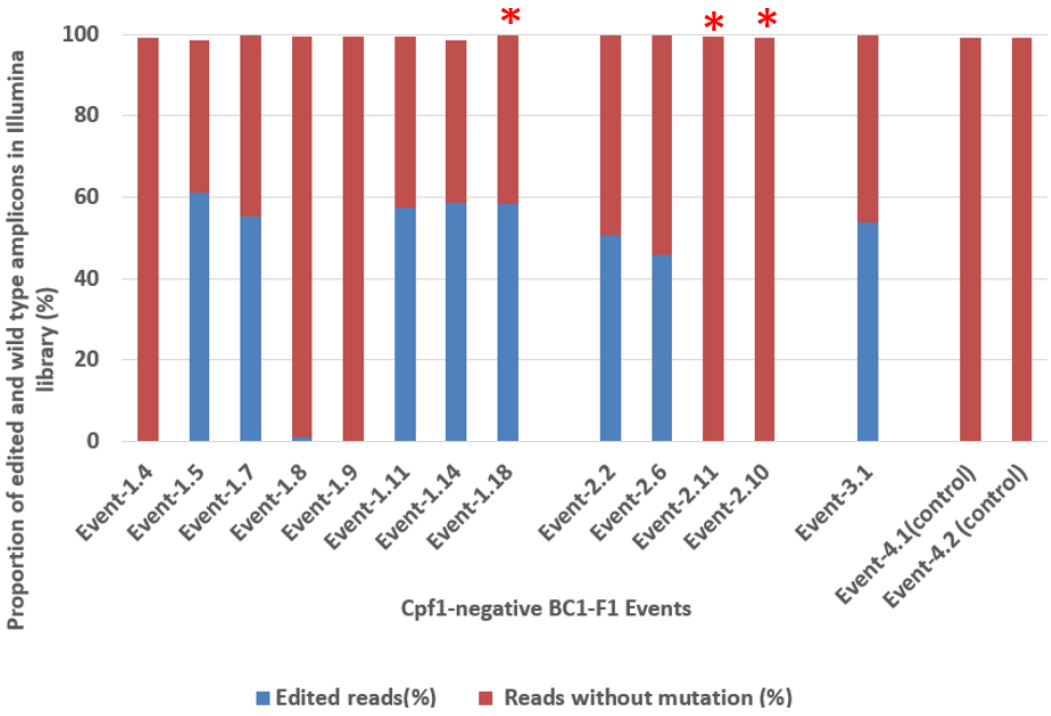


Figure 2: Result of genotyping to evaluate recombination frequency at the target sites gRNA1 and gRNA2, respectively. A) Schematic diagram of the two target regions (gRNA1 and gRNA2) and location of polymorphic SNP markers (M1-M34) used in the genotyping assays. Schematic SNP marker positions are shown as black and blue triangles. The blue triangles represent SNP markers that were shared in two different genome editing experiments. The physical genome coordinates of the gRNA target sites are based on B73 genome reference public assembly: Zm-B73-REFERENCE-NAM-5.0. Physical position of each SNP marker can be found in Supplementary file (SNP_genome_positions.xlsx). **B**, **C**, and **D**) Result of genotyping analysis and identification of BC1-F1 seeds with targeted CO in the first genome editing experiment using gRNA1. **Table B**: Genotyping analysis of edited and control BC1-F1 seeds populations. **Table C**: Identified recombinant BC1-F1 seeds grouped by F1-T0 plants (events). **D**: Genotypes of all BC1-F1 seeds with targeted CO is shown. **E**, **F**, and **G**) Result of genotyping analysis and identification of BC1-F1 seeds with targeted CO in the second genome editing experiment using gRNA2. **Table E**: showing genotyping analysis of edited and control BC1-F1 seeds populations. **Table F**: showing identified recombinant BC1-F1 seeds derived from a single

F1-T0 plant (event). **G**: Genotypes of the BC1-F1 recombinant seeds (subset of 10 out of 175 BC1-F1 examples for each reciprocal crossover is shown).

A.

B.



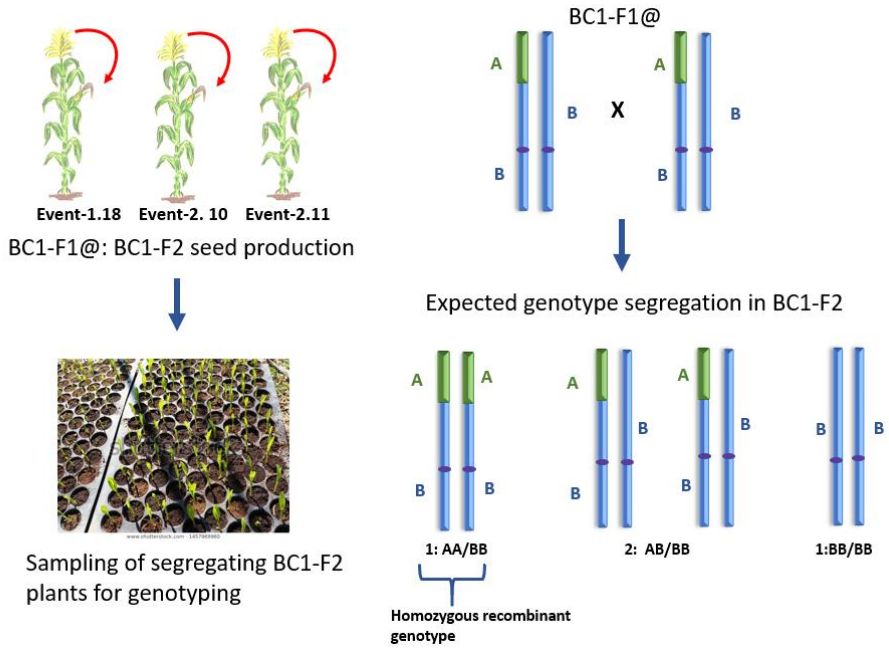
C.

BC1-F1 Event #	Construct	Edited read(%)	Reads without mutation(%)	Editing pattern
Event-1.4	Editing	0	99	TTTCATAGTACACGTCTGGGTAGATTCACAGTACACGTCCGGGTAGATCCTATGGAA
Event-1.5	Editing	61	38	TTTCA----------CAGTACACGTCCGGGTAGATCCTATGGAA
Event-1.7	Editing	55	44	TTTCATAGTACACGTCTG---------ACAGTACACGTCCGGGTAGATCCTATGGAA
Event-1.8	Editing	1	98	TTTCATAGTACACGTCTGGGTAGATTCACAGTACACGTCCGGGTAGATCCTATGGAA
Event-1.9	Editing	0	99	TTTCATAGTACACGTCTGGGTAGATTCACAGTACACGTCCGGGTAGATCCTATGGAA
Event-1.11	Editing	57	42	TTTCATAGTACACGTC----------CGGGTAGATCCTATGGAA
Event-1.14	Editing	59	40	TTTCA----------CAGTACACGTCCGGGTAGATCCTATGGAA
Event-1.18	Editing	58	41	TTTCA----------CAGTACACGTCCGGGTAGATCCTATGGAA
Event-2.2	Editing	51	49	TTTCATAGTACACGTCTGGGTAG---CACAGTACACGTCCGGGTAGATCCTATGGAA
Event-2.6	Editing	46	54	TTTCATAGTACACGTCTGGGTAG---CACAGTACACGTCCGGGTAGATCCTATGGAA
Event-2.10	Editing	0	99	TTTCATAGTACACGTCTGGGTAGATTCACAGTACACGTCCGGGTAGATCCTATGGAA
Event-2.11	Editing	0	99	TTTCATAGTACACGTCTGGGTAGATTCACAGTACACGTCCGGGTAGATCCTATGGAA
Event-3.1	Editing	54	46	TTTCATAGTACACGTCTGG----------CGTCCGGGTAGATCCTATGGAA
Event-4.1	Control	0	99	TTTCATAGTACACGTCTGGGTAGATTCACAGTACACGTCCGGGTAGATCCTATGGAA
Event-4.2	Control	0	99	TTTCATAGTACACGTCTGGGTAGATTCACAGTACACGTCCGGGTAGATCCTATGGAA

Figure 3. Analysis of DNA editing patterns at the site of targeted CO in BC1-F1 plants identified in the first genome editing experiment (gRNA1) with Illumina sequencing. A. The gRNA1 target site showing lbCpf1 cutting pattern: the PAM sequence (red); gRNA target

sequence (italic green); a fragment of the downstream DNA sequence (black). **B.** The graph shows the proportion of edited (blue) and non-edited (magenta) reads in 13 Cpf1 negative BC1-F1 plants with targeted CO and two control plants. BC1-F1 plants selected for further characterization are highlighted with red stars (*) **C.** The table shows specific DNA editing patterns identified in each BC1-F1 plant.

A.



B.

Genotype		M1	M2	M3	M4	M5	M6	M7	M8	gRNA	M9	M10	M11	M12	M13	M14	M16	M17	M33	Cpf1
	1	TT	GG	GG	CC	TT	CC	AA	AA		GG	CC	TT	CC	GG	..	..	AA	CC	NEG
	2	TT	GG	GG	CC	TT	CC	AA	AA		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
BB/BB	3	TT	GG	GG	CC	TT	CC	AA	AA		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
	4	TT	GG	GG	CC	TT	CC	AA	AA		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
	5	TT	GG	GG	CC	TT	CC	AA	AA		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
	6	TT	GG	GG	CC	TT	CC	AA	AA		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
	1	CT	AG	AG	CT	AT	CG	AT	AG		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
	2	CT	AG	AG	CT	AT	CG	AT	AG		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
	3	CT	AG	AG	CT	AT	CG	AT	AG		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
	4	CT	AG	AG	CT	AT	CG	AT	AG		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
	5	CT	AG	AG	CT	AT	CG	AT	AG		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
AB/BB	6	CT	AG	AG	CT	AT	CG	AT	AG		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
	7	CT	AG	AG	CT	AT	CG	AT	AG		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
	8	CT	AG	AG	CT	AT	CG	AT	AG		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
	9	CT	AG	AG	CT	AT	CG	AT	AG		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
	10	CT	AG	AG	CT	AT	CG	AT	AG		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
	1	CC	AA	AA	TT	AA	GG	TT	GG		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
	2	CC	AA	AA	TT	AA	GG	TT	GG		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
AA/BB	3	CC	AA	AA	TT	AA	GG	TT	GG		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG
	4	CC	AA	AA	TT	AA	GG	TT	GG		GG	CC	TT	CC	..	AA	GG	AA	CC	NEG
	5	CC	AA	AA	TT	AA	GG	TT	GG		GG	CC	TT	CC	GG	AA	GG	AA	CC	NEG

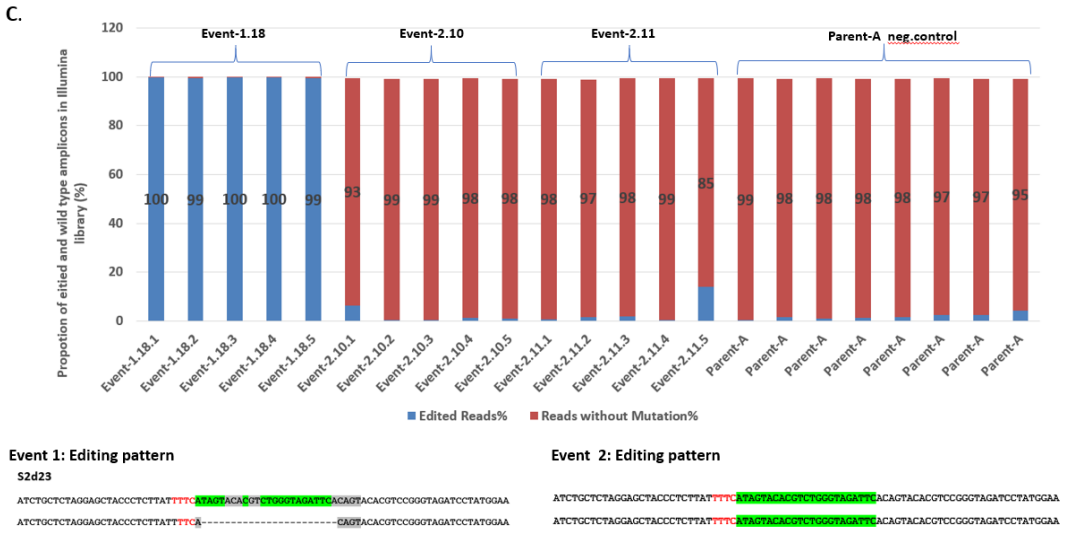


Figure 4. Segregation analysis of BC1-F2 populations and identification of homozygous recombinant plants. A. Schematic diagram of the segregation analysis of the BC1-F2 populations; **B.** The result of the genotyping analysis of the BC1-F2 segregating population. Example of the segregation analysis of the SNP markers in 21 BC1-F2 plants from Event-2.11 is shown; **C**. Confirmation of DNA editing pattern at the site of guided CO in homozygous BC1-F2 plants with Illumina sequencing**.** The proportion of edited (blue) and non-edited (magenta) amplicons in homozygous recombinant plants from 3 evaluated populations is shown on the graph. The pictures under the graph represent alignments of the sequences at the site of targeted CO with the inbred reference. gRNA1 target sequence is highlighted in green. The grey highlights nucleotides microhomology that could contribute to repair of DSBs. S2d23 is abbreviated description of the DNA editing pattern. S2 indicates nucleotide position where deletion starts; d23 indicates the length of the deleted nucleotides.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SNPgenomepositions.xlsx
- Primers.xlsx
- SupplementaryNaturePlants.pdf