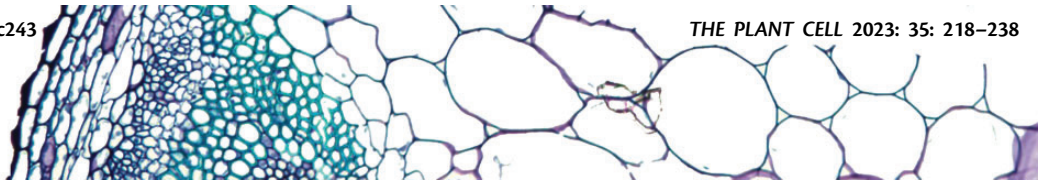
https://doi.org/10.1093/plcell/koac243 THE PLANT CELL 2023: 35: 218–238



Large-Scale Biology

BREEDIT: a multiplex genome editing strategy to improve complex quantitative traits in maize

Christian Damian Lorenzo [1,2], Kevin Debray [1,2], Denia Herwegh [1,2], Ward Develtere [1,2], Lennert Impens [1,2], Dries Schaumont [3], Wout Vandeputte [1,2], Stijn Aesaert [1,2], Griet Coussens [1,2], Yara De Boe [1,2], Kirin Demuynck [1,2], Tom Van Hautegem [1,2], Laurens Pauwels [1,2], Thomas B. Jacobs [1,2], Tom Ruttink [3], Hilde Nelissen [1,2] and Dirk Inzé [1,2,*,†]

1 Center for Plant Systems Biology, VIB, B-9052 Gent, Belgium
2 Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Gent, Belgium
3 Flanders Research Institute for Agriculture, Fisheries and Food (ILVO), B-9820 Merelbeke, Belgium

*Author for correspondence: dirk.inze@psb.ugent.be
†Senior author
These authors contributed equally (C.D.L. and K.D.)

C.D.L. extracted DNA samples, designed and performed the phenotyping experiments, performed crosses, and generated the plant material after the T1 generation. K.D. analyzed the data and created the visualizations. C.D.L., K.D., and D.I. drafted the manuscript. T.V.H., H.N., D.H., and D.I. selected the candidate genes. W.D., D.H., and T.B.J. designed gRNAs and primers. T.R. curated gene models and initiated the HiPlex sequencing. W.D. and T.B.J. designed, constructed, and cloned the SCRIPTs. L.I. and L.P. transformed B104 inbred line to create the EDITOR and SCRIPT lines. D.S., K.D., and T.R. created bioinformatics pipelines to process and analyze HiPlex sequencing reads. S.A. and G.C. performed the plant transformations, performed initial plant crosses, and generated plant material for T0–T1 generation. W.V., Y.d.B., and K.D. assisted with phenotyping, DNA extractions, and greenhouse organization. H.N., T.B.J., T.R., L.P., and D.I. initiated and supervised the project.

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 Author (https://academic.oup.com/plcell) is: Dirk Inzé (dirk.inze@psb.vib-ugent.be).

Abstract

Ensuring food security for an ever-growing global population while adapting to climate change is the main challenge for agriculture in the 21st century. Although new technologies are being applied to tackle this problem, we are approaching a plateau in crop improvement using conventional breeding. Recent advances in CRISPR/Cas9-mediated gene engineering have paved the way to accelerate plant breeding to meet this increasing demand. However, many traits are governed by multiple small-effect genes operating in complex interactive networks. Here, we present the gene discovery pipeline BREEDIT, which combines multiplex genome editing of whole gene families with crossing schemes to improve complex traits such as yield and drought tolerance. We induced gene knockouts in 48 growth-related genes into maize (*Zea mays*) using CRISPR/Cas9 and generated a collection of over 1,000 gene-edited plants. The edited populations displayed (on average) 5%–10% increases in leaf length and up to 20% increases in leaf width compared with the controls. For each gene family, edits in subsets of genes could be associated with enhanced traits, allowing us to reduce the gene space to be considered for trait improvement. BREEDIT could be rapidly applied to generate a diverse collection of mutants to identify promising gene modifications for later use in breeding programs.

Received May 2, 2022. Accepted July 30, 2022. Advance access publication September 6, 2022

IN A NUTSHELL

Background: Plant breeding is the process of generating superior crop varieties by crossing different plants with positive agronomic traits, such as improved yield or drought tolerance, which are regulated by complex gene regulatory networks. So far, conventional breeding has mostly focused on traits improved by natural mutations on particular key genes. Nonetheless, gene redundancy (genes sharing the same function) has limited further improvement because several genes need to be mutated and combined to produce stronger effects. As the threat of climate change is forcing us to develop more productive crops with fewer resources, novel gene combinations need to be discovered to overcome this problem in the future.

Question: Is it possible to achieve improvements in complex traits using genome editing to target whole gene families in crop species? Can these different combinations of gene family edits be further combined to generate even better characteristics?

Findings: Using maize as a model species, we developed a pipeline called BREEDIT to rapidly generate a diverse collection of multiplex gene-edited plants. By using CAS9-expressing maize lines (EDITOR), we simultaneously targeted up to 12 candidate growth-related genes with one SCRIPT construct. This was followed by selfing and/or crossings of plants transformed with same or different SCRIPTs to obtain lines with improved agronomical traits such as increased leaf size and drought tolerance when tested in growth chamber conditions. Furthermore, edits at specific genes were associated with improved characteristics, showing that the number of genes that need to be mutated in order to generate improved plant traits can be reduced. BREEDIT allows to quickly identify these promising gene modifications for later use in breeding programs.

Next steps: Having proved the power of this pipeline to positively reshape plant growth, we now aim to perform field trials with the most promising lines to check for improved performance under field conditions.

Introduction

The production of enough food to feed the increasing global population is becoming quite challenging due to climate change. Extreme temperature ranges, reduced water availability, and the limited use of arable land are all expected to converge and cause a significant drop in crop yields (Zhang and Cai, 2011; Long et al., 2015; Brás et al., 2021). During the past century, conventional breeding has substantially helped adapting crops to local environments, while increasing yields under stress conditions (Nuccio et al., 2018; Snowdon et al., 2021). Genomics-assisted breeding has greatly contributed to the generation of new crop varieties by incorporating haplotype information in breeding programs (Bhat et al., 2021). Nonetheless, we are slowly approaching a plateau in crop improvement using conventional breeding, since gene discovery and the introgression of favorable alleles cannot be implemented quickly enough to cope with the losses caused by environmental stress.

Therefore, innovative strategies need to be implemented to bridge the gap between conventional breeding and the knowledge acquired through plant molecular biology to further improve complex traits such as yield. Crop yield is determined by the complex interactions of the (a)biotic environment with the genetically determined growth and developmental processes that drive the plant's lifecycle (Elias et al., 2016). There are numerous yield-related traits, such as early seedling vigor, root and shoot architecture, biomass allocation, resource use efficiency, senescence, seed filling, and so on. In some cases, such as disease resistance, few causative genes control the expression of the trait (Poland and Rutkoski, 2016). However, for many yield- and growth-related quantitative traits (e.g. organ growth and tolerance to abiotic stress such as drought), numerous small-effect genes contribute to the trait (Mickelbart et al., 2015). Traditionally, yield improvement has been tackled from two distinct angles. Breeding aims at producing genetic combinations with better performance, whereas molecular biology works to understand the mode of action of yield-related genes. These two fields operate at very different scales: breeding recombines chromosomal segments toward a favorable genome constitution, whereas molecular biology only deals with a limited number of genes. In crop breeding programs, phenotypes (e.g. seed yield) are collected from many individuals and multi-year/multi-location field trials. By correlating the phenotypes with the genotypic diversity of individuals, genetic variants associated with the improved trait values can be identified (Rasheed et al., 2017). Using this approach, many quantitative agronomic traits were found to be determined by numerous small-effect loci, with the underlying genomic regions known as quantitative trait loci or QTLs. Such QTLs are generally searched for in segregating mapping populations of recombinant inbred lines (RILs) obtained from two or more parents. A more recent variant of this approach is genome-wide association study (GWAS), in which numerous genome-wide markers are assayed in many diverse genotypes to associate loci with the phenotypic trait (Wang and Qin, 2017). Furthermore, the combination of phenotypic trait data with the availability of

a high number of genomic markers, or even the entire genome sequence, can be used for genomic prediction to increase the predictability of the breeding value of new material (Voss-Fels and Snowdon, 2016). Although these marker-assisted breeding technologies have had a major impact on the accuracy and speed of crop breeding, the genes underlying the QTLs are in many cases unknown. In recent years, technological advances have combined GWAS with molecular -omics phenotypes that go beyond genomic information, allowing molecular networks to start to emerge in molecular breeding (Baute et al., 2015, 2016; Xiao et al., 2016; Miculan et al., 2021).

Over the past four decades, there has been tremendous progress in elucidating the molecular basis of many different plant processes. The use of model organisms such as *Arabidopsis thaliana* and rice (*Oryza sativa*) has been a driving force. A vast amount of research has delivered insights into the molecular pathways steering seed development, root growth, leaf development, plant architecture, tolerance to severe drought stress, cold tolerance, flooding, and many more agronomic traits. Combined, this information reinforced the idea that plant growth and possibly crop yield may be improved by altering the expression of specific (regulatory) genes. Indeed, many reports have shown that positive effects on yield-related traits could be obtained by modifying the expression of individual genes. In Arabidopsis, >60 genes were identified that, when ectopically expressed or downregulated, increase leaf size and in many cases also the sizes of other organs, including seeds (for reviews: Gonzalez et al., 2012; Czesnick and Lenhard, 2015; Vercruysse et al., 2020). Likewise, numerous genes that can be used to improve seed yield and size in rice have been described (Li and Li, 2016). Based on these observations, agro-biotech companies initiated large-scale programs in the beginning of the 21st century to investigate the effects of numerous selected genes on agronomic traits in crops of interest, mainly maize (*Zea mays*) and rice. The conclusion of these studies was that although positive effects were often noticed in the greenhouse and even in field trials, the observed changes were often too small and too much dependent on the genotype and the environment to justify further investments in pursuing this high-throughput screening approach (Paul et al., 2018; Simmons et al., 2021).

Why is it so challenging to translate basic insights in molecular networks and genes into improved crops? In breeding, the phenomenon of expressivity is well-known. Expressivity measures the extent to which a given genotype is expressed at the phenotypic level. The concept of expressivity is best explained by the notion that genes often work in complex networks with many different levels of regulation. Such higher-order regulation is typically exerted on complex and essential processes, such as growth, which need to integrate a panoply of endogenous, genetically determined signals as well as environmental cues. Single-point perturbations of networks often have a limited effect, because other components of the network take over to buffer the system. However, in many cases, the combination of perturbations of a network makes phenotypes much more visible. For example, the pairwise combinations of 13 Arabidopsis growth-related genes (GRGs), each enhancing leaf size on their own when ectopically expressed or mutated, led to additive or synergistic effects on leaf size in >80% of the combinations examined (Vanhaeren et al., 2014, 2017). Moreover, a triple combination of three different mutants of GRGs increased the size of leaves, flowers, seeds, and even roots of Arabidopsis in a spectacular manner (Vanhaeren et al., 2017). Also in maize, albeit with fewer genes, pairwise combinations of specific alleles of growth-enhancing genes result in additive effects (Sun et al., 2017; Liu et al., 2021). This concept is also clearly observed during breeding, when yield traits are most often determined by many small-effect loci that need to work in concert to obtain a maximal output.

Despite the spectacular advances made by systems biology in integrating large data sets, the mechanisms behind the control of plant developmental processes are so complex that predicting which combination of genes would provide the optimal effect on yield remains virtually impossible. Understanding the mode of action might be the best way forward to estimate the combinability of genes (Vanhaeren et al., 2014; Sun et al., 2017). However, even when dealing with a relatively small number of genes, testing all possible pairwise gene combinations remains cumbersome and resource intensive. The investments become even more important when triple or higher-order gene combinations have to be tested, which is necessary to achieve stable yield increases of 10% or higher.

Clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein 9 (Cas9) technology has emerged as a powerful tool for simultaneously multiplex-targeting several GRGs, easily generating genetic variability in a broad set of targets and thus enabling a plethora of combinatorial mutations to be analyzed (Knott and Doudna, 2018; Zhang et al., 2019). Several studies have shown how CRISPR could be used to reshape plant architecture and target complex traits in multiple species like tomato (*Solanum lycopersicum*) (Rodríguez-Leal et al., 2017; Wang et al., 2021), wheat (Li et al., 2020), rice (Meng et al., 2017), and maize (Doll et al., 2019). As a broader application, large-scale CRISPR screens have been carried out in rice (Lu et al., 2017), cotton (*Gossypium hirsutum*) (Ramadan et al., 2021), maize (Liu et al., 2020; Gong et al., 2022), tomato (Jacobs et al., 2017), oilseed rape (*Brassica napus*) (Li et al., 2018), and soybean (*Glycine max*) (Bai et al., 2020).

Here, we designed an experimental approach to bridge the gap between conventional breeding and genetic engineering of multiple genes by combining multiplex CRISPR-mediated genome editing with crossing schemes to observe favorable phenotypes. We named this approach BREEDIT, a contraction of breeding and gene editing, and propose this strategy as a powerful technique to engineer complex traits by knocking out a large number of key players in gene

families and pathways. In just two generations, we generated a list of putative gene knockouts (KOs) required to evoke clear yield-related phenotypes in maize. BREEDIT could therefore be used to rapidly identify a subset of genes involved in the expression of a complex trait and find targets for plant breeding programs.

Results

Development of a CRISPR/Cas9 multiplex genome editing pipeline in maize: general outline

The aim of this study was to develop a flexible pipeline that combines multiplex gene editing and different crossing schemes to generate plants with modified traits (Figure 1). First, candidate GRGs in the target species are selected based on the literature or in-house knowledge; we selected 48 candidate GRGs in maize (Table 1) based on this knowledge combined with information from other model organisms, that is, Arabidopsis and rice (Figure 1A). In particular, negative growth regulators whose inactivation is likely to result in positive effects on growth are suitable GRG candidates. Guide RNAs (gRNAs) targeting these GRGs are then designed and cloned into multiplex gene editing vectors (referred to as SCRIPTs), which are then used to transform *Cas9*-expressing lines (named EDITOR lines), resulting in supertransformed lines that harbor both *Cas9* and a SCRIPT containing 12 gRNAs (Figure 1B and Supplemental Figure S1).

The BREEDIT pipeline then uses highly multiplex (HiPlex) amplicon sequencing combined with the stack mapping anchor points (SMAP) haplotype-window bioinformatics workflow to routinely monitor gene edits at gRNA cutting sites (Schaumont et al., 2022). Amplicon sequencing at great depths allows haplotype sequences and their respective frequencies to be determined. Both types of information can be used to assess the effects of mutations on the encoded protein function or activity and to assign a genotype to the plant for a specific locus. Per sample and per locus, the length difference between a mutated haplotype and the reference haplotype is used to classify the mutated haplotypes into two categories: haplotype$_{KO}$, which corresponds to haplotypes containing out-of-frame insertions or deletions (indels), leading to a gene KO and a nonfunctional protein; and haplotype$_{REF}$, including haplotypes with only single-nucleotide polymorphisms (SNPs) outside the cutting site or in-frame indels thought to have less of an impact on the translated protein that may still behave as the reference protein. In CRISPR/Cas9 experiments, one plant may contain more haplotypes than its ploidy level because it contains mosaic tissues due to the initial (T0) or ongoing (T1, T2) Cas9 activity, thus complicating the genotyping. To interpret complex haplotype constitutions, the relative fraction of all haplotype$_{KO}$ is summed per locus per sample. The resulting aggregation is interpreted as a gene loss-of-function (LOF) dosage, which is further divided into three categories: LOF$_{0/2}$ (neither of the two chromosomes is affected by a set of haplotype$_{KO}$), LOF$_{1/2}$ (one of the two chromosomes is affected by a set of haplotype$_{KO}$), and LOF$_{2/2}$ (both chromosomes are affected by a set of haplotype$_{KO}$). The three dosage categories are used in genotype-to-phenotype associations.

After selecting transgenic lines, T0 lines are genotyped, and the T0 plants with the highest number of gene KOs (either partial [LOF$_{1/2}$] or complete [LOF$_{2/2}$]) are crossed to obtain material for phenotyping (Figure 1, C and D). Different crosses can be performed to maximize the number of edited genes and to fix combinations of gene edits. Self-crosses serve to fix edits in parallel to maximize the phenotypic readout, while backcrosses to the original line provide heterozygous lines that can later be self-crossed and phenotyped in the T2 generation. Additional specific crosses can be performed to further enrich edit diversity. Plants harboring the same SCRIPT but containing edits at different genes from that SCRIPT can be crossed to increase the number of gene edits (up to 12) in the corresponding gene family or pathway. Such crosses are referred to as intra-script crosses. Furthermore, plants transformed with different SCRIPTs can be crossed to maximally combine mutants in genes covered by different families or pathways. These crosses are referred to as inter-script crosses. Our pipeline was also designed to keep Cas9 active in all subsequent generations, a strategy that will produce further gene edits in wild-type alleles with a very low risk of generating off-targeted mutagenesis (Lee et al., 2019; Bessoltane et al., 2022). These new trans-generational edits (Impens et al., 2022) are expected to accumulate and possibly saturate all targeted loci, resulting in a large collection of higher-order mutants (up to 24 gene edits when two SCRIPTs are combined) in different segregating states (i.e. LOF$_{0/2}$, LOF$_{1/2}$, or LOF$_{2/2}$).

Because several plants are generated with the BREEDIT approach, easy-to-measure quantitative traits are used to maximize the throughput of the phenotyping steps. Despite the high number of plants generated, each individual likely has a unique genotypic profile given the many combinations of indels and dosage that can happen in a set of 12 genes or more. Therefore, repetitions of the same genotypic combinations cannot be used for statistical analysis in BREEDIT. The effects of combinations of gene edits on traits are better appraised at the population level, although the specific causative gene combination cannot be deduced. However, the effect of a single gene on a trait can still be evaluated considering that multiple observations of a single-gene KO would conceal the putative noise brought by mutations in other genes. The framework for phenotyping experiments consists of several (minimum of two) independent trials to test the performance of independent mutated populations compared with the EDITOR line. Single-gene associations to a trait are then conducted per experiment per population. The number of times a gene KO is significantly associated with a trait across different independent populations and experiments is a measure of the importance of that gene in the expression of the trait. At the end of the BREEDIT

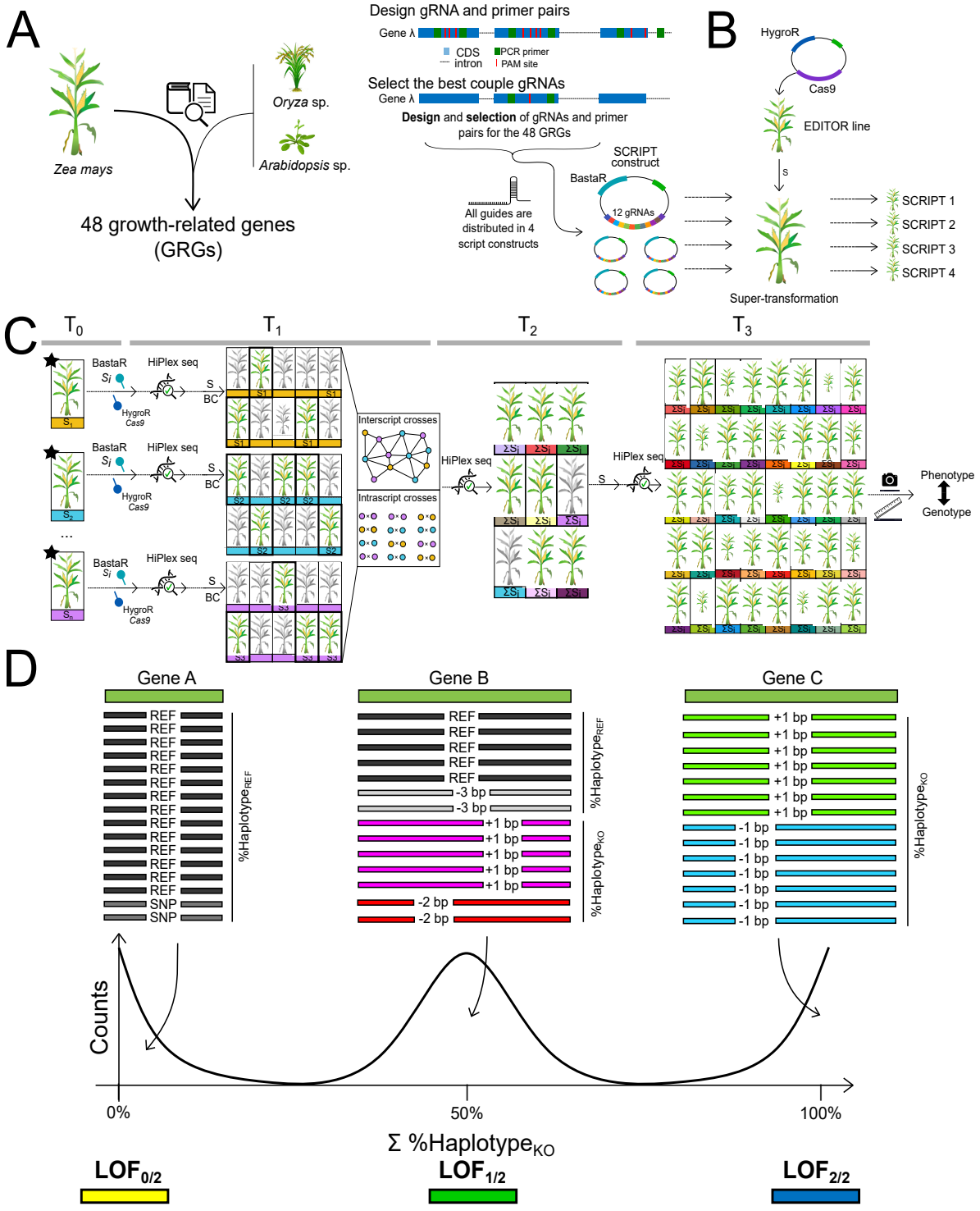


Figure 1 The multiplex gene editing strategy of BREEDIT. A, Selection of GRGs based on published and in-house research performed in Arabidopsis, rice, or maize. B, After gene selection, gRNAs with NGG protospacer adjacent motif (PAM) sites are selected for each gene, and PCR primer pairs are designed to resequence gRNA target sites and flanking regions via HiPlex amplicon sequencing. For each gene, the best set of gRNAs and flanking primer pairs is selected. Twelve gRNAs are cloned in multiplex gene editing vectors named SCRIPTs. Next, the SCRIPT constructs are transformed in a Cas9-expressing maize line named EDITOR. C, Vigorous T0 plants containing the SCRIPT (BASTA resistant) and the Cas9 EDITOR construct (hygromycin resistant) are further genotyped using HiPlex amplicon sequencing. Based on the genotypes, plants are selected for crossing with B104 (BC: backcross), with plants with complementary mutations caused by the same SCRIPT (intra-script crosses), or with plants containing a different SCRIPT and therefore mutations in genes from a different family or pathway (inter-script crosses). These crosses aim at maximizing the mutation landscape and diversity. Finally, self-crosses (S) of lines generate a segregating progeny for high-throughput phenotyping of selected traits, which later can be associated with (combinations of) genes. D, From continuous read depth to discrete LOF genotypic classes. Sequencing reads are mapped to the B104 reference loci. Two read categories are derived, namely haplotype$_{REF}$ and haplotype$_{KO}$.

(continued)

pipeline, genes can be ranked to delineate a minimal set of candidate genes with maximal effects on trait expression, thus reducing the gene space to be considered for further research.

Table 1 List of the 48 GRGs targeted by different SCRIPTs

SCRIPT	Position	Gene	Gene family/pathway	B73 V3 gene ID	References
1	1	*ZmGa2ox2*	GA2-oxidases	GRMZM2G006964	(Huang et al., 2015; Li et al., 2021)
1	2	*ZmGa2ox4*	GA2-oxidases	GRMZM2G153359	
1	3	*ZmGa2ox5*	GA2-oxidases	GRMZM2G176963	
1	4	*ZmGa2ox7*	GA2-oxidases	GRMZM2G427618	
1	5	*ZmGa2ox8*	GA2-oxidases	GRMZM2G155686	
1	6	*ZmGa2ox9*	GA2-oxidases	GRMZM2G152354	
1	7	*ZmGa2ox13*	GA2-oxidases	GRMZM2G031432	
1	8	*D8*	DELLA/GRAS family	GRMZM2G144744	(Winkler and Freeling, 1994; Lawit et al., 2010)
1	9	*D9*	DELLA/GRAS family	GRMZM2G024973	
1	10	*ZmSLRL1-1*	DELLA/GRAS family	GRMZM5G826526	(Ikeda et al., 2001; Itoh et al., 2005; Liu et al., 2021)
1	11	*ZmSLRL2*	DELLA/GRAS family	GRMZM5G874545	
1	12	*ZmSPY*	GA signalling	GRMZM2G357804	(Qin et al., 2011)
2	1	*ZmCKX-2*	Cytokinin oxidases	GRMZM2G050997	(Ashikari et al., 2005; Bartrina et al., 2011)
2	2	*ZmCKX-3*	Cytokinin oxidases	GRMZM2G167220	
2	3	*ZmCKX-4*	Cytokinin oxidases	GRMZM5G817173	
2	4	*ZmCKX-4B*	Cytokinin oxidases	GRMZM2G024476	
2	5	*ZmCKX-5*	Cytokinin oxidases	GRMZM2G325612	
2	6	*ZmCKX-6*	Cytokinin oxidases	GRMZM2G404443	
2	7	*ZmCKX-7*	Cytokinin oxidases	GRMZM2G134634	
2	8	*ZmCKX-8*	Cytokinin oxidases	GRMZM2G162048	
2	9	*ZmCKX-9*	Cytokinin oxidases	GRMZM2G303707	
2	10	*ZmCKX-10*	Cytokinin oxidases	GRMZM2G348452	
2	11	*ZmCKX-11*	Cytokinin oxidases	GRMZM2G122340	
2	12	*ZmCKX-12*	Cytokinin oxidases	GRMZM2G008792	
3	1	*ZmKRP1;1*	ICK/KRP cyclin-dependent kinase	GRMZM2G101613	(Cheng et al., 2013; Cao et al., 2018)
3	2	*ZmKRP1;2*	ICK/KRP cyclin-dependent kinase	GRMZM2G084570	
3	3	*ZmKRP1;3*	ICK/KRP cyclin-dependent kinase	GRMZM2G343769	
3	4	*ZmKRP3*	ICK/KRP cyclin-dependent kinase	GRMZM2G154414	
3	5	*ZmKRP4;2A*	ICK/KRP cyclin-dependent kinase	GRMZM2G037926	
3	6	*ZmKRP4;2B*	ICK/KRP cyclin-dependent kinase	GRMZM2G116885	
3	7	*ZmKRP5;1*	ICK/KRP cyclin-dependent kinase	GRMZM2G358931	
3	8	*ZmKRP5;2*	ICK/KRP cyclin-dependent kinase	GRMZM2G157510	
3	9	*ZmPP2C-A9*	ABA signal transduction	GRMZM2G134628	(He et al., 2019)
3	10	*ZmPP2C-A11*	ABA signal transduction	GRMZM2G159811	
3	11	*ZmHB124B*	Homeobox transcription factor family	GRMZM2G023291	(McConnell et al., 2001)
3	12	*ZmHB124C*	Homeobox transcription factor family	GRMZM2G178102	
4	1	*ZmTCP3*	TCP-CIN clade	GRMZM2G115516	(Koyama et al., 2017; Sarvepalli and Nath, 2018; Lan and Qin, 2020)
4	2	*ZmTCP8*	TCP-CIN clade	GRMZM2G020805	
4	3	*ZmTCP9*	TCP-CIN clade	GRMZM2G589470	
4	4	*ZmTCP10*	TCP-CIN clade	GRMZM2G166946	
4	5	*ZmTCP22*	TCP-CIN clade	GRMZM2G120151	
4	6	*ZmTCP25*	TCP-CIN clade	GRMZM2G035944	
4	7	*ZmTCP42*	TCP-CIN clade	GRMZM2G180568	
4	8	*ZmGRF4*	Growth-regulating factor clade	GRMZM2G004619	(Nelissen et al., 2012; Voorend et al., 2016; Liebsch and Palatnik, 2020)
4	9	*ZmGRF10*	Growth-regulating factor clade	GRMZM2G096709	
4	10	*ZmGRF17*	Growth-regulating factor clade	GRMZM2G124566	
4	11	*ZmPHD8*	SET domain transcription factor	GRMZM2G409224	–
4	12	*ZmBPC6*	GAGA-binding protein	GRMZM2G118690	(Gong et al., 2018)

Figure 1 Continued

Haplotype$_{REF}$ corresponds to the aggregated fraction of reads containing only SNPs, in-frame indels, or the reference haplotype. Haplotype$_{KO}$ refers to the aggregated fraction of reads with out-of-frame indels. A tri-modal distribution is expected for haplotype$_{KO}$, with local maxima around 0%, 50%, and 100%, each corresponding to a fraction of the genome being edited at the locus. Haplotype$_{KO}$ is therefore divided into three classes of LOF: LOF$_{0/2}$ (the genome is not edited with out-of-frame indels, i.e. 0 chromosome out of two in a diploid organism), LOF$_{1/2}$ (half of the genome is edited with out-of-frame indels, i.e. one chromosome out of two in a diploid organism), LOF$_{2/2}$ (the whole genome is edited with out-of-frame indels, i.e. two chromosomes out of two in a diploid organism).

Applying the BREEDIT strategy

To test the BREEDIT strategy, we selected 48 maize GRGs (Table 1) with potential positive effects on growth when mutated, individually or in combination, as described above. The gRNAs targeting the 48 genes were distributed over SCRIPT 1 to SCRIPT 4, and were grouped per gene family when possible. This distribution primarily aimed to simultaneously KO multiple members of the same gene family/pathway to overcome the potential functional redundancy of paralogs. In addition, grouping by family can generate segregating mutants with a range of gene KOs, which may help to untangle complex relationships in gene regulatory networks that might be overlooked when only single or double mutants are considered. Additionally, the chromosomal positions of the GRGs were taken into consideration to spread the distribution of genes belonging to a same SCRIPT over chromosomes when possible (Supplemental Figure S2). The 12 genes targeted in SCRIPT 1 are major players in gibberellin catabolism and signaling; those targeted in SCRIPT 2 are maize putative orthologs of genes encoding cytokinin oxidases (*CKXs*), which are key regulators of cytokinin catabolism. SCRIPT 3 contains gRNAs for eight *ICK/KRP* genes encoding the family of inhibitors of cyclin-dependent kinase/Kip-related proteins, as well as four genes expected to encode negative regulators of growth under drought conditions: two maize *PP2C* putative orthologs (*ZmPP2Cs*) and two *HOMEOBOX*-type genes (*HB124B* and *HB124C*), which are orthologs of the Arabidopsis genes *PHABULOSA* and *PHAVOLUTA* (McConnell et al., 2001). Finally, SCRIPT 4 contains gRNAs for seven putative orthologs of class II *CINCINNATA-TEOSINTE BRANCHED 1/CYCLOIDEA/PROLIFERATING CELL FACTOR* (*CIN-TCP*) and three members of the *GROWTH REGULATING FACTORS* (*GRF*) genes, which are major regulators of cell division, leaf shape, and leaf size determination. Additionally, gRNAs targeting an ortholog of the GAGA-binding protein-encoding gene *BASIC PENTACYSTEINE 6* (*ZmBPC6*) and a gene encoding a plant homeodomain (PHD)-finger protein (*ZmPHD8*) were included in SCRIPT 4.

Generation of highly edited maize populations for all SCRIPTs

We developed a set of three independent homozygous EDITOR lines that constitutively express the Cas9 protein in the maize inbred line B104 background (Supplemental Figure S3) to execute editing at loci targeted by arrays of 12 gRNAs expressed from the SCRIPT vector. EDITOR 1 and EDITOR 3 were supertransformed with SCRIPT 1 for a preliminary evaluation of gene editing. After transformation, the EDITOR 1 and EDITOR 3 supertransformed populations showed similar editing profiles (Supplemental Figure S4). At T0, 6 out of the 12 targeted genes showed $LOF_{1/2}$ or $LOF_{2/2}$ in both EDITOR lines, and the number of mutant alleles at each locus was comparable between both EDITOR backgrounds. The same gRNAs were active in both EDITOR backgrounds, but 4 genes out of the 6 that were commonly edited in both EDITOR backgrounds showed $LOF_{2/2}$ in EDITOR 1, whereas two genes showed $LOF_{2/2}$ in EDITOR 3 (Supplemental Figure S4).

We selected EDITOR 1 as the genetic background for further experiments and supertransformed this line with the three remaining scripts. Like for SCRIPT 1, we monitored gene edits in T0 plants and all subsequent generations using HiPlex amplicon sequencing. Indels in haplotype sequences ranged from −90 to +92 bp. Insertions of one nucleotide (+1 bp) were the most commonly represented type of mutation, but overall, more deletions were present than insertions (Figure 2A). The largest insertions showed sequence similarity to genomic fragments located up to 1 kb upstream or downstream of the expected cutting site. At T0, we detected $haplotype_{KO}$ in 11, 12, 8, and 12 out of the 12 target sites for SCRIPT 1, 2, 3, and 4, respectively (Figure 2B). Across all T0 SCRIPT populations, a large diversity of haplotypes (109 haplotypes with in-frame indels and 407 haplotypes with out-of-frame indels) could be identified (Supplemental Figure S5). Some $haplotype_{KO}$ were initially not detected at T0 but appeared in T1 populations (Figure 2B) of both intra-script and inter-script crosses, revealing either ongoing gene editing in subsequent generations or overlooked edits due to mosaic tissues in T0. Overall, from T0 to T2, mutations could be found in all 48 targeted genes except one (*SPY* in SCRIPT 1). We focused on $haplotype_{KO}$ and observed a diversity of $haplotype_{KO}$ combinations per locus per sample (mono-, bi-, and multi-allelic) in the T0 to T2 samples, which were all expected to lead to a gene LOF, either partial ($LOF_{1/2}$) or complete ($LOF_{2/2}$) (Figure 2C). We observed a typical tri-modal distribution for the aggregated fraction of $haplotype_{KO}$ that could be roughly divided into three areas with higher counts, each corresponding to a discrete genotypic class ($LOF_{0/2}$, $LOF_{1/2}$, and $LOF_{2/2}$; Supplemental Figure S6).

From haplotype frequencies to genotypic information

We used the aggregated fraction of $haplotype_{KO}$ in sequencing reads as a proxy to characterize partial ($LOF_{1/2}$) and complete ($LOF_{2/2}$) gene KOs (Figure 3). Our approach for the detection of gene edits using HiPlex amplicon sequencing combined with SMAP haplotype-window analyses successfully captured haplotype sequences in 96% of the cases, encouraging us to use this technique to monitor edits in the offspring (Figure 3A). At T0, 73% (35/48) of the target loci showed $LOF_{1/2}$ or $LOF_{2/2}$, with SCRIPT 1 and SCRIPT 3 showing worse performance than SCRIPT 2 and SCRIPT 4 (Figure 3B). At T1, of the 13 remaining genes not edited at T0, 12 (92%) were de novo edited. No $haplotype_{KO}$ was observed at the last remaining non-edited locus (*SPY*) at T2. Also, all the transgenerational de novo edits were only heterozygous mono-allelic mutations (Figure 3B). Considering both T0 and T1 materials, we observed plants stacking up to nine $LOF_{1/2}$ or $LOF_{2/2}$ gene KOs in both SCRIPT 1 and SCRIPT 3 and 11 $LOF_{1/2}$ or

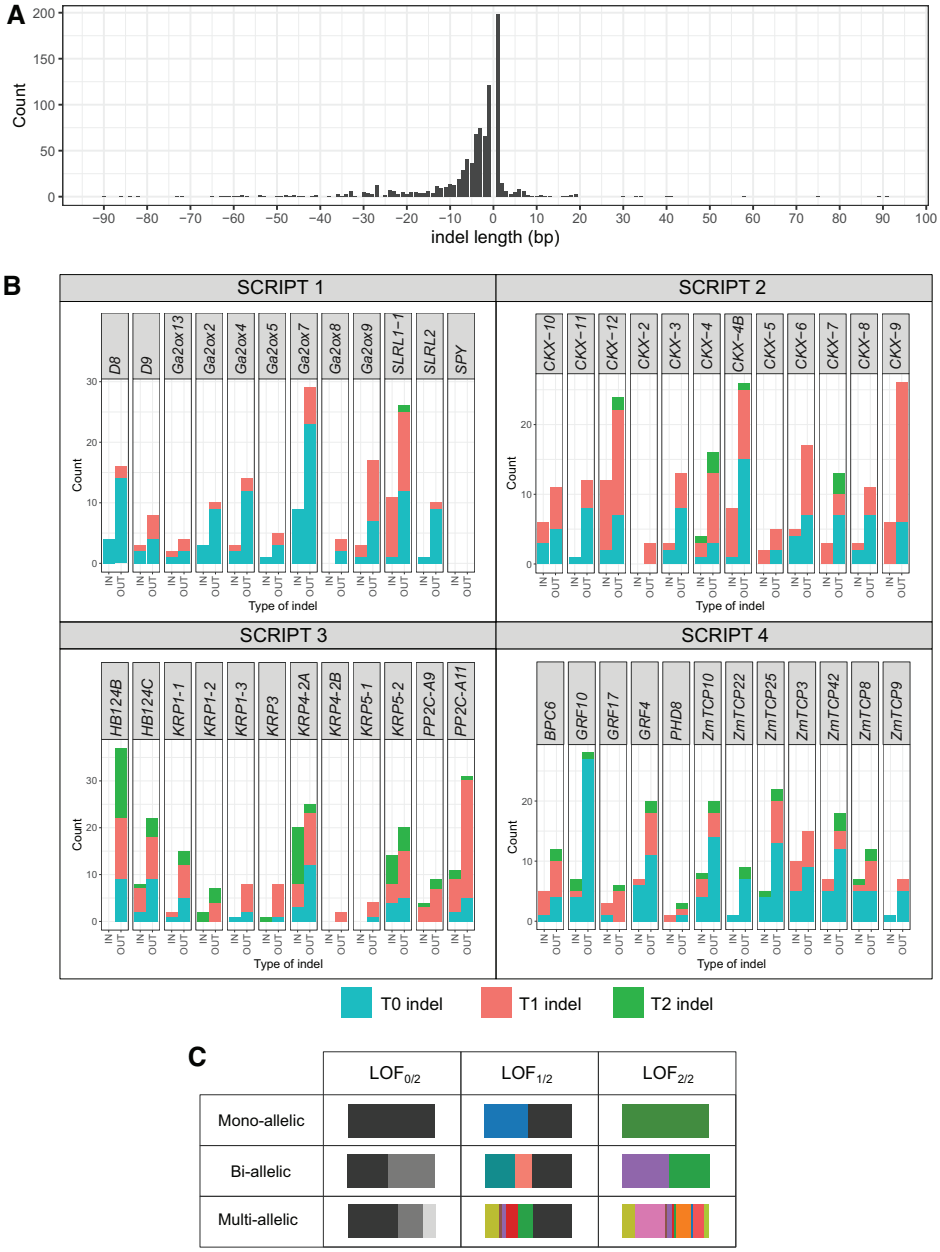


Figure 2 Diversity of mutated haplotypes obtained after CRISPR/Cas9 genome editing. A, Distribution of indel length. B, Number of different haplotypes with indels first observed at T0, T1, and T2. Any haplotype with indels with >1% relative frequency in the sequencing reads per locus per sample is included. IN, in-frame indel; OUT, out-of-frame indel. C, Different haplotype combinations in plants can all lead to a gene LOF, either partial ($LOF_{1/2}$) or complete ($LOF_{2/2}$). Each colored horizontal stacked bar corresponds to a different $haplotype_{KO}$. The bar's length is proportional to the fraction of sequencing reads per locus containing the $haplotype_{KO}$. The black fraction corresponds to the aggregation of alleles assigned to the wild-type haplotype ($haplotype_{REF}$). For an overview of the different $haplotype_{KO}$ found in T0 plants harboring the different SCRIPTs, see Supplemental Figure S5.

$LOF_{2/2}$ gene KOs in SCRIPT 2 and SCRIPT 4 (Figure 3C). Because of sterility issues, it was difficult to generate progeny of SCRIPT 1 by crossing, resulting in the low numbers of T2 plants for that SCRIPT (Figure 3C). We also studied progeny resulting from inter-script crosses involving two SCRIPTs (2 × 12 target loci) and observed that, on average, 40% of the loci showing edits in the progeny presented transgenerational editing patterns (result of ongoing Cas9 activity through a second generation) and 25% were completely de novo edited, meaning that edits at these loci were not observed in the parental lines (Supplemental Figure S7). Per locus across all populations, an average of 7% of the progeny were affected by transgenerational edits inducing $LOF_{1/2}$ at the target sites.

In conclusion, the approach of supertransforming EDITOR lines with SCRIPT constructs generated a high frequency of

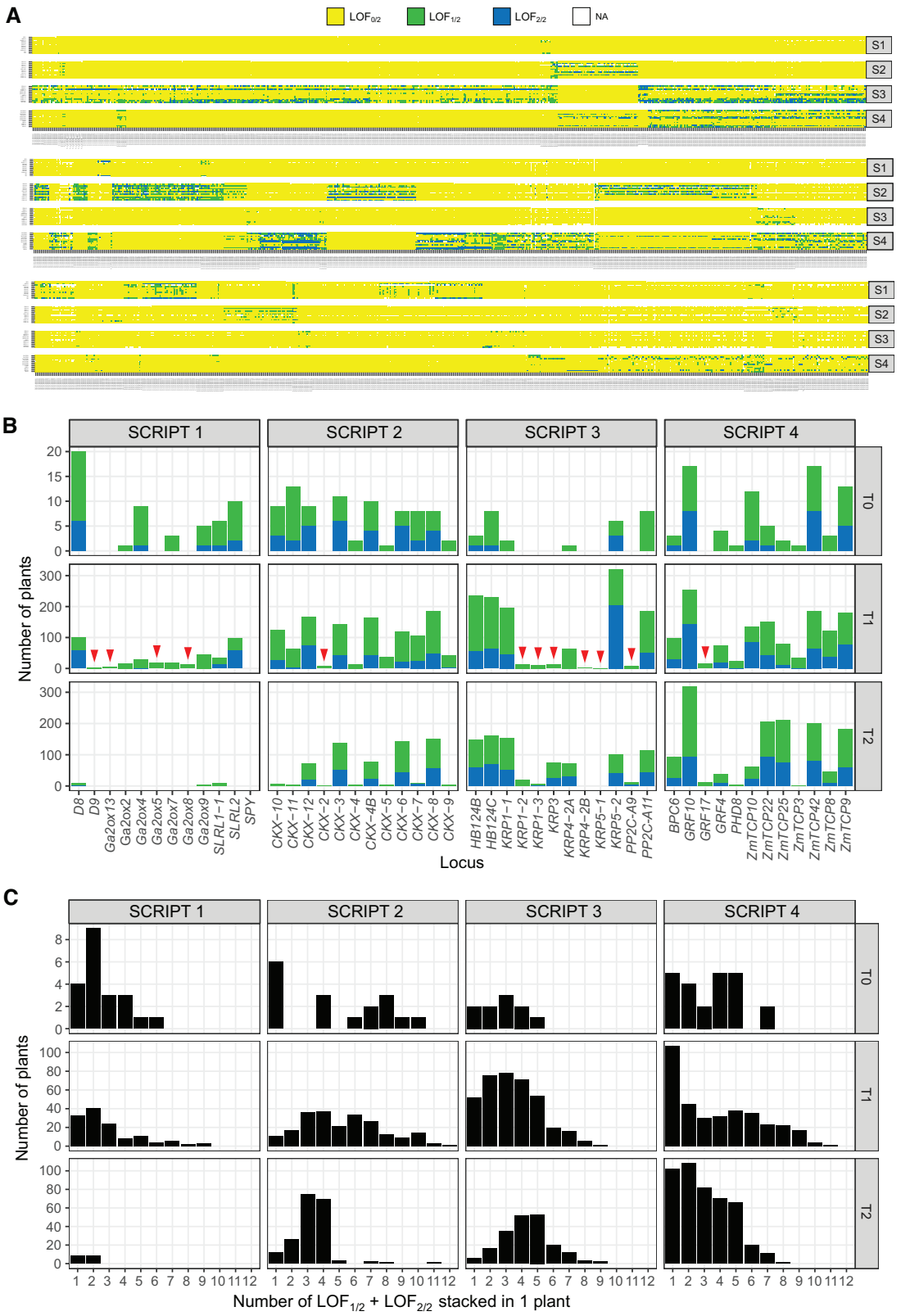


Figure 3 Distribution of LOF in genes across the entire set of samples. Only haplotype$_{KO}$ were considered for genotype calling. The fractions of reads containing haplotype$_{KO}$ were summed per sample per locus. A, Overview of the classes $LOF_{0/2}$, $LOF_{1/2}$, or $LOF_{2/2}$ obtained in the entire sample set for the four SCRIPTS (S1–S4). Samples are on the *x*-axis and are distributed over three rows. B, Distributions of $LOF_{1/2}$ and $LOF_{2/2}$ across the four SCRIPTs throughout the generations. The top, middle, and bottom panels show T0, T1, and T2 plants, respectively. Triangles indicate new LOFs that appeared at T1. C, Stacking LOFs at multiple genes within plants.

heritable edits in the T0 generation and additional transgenerational edits in the T1 and T2 generations.

T1 single-SCRIPT multiple-edited populations display phenotypic variability in seedling growth-related traits

After we generated the single-SCRIPT populations of edited plants, we studied the effects of multiple gene edits on plant growth by phenotyping T1 maize seedlings derived from T0 selfings of each SCRIPT at the vegetative 3 (V3) stage. To facilitate high-throughput phenotyping of several populations, we scored easy-to-measure parameters such as the final leaf length and width of leaf 3 (FLL3 and FLW3, respectively) and integrative parameters such as the fresh weight (FW), dry weight (DW), and moisture content of plants grown under well-watered (WW) and water-deficient (WD) conditions. We scored populations derived from independent transgenic events to analyze the effects of combinations of LOF dosages resulting from different haplotype$_{KO}$ on trait expression (Figure 4, gradient of edits displayed in orange). Detailed information about the different populations that were phenotyped is provided in Supplemental Table S1.

SCRIPT 1 plants were tested in two independent WW experiments (WW001 and WW008) (Figure 4, A and B) and displayed conspicuous phenotypes such as a slender shoot architecture (Figure 4C) with longer and narrower leaves (Figure 4, A, B, and E) compared with the EDITOR 1 control. The most conspicuous phenotypes were observed in population P013, which included individuals with a partial or complete LOF of 11 out of 12 genes (Figure 4, A–E). Additionally, some SCRIPT 1 plants displayed abnormal tassel development with a lack of florets or pollen and the formation of silks in the anthers (Supplemental Figure S8), leading to male sterility.

For SCRIPT 2 and SCRIPT 3, when tested in experiment WW001, significant increases of approximately 5% compared with the controls were detected only for FLL3 and only in one of the two populations of each group (P108 for SCRIPT 2 and P033 for SCRIPT 3), while FLW3 remained unaffected in SCRIPT 2 populations or decreased for both populations of SCRIPT 3 (Supplemental Figure S9, B and C). Because the genes targeted in SCRIPT 2 are involved in cytokinin metabolism, which was previously implicated in drought tolerance (Rida et al., 2021), and some of the genes targeted in SCRIPT 3 are drought responsive (Li et al., 2016; Hai et al., 2020), we phenotyped these populations under WD conditions (Supplemental Figure S9, B and C). Under WD, the SCRIPT 2 populations showed enhanced growth (Supplemental Figure S10A), as reflected by a significant increase in FLL3, FLW3, FW, and DW compared with the EDITOR 1 control (Supplemental Figures S9B and S10, A–C). For SCRIPT 3, all tested populations displayed enhanced growth traits (Supplemental Figure S11), but only significant increases in FLL3 compared with the EDITOR 1 control were observed (Supplemental Figure S9C). Moreover, population P034 presented a significant increase in FW compared with the EDITOR 1 control (Supplemental Figures S9C and S11).

Changes in leaf morphology were also observed for SCRIPT 4 plants (TCP, GRF family genes). Individuals that segregated LOF dosages in all 12 and 9 out of 12 genes were observed in populations P059 and P060, respectively (Figure 4, F and G, gradient of edits in orange; Supplemental Table S1). Both populations presented significantly longer FLL3 (Figure 4F) alongside a $>15\%$ increase in FLW3 compared with the EDITOR 1 control (Figure 4, G–I). An increase in FLL3 was not detected in population P054, P079, or P130 (Figure 4F and Supplemental Table S1), but a significant rise in FLW3 was detected in all populations (Figure 4G).

Crossing plants with different SCRIPTs allows phenotypes to be combined in T2 plants

After focusing on single-SCRIPT populations, we phenotyped inter-script populations that stacked edits in genes from different SCRIPTs after crossing. For this analysis, T0 plants with different scripts were crossed (inter-script crosses) and the resulting T1 plants were self-crossed. Of all the different combinations, we phenotyped two T2 inter-script populations that presented different profiles of edits in crosses between SCRIPT 2 $\times$ SCRIPT 4 (P148 and P152) and SCRIPT 3 $\times$ SCRIPT 4 (P157 and P158) under WD conditions. For both populations of SCRIPT 2 $\times$ SCRIPT 4 and SCRIPT 3 $\times$ SCRIPT 4, we detected a significant increase in FLW3 (Supplemental Figure S12 and Supplemental Table S1), a phenotype observed in single-SCRIPT 4 T1 lines. For the other traits, distinct differences were observed in each population. P148 displayed an increase in FLL3, whereas P152 showed a decrease in FLL3 and significant increases in FW and moisture content compared with the EDITOR 1 control (Supplemental Figure S12A). Both P157 and P158 displayed significant increases in moisture content, and P158 displayed reduced DW, compared with the EDITOR 1 control (Supplemental Figure 12B).

Genotype-to-phenotype associations and reducing the gene space

After performing phenotypic evaluation of all SCRIPT populations, we aimed to identify the possible major-effect causative genes for the observed phenotypes. Because each individual phenotyping experiment did not allow for sufficient replication of LOF dosage combinations, we performed genotype-to-phenotype associations at the single-gene level. For each gene and trait, we compared the three classes of LOF dosages ($LOF_{0/2}$, $LOF_{1/2}$, and $LOF_{2/2}$) with the EDITOR 1 control. Such single-gene analyses were carried out separately for all experiments conducted under WW and WD, representing in total a collection of $>1{,}000$ plants that included data on selfed, inter-, and intra-script crossed lines.

Following this approach, we detected a subset of genes for each gene family that could be at least partially responsible for the observed phenotypes (Figure 5). In SCRIPT 1,

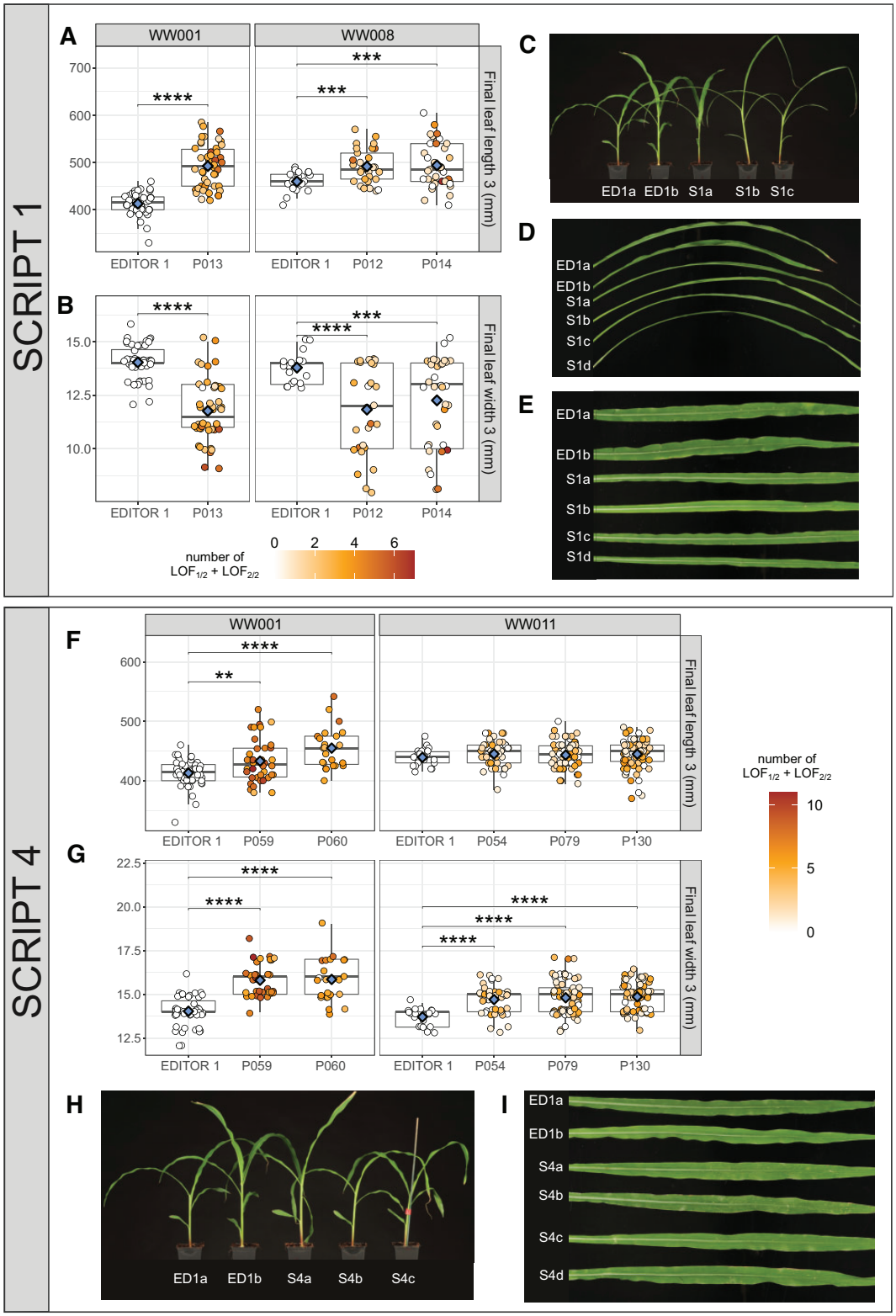


Figure 4 Phenotypes observed in multiple gene-edited populations of SCRIPT 1 and SCRIPT 4. A, B and F, G, Measurements of final length of leaf 3 (FLL3) (A, F) and final leaf width (FLW3) (B, G) of gene-edited SCRIPT 1 (A, B) and SCRIPT 4 (F, G) individuals compared with the EDITOR 1 background control. For each SCRIPT, data correspond to independent multiple gene-edited populations assayed on two different independent experiments under WW conditions. On the distributions, each dot represents one individual and is colored according to the amount of partial ($LOF_{1/2}$) and complete ($LOF_{2/2}$) LOF observed in that individual. The more orange, the higher the LOF in the individual. A pairwise Student's *t*-test was conducted between the EDITOR 1 control and mutated populations. Significant differences are displayed with *P*-values summarized as follows: $^{**}P < 0.01$, $^{***}P < 1e{-}3$, $^{****}P < 1e{-}4$. Box plot borders represent the first and third quartiles, the middle line represents the median, and whiskers represent the maximal and minimum values. Diamonds indicate the means of each distribution. C-E and H, I, Photographs of general plant architecture (C for SCRIPT 1 and H for SCRIPT 4) and final leaf 3 (D for SCRIPT 1 and I for SCRIPT 4) compared with the EDITOR 1 (ED1) background.

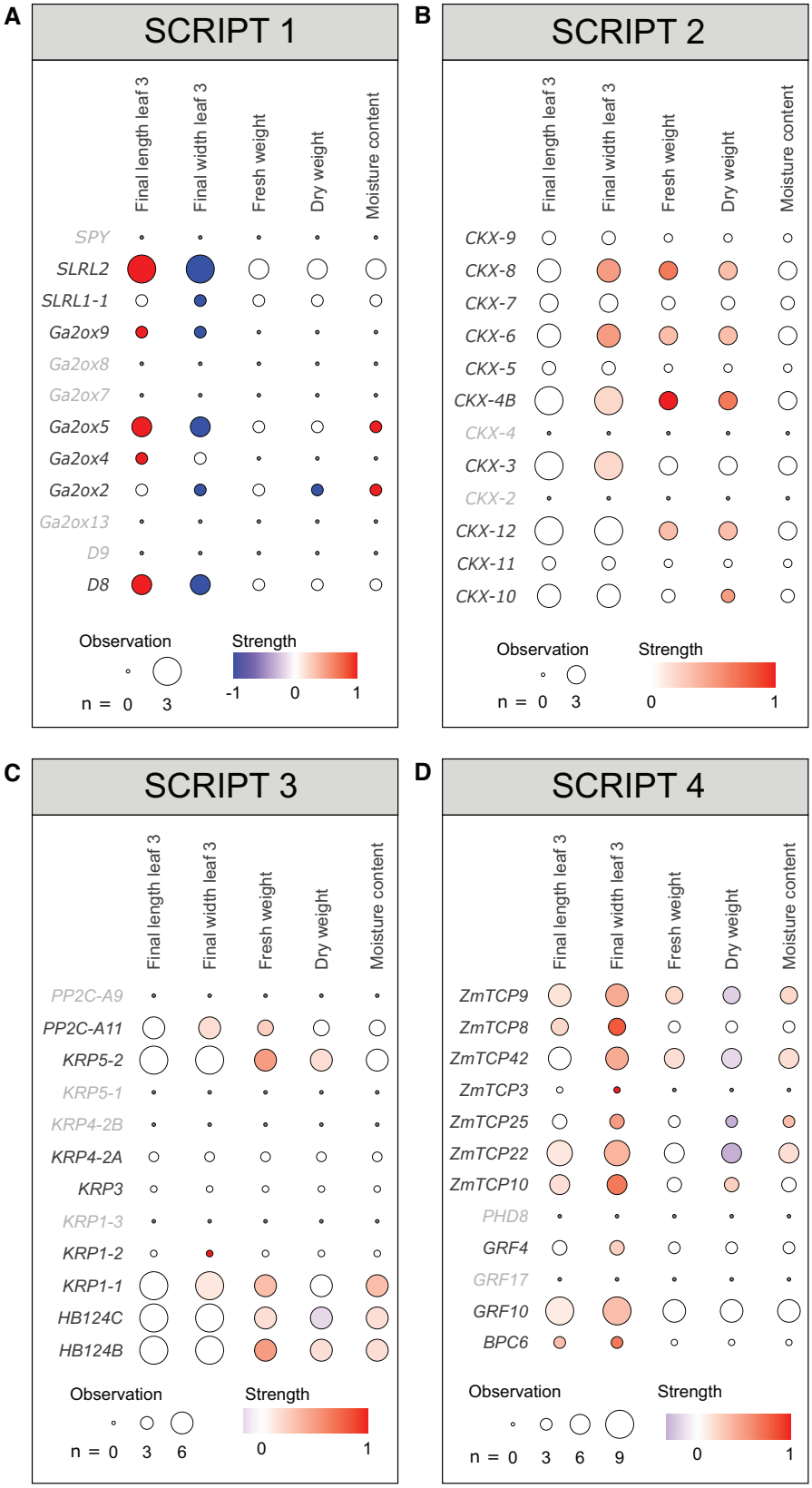


Figure 5 Aggregated association analysis of single-gene LOF and traits. Summaries of single-gene associations to traits are represented for SCRIPT 1 (A), SCRIPT 2 (B), SCRIPT 3 (C), and SCRIPT 4 (D). Single-gene associations were performed per population, in each phenotypic experiment and for all measured traits. Results are summarized per gene, per trait with two indices. (1) Observation: the number of times a given gene has been observed in a situation with sufficient genotypic and phenotypic data across populations and experiments. An observation with sufficient data corresponds to a situation where a gene displays at least one LOF group between $LOF_{1/2}$ and $LOF_{2/2}$ represented by at least six individuals with

(continued)

increases in FLL3 and decreases in FLW3 were associated with edits in DELLA putative orthologs *D8* and *ZmSLR2*, as well as *ZmGa2ox5* (Figure 5A). For SCRIPT 2, edits in *ZmCKX4B*, *ZmCKX6*, and *ZmCKX8* were related to changes in FW, FLW3, and DW (Figure 5B). In SCRIPT 3, LOFs in *ZmKRP5-2* and *ZmPP2C-A11* were associated with increases in FW and DW, while LOFs in Zm*KRP1-1*, *ZmHB124B*, and *ZmHB124* were associated with increases in biomass moisture (Figure 5C). Finally, for SCRIPT 4, the main genes involved in increases in FLW3 were *ZmTCP8*, *ZmTCP9*, *ZmTCP10*, *ZmTCP22*, and *ZmTCP42* (Figure 5D). In particular, LOFs in *ZmTCP22*, *ZmTCP42*, and *ZmTCP9* were associated with concomitant increases in FW and moisture content, and therefore decreases in DW. *ZmGRF10* and *ZmGRF4* were associated with increases in FLL3.

To further validate the rationale used for the associations, we analyzed population P012 of SCRIPT 1 in detail (Figure 6). In this population, *D8*, one of the selected genes associated with increases in FLL3 (Figure 5A), showed two haplotype$_{KO}$, each with an out-of-frame indel (−1 and +1 bp), and an haplotype$_{REF}$ with an in-frame indel (−3 bp) (Figure 6A). In the progeny, the haplotypes segregated, resulting in different LOF dosage combinations. Within that population, plants containing only a LOF$_{1/2}$ in *D8* presented similar phenotypes of FLL3 and FLW3 compared with the EDITOR 1 control, whereas plants with a LOF$_{2/2}$ in *D8* displayed longer and narrower leaves (Figure 6, B and C).

Using the data obtained from this genotype-to-phenotype association study, we built a putative regulatory network that integrates all the single-gene effects and their impact over the different measured traits (Figure 7). In this network, central genes (such as *ZmCKX4B*, *ZmCKX48*, and *ZmCKX46* and *ZmTCP9*, *ZmTCP10*, *ZmTCP22*, and *ZmTCP42*) act as nodes connected to several traits, implying a possible broader role in regulation. The genes located at the edge of the network may play a more defined role, connecting to just one or two traits (such as *ZmGRF10*, *ZmGRF4*, *ZmCKX3*, and *ZmTCP8*). Finally, other genes exhibited specific patterns, such as *D8*, *SLRL2*, and *Ga2ox5*, whose LOF strongly increased the FLL3 while also strongly decreasing the FLW.

Discussion

Complex agronomic traits, such as yield or tolerance to a particular (a)biotic stress, are governed by a large network of genes that together determine a specific phenotype. Understanding the complexity of such networks is the central goal of systems biology. Here, we developed an experimental approach, named BREEDIT, to study gene networks affecting complex quantitative traits by combining multiplex CRISPR-mediated gene editing of whole gene families with specific crossing schemes. In BREEDIT, a Cas9-expressing line (EDITOR) is supertransformed with vectors containing 12 gRNAs (SCRIPTs) targeting a set of GRGs. Gene edits are further stacked in plants using crossing schemes.

We evaluated the BREEDIT strategy by targeting putative players in major plant gene families or pathways involved in growth regulation. The success rate of the multiplex gene editing approach in maize was very high, with >97% of the genes showing at least partial or complete LOF in T1 plants. In just two generations, BREEDIT created multiple gene KOs leading to a diverse collection of genetic profiles, from low-order mutants with one, two, or three gene KOs to higher-order mutants stacking mutations in up to 11 genes out of the 12 within a single SCRIPT. Additional levers could be used to further increase the number of gene KOs stacked in one plant, namely inter-script crosses and the ongoing Cas9 activity. Regarding this latter point, our strategy is based on the notion that the ongoing Cas9 activity through generations would help to quickly reach a high saturation of LOF$_{2/2}$ edits. However, contrary to this idea, we observed that transgenerational gene editing did not result in such an outcome, as an average of only 7% of the progeny displayed new LOF$_{1/2}$ type edits per locus. This extra gene editing was mostly useful for generating de novo edits at loci not edited in the T0 generation. Therefore, the ongoing Cas9 activity can contribute to the effort to reach saturation of LOF in loci, but additional selfing steps might be needed to increase the number of gene KOs stacked in one plant. We also showed that higher-order mutants could be obtained by crossing plants already containing a high number of gene KOs at the single-SCRIPT level.

Interestingly, while in T0 plants, both copies of the target genes often carried the same or a different mutation (bi-allelic), genes newly edited in the T1 generation all showed heterozygous mutations, suggesting that only one chromosome of the two was edited. We suggest that chromatin condensation might influence DNA accessibility for the CRISPR/Cas9 machinery, possibly by imprinting (Borg and Berger, 2015). Further research is needed to elaborate on the mechanisms.

We obtained >1,000 plants, often with different unique LOF profiles, to score for phenotyping. The high sensitivity of HiPlex amplicon sequencing enabled us to capture complete sets of haplotypes with CRISPR/Cas9 mutations in large arrays of samples and loci. We used the haplotype sequence to focus on haplotypes thought to have major effects (haplotype$_{KO}$) on the function or activity of the translated protein. The experimental set-up based on multiple observations of significant single-gene KO associations with phenotypes across different populations and

Figure 5 Continued

phenotypic information for a specific trait. In such cases, the mean phenotypic value of each genotypic group could be statistically compared with that of the EDITOR 1 control. (2) Strength: for each gene, we calculated the weighted sum of observations in which the genotypic group with the highest mean phenotypic value is 10% above (weight: +1) or below (weight: −1) the mean phenotypic value of the EDITOR 1 control. The resulting sum was divided by the total number of observations (*n*). Associations displaying highest strength, either positive or negative, along with a large total number of observations indicate strong evidence for the effect of a gene on the trait.

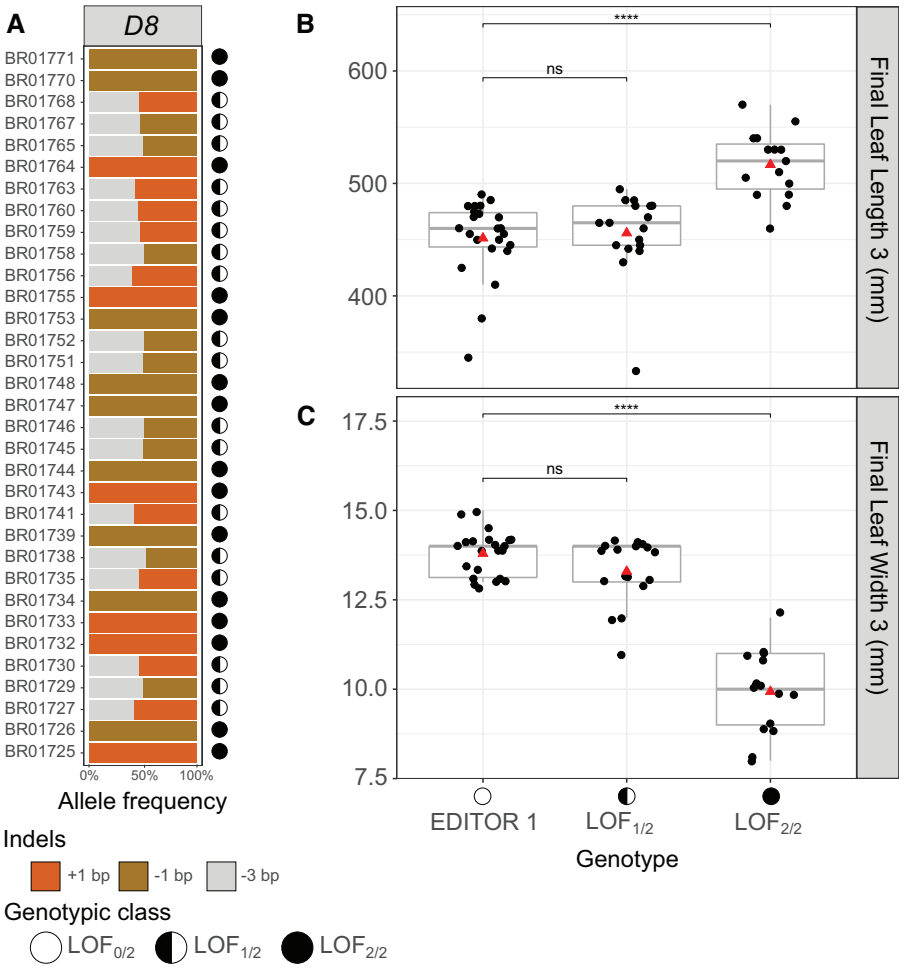


Figure 6 LOF dosages in *D8* and leaf shape parameters. A, Haplotype profiles at gene *D8* of T1 segregants from population P012. Three haplotypes were detected, with two containing out-of-frame indels (−1 bp and + 1 bp) and one containing an in-frame (−3 bp) deletion. This results in a collection of plants with *D8* either partially ($LOF_{1/2}$) or completely ($LOF_{2/2}$) knocked out. The resulting two classes of LOF dosages are compared with the EDITOR 1 control for final leaf length 3 (B) and final leaf width 3 (C). Significant differences (pairwise Student's *t*- test) are displayed with *P*-values summarized as follow: ****$P < 1e{-}4$, ns: not-significant. Box plot borders represent the first and third quartiles, the middle line represents the median, and whiskers represent the maximal and minimum values. Triangles indicate the mean of each distribution.

experiments enabled us to identify significant phenotypic responses in growth traits for all SCRIPTs. In the case of SCRIPT 1, we observed previously known effects of elevated gibberellic acid (GA) levels, such as plants with long and narrow leaves (Nelissen et al., 2012; Voorend et al., 2016), as well as male sterility, a trait that was previously associated with the effect of GA on tassel development (Colombo and Favret, 1996). SCRIPT 4 plants displayed an increased FLW3 (along with milder increases in FLL3), which affected the FW in some populations. Lastly, for SCRIPT 2 and SCRIPT 3, the most pronounced phenotypes observed were increased FLL3 and FW, particularly under WD conditions. If we consider these single-SCRIPT lines as building blocks, the possibility of stacking several different combinations of SCRIPTs by crossing paves the way toward creating higher-order mutant lines that display even stronger or additive phenotypes. The inter-script lines we generated displayed more than 12 higher-order mutations and inherited traits observed in parental single-SCRIPT lines. Although some of the expressed traits (e.g. increases in FW) were not observed in all different inter-script populations of the same type (probably because different haplotypes segregated in different populations), some other traits (such as increases in FLW) were consistently observed in all generations of lines containing SCRIPT 4, which further validated the consistency of the results analyzed at the single-SCRIPT level.

Another important outcome of the BREEDIT strategy is the possibility to screen large sets of genes that are then ranked and prioritized to delineate a minimal set of LOF required to induce a maximal phenotypic effect. Further inspection of the selected 48 GRGs showed that certain subsets of genes are strongly associated with specific traits (or a combination thereof). Therefore, using single-gene KO–trait associations, we identified subsets of genes per family whose LOF, alone or in combination, may be responsible for the observed phenotypes. Furthermore, some of these genes were already shown to play roles in modifying agronomic traits in other studies. A gain-of-function

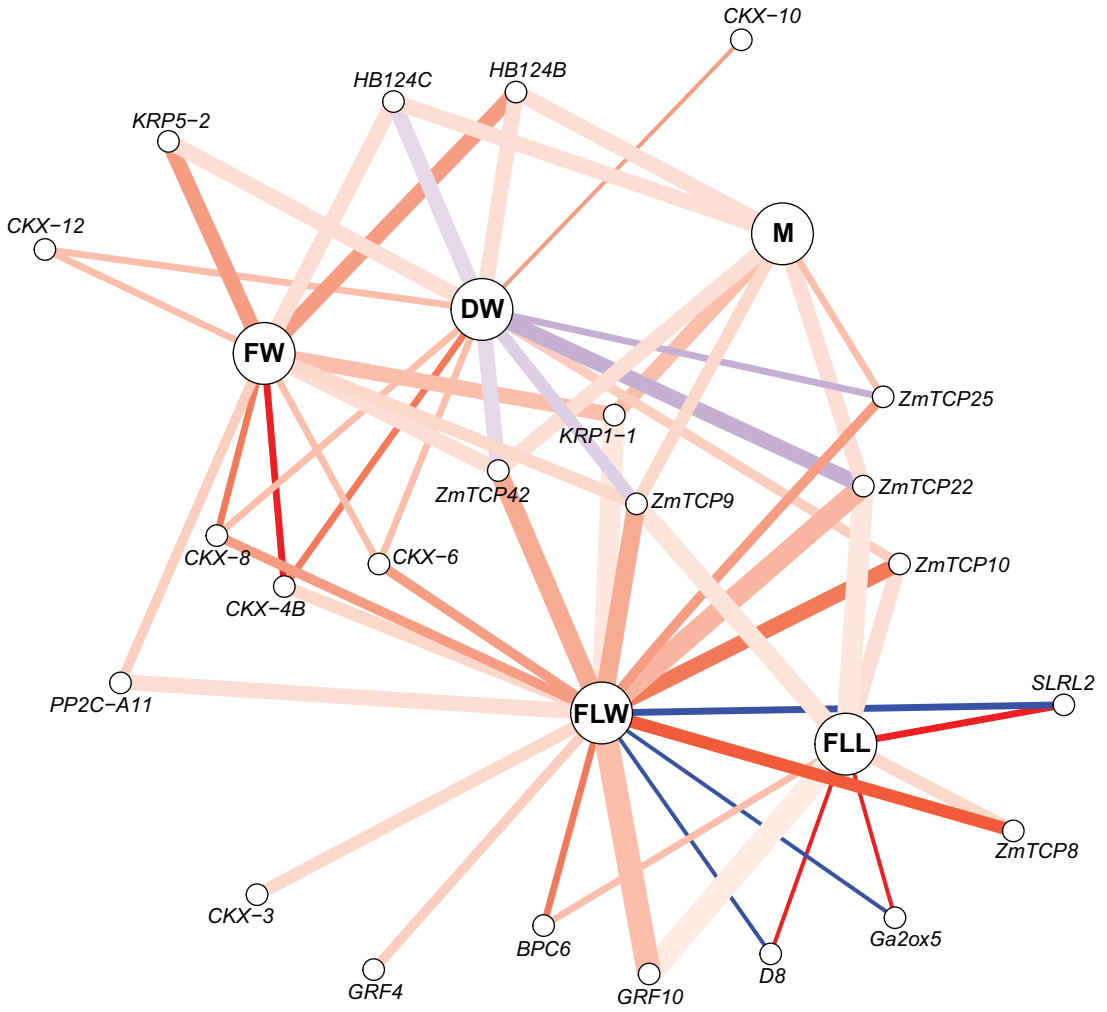


Figure 7 Network representation of single-gene effects on growth-related traits. Traits are displayed in bold (FLL, final leaf length; FLW, final leaf width; and M, moisture content). Genes associated at least once with a trait are displayed. Lines indicate connections between genes and traits. Line width is proportional to the number of times the underlying dataset to detect a gene KO–trait association in different experiments and/or populations contained sufficient data for statistics (i.e. a minimum of one LOF class between $LOF_{1/2}$ and $LOF_{2/2}$ with at least six individuals with phenotypic information). Line color represents the weighted fraction of gene KO–trait associations that significantly outperformed the EDITOR 1 control by 10% (ANOVA test; $P < 5\%$), either positively (weight: + 1, more red) or negatively (weight: −1, more blue), over the number of times a gene KO–trait association could have been observed due to sufficient data points.

mutation in *D8* (SCRIPT 1), encoding a DELLA maize protein ortholog, is causative of dwarf phenotypes (Winkler and Freeling, 1994; Lawit et al., 2010). Zm*CKX-4B* (SCRIPT 2) plays a role in both drought (Rida et al., 2021) and heat shock stress (Wang et al., 2020). Downregulation of *ZmHB124B* and *ZmHB124C* (SCRIPT 3) induces the formation of additional protoxylem files in the vasculature (Bloch et al., 2019), which prevent vascular embolism and water retention under water-limiting conditions (Hwang et al., 2016). *ZmGRF10* (SCRIPT 4) overexpression in maize leads to a decrease in leaf length and plant height (Wu et al., 2014; Nelissen et al., 2015). Although these genes were highlighted in our single-gene KO–trait associations, we cannot exclude the possibility that other genes from the original pool may play minor roles, either individually or in combination, but end up being masked by the effects of major gene KOs.

Nonetheless, this subset selection provides a pool of valuable material for further research to tackle specific traits (e.g. leaf width and enhanced growth under WD). The scored phenotypes were observed in plants grown in growth rooms under controlled conditions. Although seedling and leaf growth trait translate well from growth chambers into the field (Nelissen et al., 2020), less favorable phenotypes resulting from higher-order mutations might be observed at later developmental stages. One example is the male sterility phenotype observed in SCRIPT1 plants. Ideally, large populations of multiplex edited plants should be evaluated under field conditions but, at least in Europe, highly restrictive legislation prevents this type of analysis (Dima and Inzé, 2021).

While applying the BREEDIT strategy to our case study in maize, we identified an important limitation to this approach: the inability to fully uncouple complex gene interactions. In plant models where transformation and

regeneration are efficient, the possibility for massive gene editing grows more rapidly than the capacity to phenotypically analyze the resulting collections of mutants. For complex quantitative traits, large populations of lower-order mutants need to be screened accurately to decipher the complex mechanisms underlying plant development (Liu et al., 2020). To illustrate this, we developed the following multiplex edited scenario (Supplemental Figure S13). If one is interested in exhaustively capturing both additive and synergistic gene effects, all the gene KO combinations have to be generated and analyzed. Considering n genes to be targeted, the number of different genetic combinations that have to be produced amounts to 2^n in the case of two-state genes (homozygous wild-type or homozygous mutant) (Supplemental Figure S13A) and 3^n if the heterozygous stage has to be considered as well. Given that at least 10 replicates per genetic profile (combination) are required to statistically demonstrate a 10% significant difference in FLL3 with enough statistical power (Supplemental Figure S13B), the final number of plants to be processed increases dramatically as the number of genes in the study grows. The statistical power could be increased by performing large-scale phenotyping/genotyping in field conditions in order to detect combinatorial gene effects that govern agronomic traits, including seed yield.

Several approaches could be used to further study the causative genes. First, plants with a phenotype of interest but still containing Cas9 could be crossed with wild-type plants, and Cas9-negative plants with the desired trait could be identified in the T2 generation. Another approach is to use haploid induction (Chaikam et al., 2019; Jacquier et al., 2020), a promising technique applied to create homozygous mutations, thus removing the need to consider heterozygous material. This could be particularly interesting as a follow up to BREEDIT as a strategy to fix the new heterozygous gene edits that were observed in the T1 generation. Another approach is to preselect plants to be phenotyped based on their genetic constitution by predicting gene effects with statistical models in the same fashion as for genomic selection. Such predictions can be combined with the use of non-destructive seed chipping (Mills et al., 2020) to select specific gene combinations before sowing and therefore reduce the number of plants to be tested. Once the gene space is lowered, the BREEDIT pipeline can be followed again to design validation constructs by engineering a vector containing gRNAs targeting only the genes retained in the selected subset.

In this study, we developed the BREEDIT strategy to rapidly generate a large collection of mutants in specific gene families, pathways, or networks. We foresee great potential for BREEDIT combined with existing and more recent breeding approaches, such as marker-assisted breeding, haploid induction, and genomic selection. Effectively implementing the concept of breeding by editing using the BREEDIT pipeline still requires some practical obstacles to be overcome, such as the ability to transform and regenerate the plant material, obtain the desired gene KOs, and segregate out the original transgene construct. When these conditions are met, applying the BREEDIT pipeline allows many lines to be generated with specific combinations of gene KOs able to modify particular traits of interest. These engineered lines could be directly introduced in a hybrid breeding pipeline by crossing to elite material. Furthermore, the impact of favorable allele combinations on complex traits could be evaluated in different genetic backgrounds and across several generations to assess their heritability. Thus, BREEDIT could significantly speed up pre-breeding activities, which usually involve screening pools of diverse materials (wild species, landraces, and commercial varieties) for promising mutations and phenotypes (Teixeira and Guimarães, 2021) that must then be transferred into an intermediate set of materials that breeders can use to create new varieties. Introgression of alleles from a divergent pool of materials is often cumbersome due to cross incompatibility, low seed yield quantity and quality, or persistence of a deleterious linkage drag. Provided that elite materials can be transformed and regenerated, the reverse-genetics approach developed in the BREEDIT pipeline can circumvent the long and tedious step of introgression and save time in the development of new commercial varieties. An additional benefit of BREEDIT is that it could be used to generate large collections of plants mutated in coding or non-coding genome areas using other novel CRISPR technologies, such as base editing and promoter bashing, to further extend the repertoire of allele variability and phenotypic responses (Vats et al., 2019; Anzalone et al., 2020; Gaillochet et al., 2021).

Materials and methods

Plant material and DNA extraction

The original line used for all transformation procedures was the maize (*Z. mays*) inbred line B104. DNA was extracted from the samples following an adapted protocol from Berendzen et al. (2005) coupled with a magnetic bead purification. A 1–2-cm piece of leaf 1 was ground in 8-strip 2-mL capacity tubes (National Scientific Supply Co.). After grinding and centrifugation, the supernatant was mixed with magnetic beads (CleanNA), washed in 80% ethanol, and dried for further processing.

Selection of GRGs and curation of gene models

We selected 48 GRGs based on the literature, in house knowledge, and orthology searches (see "Results") in version 4 of the reference maize B73 genome sequence (Jiao et al., 2017). The integrative orthology viewer in PLAZA v4.5 (Van Bel et al., 2022) was used to identify most orthologous genes, including finding gene families from other species and identifying the corresponding maize B104 genes. When required, B104 gene models were manually curated using ORCAE, an online genome annotation resource (https://bioinformatics.psb.ugent.be/orcae/). Sequences in maize lines B104 and B73 were compared by pairwise alignment using Geneious Prime 2020.1.2 (https://www.geneious.com/prime/).

Design of the amplicons and gRNAs was performed in Geneious Prime. The maize B73 genome version 4 was used to identify gRNA on-target and off-target sites. gRNAs were selected with specificity score ⩾80%–85% and with no stretches of Ts (>4) or internal *Bsa*I or *Bbs*I restriction sites, which would interfere with gRNA expression or vector construction, respectively.

Monitoring CRISPR/Cas9 edits by HiPlex amplicon sequencing

Geneious Prime was used to design primers to amplify the genomic regions surrounding the gRNA cutting sites. Two amplicons per gene with a size range of 120–150 bp were manually selected. Each amplicon contained at least two gRNAs separated from either primer by at least 15 bp. The amplicons were selected to target the middle of the coding sequence with no overlap. The specificity of primers was checked in the maize B73 genome version 4, and only specific primers were retained (Supplemental Table S2). All primers were pooled in a HiPlex amplicon sequencing assay to sequence each locus in each plant simultaneously. HiPlex library preparation was performed by the Floodlight Genomics facility (Knoxville, Tennessee, USA) using MonsterPlex technology. Pilot runs of HiPlex amplicon sequencing were conducted to select the best amplicon per gene (out of the two). The selection was based on amplification efficiency in the HiPlex assay measured as read counts and unambiguous read-reference mapping. For each gene, the overlapping gRNA was selected for cloning into the expression vector.

We used SMAP haplotype-window analyses (Schaumont et al., 2022) to trim sequencing reads, identify haplotypes at each locus, and calculate the respective haplotype frequency per locus per sample. SMAP haplotype-window bioinformatics extracts haplotypes from HiPlex sequencing reads as the entire DNA sequence between the HiPlex primers per locus. Any unique DNA sequence is considered to be a haplotype. The total haplotype count is recorded per locus per sample, and the relative haplotype frequency per locus per sample is calculated. A haplotype detection threshold of at least 1% relative read depth per locus per sample was set to remove possible spurious haplotypes derived from amplification and/or sequencing artifacts. The nucleotide length difference between the haplotype sequence and the B104 reference sequence (length difference rate, LDR) was used to classify the mutations into three classes: SNPs (LDR = 0 but sequences are different), insertions (LDR > 0, the mutated haplotype is longer than the reference haplotype), and deletions (LDR < 0, the mutated haplotype is shorter than the reference haplotype). We defined haplotypes whose indel length is not a multiple of three nucleotides as $haplotype_{KO}$ because they generate a frame shift in the open-reading frame that likely leads to the translation of the wrong amino acid sequence downstream of the mutation and/or creates a premature stop codon, both of which could disrupt the protein's function or activity. Haplotypes with SNPs outside the cutting site and in-frame indels are referred to as $Haplotype_{REF}$ to denote possible minor impacts of their mutations on the resulting protein, which may still behave as the reference protein.

Maize is a diploid organism in which each gene has two alleles per nucleus, each derived from one of the two parents. In plant material that stably expresses CRISPR/Cas9 and gRNAs, continuously driving gene editing, one may expect to observe mosaic tissues, that is, patches of tissues within an organ that contain different genome sequences due to non-uniform gene editing. Mosaic tissues may occur both in primary transformants and subsequent generations due to the initial and ongoing Cas9 activity, respectively. A single leaf sample used for DNA preparation may therefore contain cells with different gene edits, resulting in the scoring of one individual with more than two alleles. The allele dosage is also affected by mosaicism. Multi-allelism resulting from mosaic tissues blurs the expected 50:50 read depth ratio commonly observed between the two alleles of a diploid organism. In addition, bi-allelism can be observed in non-mosaic tissues, with a plant harboring two indels of the same or different nature (in-frame or out-of-frame) following a 50:50 read depth ratio. Genotype-to-phenotype statistical associations require discrete genotypic classes (absence/presence, or homozygous wild type, or heterozygous or homozygous mutant). We therefore summed the relative fraction of $haplotype_{KO}$ per locus per sample to quantify how much the locus is affected by mutations leading to a LOF. The resulting aggregation ($\Sigma haplotype_{KO}$) is divided into three genotypic classes representing three dosages of $haplotype_{KO}$: $LOF_{0/2}$ (<15% of the read depth per locus per sample contain $haplotype_{KO}$), $LOF_{1/2}$ (40%–60% of the read depth per locus per sample contain $haplotype_{KO}$), and $LOF_{2/2}$ (>85% of the read depth per locus per sample contain $haplotype_{KO}$). Because distinguishing among these three groups is critical for analyzing dosage effects associated with a particular trait, any value outside of these three ranges was scored as a missing genotype call during the genotype-to-phenotype association analyses.

Construction and cloning of SCRIPT vectors

The gRNA entry vectors were constructed by PCR amplification of the entire pGG-[B-F]-OsU3-BbsI-ccdB-BbsI-[C-G] plasmids with Q5 High-Fidelity DNA polymerase (New England Biolabs) according to the manufacturer's guidelines. The primers contained an extension to insert unique linkers (Torella et al., 2014) between the scaffold and *OsU3* promoter (Supplemental Tables S3 and S4). Two of the five linkers were modified to contain *Not*I restriction sites to facilitate validation of the final expression vectors by restriction enzyme digestion (Supplemental Figure S1). Gibson assembly was performed with NEBuilder Hifi DNA Assembly Mix (New England Biolabs) to circularize the PCR products into entry vectors following the manufacturer's guidelines. The new entry vectors were confirmed by Sanger sequencing (Mix2Seq service, Eurofins Scientific). gRNA construction and Golden Gate assembly into binary vectors were

performed as previously described (Decaestecker et al., 2019). Briefly, paired gRNA entry vectors were created by PCR amplification (Red Taq DNA Polymerase Master Mix, VWR Life Science, or iProof High-Fidelity DNA Polymerase, Bio-Rad Laboratories) using the template plasmid pEN-2xTaU3 with primers containing the 20-nt spacer sequences and *Bbs*I restriction sites. Column-purified PCR products were cloned into the Golden Gate entry vectors via a Golden Gate reaction using *Bbs*I (New England Biolabs). All paired gRNA entry vectors were verified by Sanger sequencing.

Expression vectors (SCRIPT 1–4; Supplemental Figure S1) were constructed by a Golden Gate reaction with *Bsa*I (New England Biolabs) using the paired gRNA entry vectors and a destination vector as previously described (Decaestecker et al., 2019). The destination vector, pGGBb-AG, contains a GreenGate destination module (AG) and a bialaphos-resistance (*bar*) gene driven by the 35S promoter. The expression of each individual gRNA was alternatively driven by either the rice *OsU3* promoter or the wheat *TaU3* promoter (Xing et al., 2014). The SCRIPT vectors were transformed via heat-shock into *ccd*B-sensitive DH5α *Escherichia coli* cells, grown on LB medium containing 100 μg/mL spectinomycin, and extracted using a GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific). Quality control was performed by digestion with *Not*I (Promega). SCRIPTs were transformed into *Agrobacterium tumefaciens* EHA 105 cells by the freeze/thaw method and plated on YEB medium with 100 μg/mL rifampicin and 100 μg/mL spectinomycin. The gRNA entry and pGGBb-AG destination vectors can be obtained at https://gatewayvectors.vib.be/.

Generation of EDITOR maize lines

The *zCas9* coding sequence containing a *Z. mays*-codon optimized Cas9 (Xing et al., 2014) was cloned under the control of the *ZmUbiquitin1* (*ZmUbi1*) promoter (pZmUBIL) and NOS terminator in pEN-L4-AG-R1 (Houbaert et al., 2018) using GreenGate cloning (Lampropoulos et al., 2013). The transcriptional unit was recombined with pEN-L1-linker-L2 and the pHbm42GW7 destination vector (Karimi et al., 2013). The resulting construct (pXHb-pZmUBIL-zCas9-NOSt) allows maize transformants to be selected with hygromycin and is referred to as the EDITOR construct.

The EDITOR construct was introduced into maize line B104 using *Agrobacterium*-mediated transformation of immature embryos (Aesaert et al., 2022; Coussens et al., 2012) and hygromycin as a selection agent. Three independent lines (EDITOR 1–3) with a single-locus insertion event were selected and made homozygous for the T-DNA locus by self-crossing. To measure Cas9 protein levels, total proteins were extracted from leaf tissue of the EDITOR lines, separated by polyacrylamide gel electrophoresis, and blotted onto PVDF membranes. For quantification, the blots were incubated with anti-Cas9-HRP primary antibody (Abcam, 1:5,000) for 4 h and detected by chemiluminescence. Blots were also Ponceau-stained for the protein loading control. EDITOR 1 was crossed with wild-type B104 plants to yield heterozygous immature embryos for a second round of transformation (supertransformation) with each SCRIPT construct separately. Backcrosses render more seeds/embryos compared with self-crosses and facilitate the removal of Cas9 in the progeny by segregation. For each SCRIPT, at least 10 independent T0 supertransformants were obtained following BASTA selection and genotyped by HiPlex amplicon sequencing.

Experimental design and phenotyping

Maize seeds were soaked in water for 24 h and sown in 0.3-L square pots (7 × 7 × 7 cm) using "potgrond met meststof" (N.V. Van Israel) as substrate. The pots were then arrayed in groups of 24 in 48.0- × 30.5-cm trays, randomized, and placed in growth chambers with a controlled temperature (24°C), relative humidity (55%), and a 16:8 photoperiod with controlled light intensity (170–200 μmol/m^2/s photosynthetic active radiation provided by a mixture of 50/50 Radium halogen HRI-BT 400W/D Pro Daylight and Philips MASTER SON-T PIA Plus 400-W bulbs).

For WW conditions, plants were grown under a water regime of 2.4 g of water per gram of dry potting mix, while for WD assays, this was reduced to 1.1–1.4 g of water per gram of dry potting mix, with a water potential of approximately −100 kPa (Verbraeken et al., 2021). The final leaf length was measured at V3 (FLL3, when the collar of leaf 3 is fully developed) from the crown of the plant to the leaf tip, and the final leaf width (FLW3) was measured at the middle point of the leaf blade. For biomass, aerial parts of V3 seedlings were harvested and weighed for FW and then dried in a 60°C oven to estimate DW. Biomass moisture content was calculated on a DW basis as FW−DW/DW.

Statistical analysis to detect genotype-to-phenotype associations

Phenotypic datasets were trimmed to remove individuals that scored as under-developed (misshapen or not developed to the V3 stage at the moment of harvest) or overgrown (surpassed V3 at the moment of sampling) during the phenotyping trials. Within each population and experiment, one-way ANOVAs were then conducted at the single-gene level to check for differences between the control (EDITOR 1) and mutant groups ($LOF_{1/2}$ or $LOF_{2/2}$) (Supplemental Data Set S1). The minimal size of a mutant group to be considered in statistical analysis was six individuals having both phenotypic and genotypic data. Post hoc HSD Tukey's tests were then performed to assign each mutant group to a statistical group. Finally, we recorded the number of times a KO (either $LOF_{1/2}$ or $LOF_{2/2}$) of a specific gene was found to be significantly associated with a given trait while leading (on average) to a >10% increase or decrease compared with the control line (EDITOR 1). We compared that count with the number of times sufficient data were available to make a statistical conclusion about a gene KO effect and defined the resulting ratio as the strength of the association.

Accession numbers

The entire set of Illumina paired-end read sequences has been deposited in the Sequence Read Archive (DDBJ/ENA/GenBank) under BioProject accession number PRJNA815957. The gene IDs of the 48 GRGs targeted in this study are listed in Table 1.

Supplemental data

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

Supplemental Figure S1. Map of a SCRIPT construct containing 12 gRNAs.

Supplemental Figure S2. Locations of the 48 GRGs on the 10 maize chromosomes.

Supplemental Figure S3. Cas9 protein expression levels in three different EDITOR lines.

Supplemental Figure S4. Comparative summaries of the mutations observed in plants harboring SCRIPT 1 in an EDITOR 1 versus EDITOR 3 background.

Supplemental Figure S5. Haplotype overview in T0 material from the four SCRIPTs.

Supplemental Figure S6. Distribution of the aggregated frequency of haplotypes with out-of-frame indels (Σ Haplotype KO) in the ingroup samples.

Supplemental Figure S7. Appearance of new indels in inter-script crosses.

Supplemental Figure S8. Sterility phenotypes observed in tassels of SCRIPT 1 T0 plants.

Supplemental Figure S9. Phenotypes of mutated populations for the four SCRIPTs.

Supplemental Figure S10. Characteristic phenotypes observed in T1 plants of SCRIPT 2 under WD conditions.

Supplemental Figure S11. Characteristic phenotypes observed in T1 plants of SCRIPT 3 under WD conditions.

Supplemental Figure S12. Phenotypes of mutated inter-script populations.

Supplemental Figure S13. Considerations on the number of genes, states, and observations per group for multiplex gene editing experiments.

Supplemental Table S1. Detailed information about the populations used in the phenotyping assays.

Supplemental Table S2. List of gRNAs and associated primer pairs for edit detection using HiPlex amplicon sequencing.

Supplemental Table S3. Plasmid overview.

Supplemental Table S4. Primers used for plasmid building and sequencing.

Supplemental Data Set S1. Details of statistical analysis performed in this study.

Acknowledgments

The authors would like to thank Lieven Sterck for helping with gene model curation; Mansour Karimi for helping with the cloning of EDITOR constructs; and Pan Gong, Reinout Laureyns, and Ji Li for additional help with the selection of genes.

Funding

This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (H2020/2019-2025) under grant agreement No. 833866-BREEDIT.

Conflict of interest statement. The authors declare that they have no conflict of interest.

References

Aesaert S, Impens L, Coussens G, Van Lerberge E, Vanderhaeghen R, Desmet L, Vanhevel Y, Bossuyt S, Wambua AN, Van Lijsebettens M, et al. (2022) Optimized transformation and gene editing of the B104 public maize inbred by improved tissue culture and use of morphogenic regulators. Front Plant Sci **13**: 883847

Anzalone AV, Koblan LW, Liu DR (2020) Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. Nat Biotechnol **38**: 824–844

Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles ER, Qian Q, Kitano H, Matsuoka M (2005) Cytokinin oxidase regulates rice grain production. Science **309**: 741–745

Bai M, Yuan J, Kuang H, Gong P, Li S, Zhang Z, Liu B, Sun J, Yang M, Yang L, et al. (2020) Generation of a multiplex mutagenesis population via pooled CRISPR-Cas9 in soya bean. Plant Biotechnol J **18**: 721–731

Bartrina I, Otto E, Strnad M, Werner T, Schmülling T (2011) Cytokinin regulates the activity of reproductive meristems, flower organ size, ovule formation, and thus seed yield in *Arabidopsis thaliana*. Plant Cell **23**: 69–80

Baute J, Herman D, Coppens F, De Block J, Slabbinck B, Dell'Acqua M, Pè ME, Maere S, Nelissen H, Inzé D (2015) Correlation analysis of the transcriptome of growing leaves with mature leaf parameters in a maize RIL population. Genome Biol **16**: 168

Baute J, Herman D, Coppens F, De Block J, Slabbinck B, Dell'Acqua M, Pè ME, Maere S, Nelissen H, Inzé D (2016) Combined large-scale phenotyping and transcriptomics in maize reveals a robust growth regulatory network. Plant Physiol **170**: 1848–1867

Berendzen K, Searle I, Ravenscroft D, Koncz C, Batschauer A, Coupland G, Somssich IE, Ülker B (2005) A rapid and versatile combined DNA/RNA extraction protocol and its application to the analysis of a novel DNA marker set polymorphic between *Arabidopsis thaliana* ecotypes Col-0 and Landsberg *erecta*. Plant Methods **1**: 4

Bessoltane N, Charlot F, Guyon-Debast A, Charif D, Mara K, Collonnier C, Perroud P-F, Tepfer M, Nogué F (2022) Genome-wide specificity of plant genome editing by both CRISPR-Cas9 and TALEN. Sci Rep **12**: 9330

Bhat JA, Yu D, Bohra A, Ganie SA, Varshney RK (2021) Features and applications of haplotypes in crop breeding. Commun Biol **4**: 1266

Bloch D, Puli MR, Mosquna A, Yalovsky S (2019) Abiotic stress modulates root patterning via ABA-regulated microRNA expression in the endodermis initials. Development **146**: dev177097

Borg M, Berger F (2015) Chromatin remodelling during male gametophyte development. Plant J **83**: 177–188

Brás TA, Seixas J, Carvalhais N, Jägermeyr J (2021) Severity of drought and heatwave crop losses tripled over the last five decades in Europe. Environ Res Lett **16**: 065012

Cao L, Wang S, Venglat P, Zhao L, Cheng Y, Ye S, Qin Y, Datla R, Zhou Y, Wang H (2018) Arabidopsis ICK/KRP cyclin-dependent kinase inhibitors function to ensure the formation of one megaspore mother cell and one functional megaspore per ovule. PLoS Genet **14**: e1007230

Chaikam V, Molenaar W, Melchinger AE, Boddupalli PM (2019) Doubled haploid technology for line development in maize: technical advances and prospects. Theor Appl Genet **132**: 3227–3243

Cheng Y, Cao L, Wang S, Li Y, Shi X, Liu H, Li L, Zhang Z, Fowke LC, Wang H, et al. (2013) Downregulation of multiple CDK inhibitor *ICK/KRP* genes upregulates the E2F pathway and increases cell proliferation, and organ and seed sizes in Arabidopsis. Plant J **75**: 642–655

Colombo N, Favret EA (1996) The effect of gibberellic acid on male fertility in bread wheat. Euphytica **91**: 297–303

Coussens G, Aesaert S, Verelst W, Demeulenaere M, De Buck S, Njuguna E, Inzé D, Van Lijsebettens M (2012) *Brachypodium distachyon* promoters as efficient building blocks for transgenic research in maize. J Exp Bot **63**: 4263–4273

Czesnick H, Lenhard M (2015) Size control in plants—lessons from leaves and flowers. Cold Spring Harb Perspect Biol **7**: a019190

Decaestecker W, Andrade Buono R, Pfeiffer ML, Vangheluwe N, Jourquin J, Karimi M, Van Isterdael G, Beeckman T, Nowack MK, Jacobs TB (2019) CRISPR-TSKO: a technique for efficient mutagenesis in specific cell types, tissues, or organs in Arabidopsis. Plant Cell **31**: 2868–2887

Dima O, Inzé D (2021) The role of scientists in policy making for more sustainable agriculture. Curr Biol **31**: R218–R220

Doll NM, Gilles LM, Gerentes M-F, Richard C, Just J, Fierlej Y, Borrelli VMG, Gendrot G, Ingram GC, Rogowsky PM, et al. (2019) Single and multiple gene knockouts by CRISPR–Cas9 in maize. Plant Cell Rep **38**: 487–501

Elias F, Muleta D, Woyessa D (2016) Effects of phosphate solubilizing fungi on growth and yield of haricot bean (*Phaseolus vulgaris* L.) plants. J Agric Sci **8**: 204–218

Gaillochet C, Develtere W, Jacobs TB (2021) CRISPR screens in plants: approaches, guidelines, and future prospects. Plant Cell **33**: 794–813

Gong P, Bontinck M, Demuynck K, De Block J, Gevaert K, Eeckhout D, Persiau G, Aesaert S, Coussens G, Van Lijsebettens M, et al. (2022) SAMBA controls cell division rate during maize development. Plant Physiol **188**: 411–424

Gong R, Cao H, Zhang J, Xie K, Wang D, Yu S (2018) Divergent functions of the GAGA-binding transcription factor family in rice. Plant J **94**: 32–47

Gonzalez N, Vanhaeren H, Inzé D (2012) Leaf size control: complex coordination of cell division and expansion. Trends Plant Sci **17**: 332–340

Hai NN, Chuong NN, Tu NHC, Kisiala A, Hoang XLT, Thao NP (2020) Role and regulation of cytokinins in plant response to drought stress. Plants **9**: 422

He Z, Wu J, Sun X, Dai M (2019) The maize clade A PP2C phosphatases play critical roles in multiple abiotic stress responses. Int J Mol Sci **20**: 3573

Houbaert A, Zhang C, Tiwari M, Wang K, de Marcos Serrano A, Savatin DV, Urs MJ, Zhiponova MK, Gudesblat GE, Vanhoutte I, et al. (2018) POLAR-guided signalling complex assembly and localization drive asymmetric cell division. Nature **563**: 574–578

Huang Y, Wang X, Ge S, Rao G-Y (2015) Divergence and adaptive evolution of the gibberellin oxidase genes in plants. BMC Evol Biol **15**: 207

Hwang BG, Ryu J, Lee SJ (2016) Vulnerability of protoxylem and metaxylem vessels to embolisms and radial refilling in a vascular bundle of maize leaves. Front Plant Sci **7**: 941

Ikeda A, Ueguchi-Tanaka M, Sonoda Y, Kitano H, Koshioka M, Futsuhara Y, Matsuoka M, Yamaguchi J (2001) *slender* rice, a constitutive gibberellin response mutant, is caused by a null mutation of the *SLR1* gene, an ortholog of the height-regulating gene *GAI/RGA/RHT/D8*. Plant Cell **13**: 999–1010

Impens L, Jacobs TB, Nelissen H, Inzé D, Pauwels L (2022) Mini-review: transgenerational CRISPR/Cas9 gene editing in plants. Front Genome Edit **4**: 825042

Itoh H, Shimada A, Ueguchi-Tanaka M, Kamiya N, Hasegawa Y, Ashikari M, Matsuoka M (2005) Overexpression of a GRAS protein lacking the DELLA domain confers altered gibberellin responses in rice. Plant J **44**: 669–679

Jacobs TB, Zhang N, Patel D, Martin GB (2017) Generation of a collection of mutant tomato lines using pooled CRISPR libraries. Plant Physiol **174**: 2023–2037

Jacquier NMA, Gilles LM, Pyott DE, Martinant J-P, Rogowsky PM, Widiez T (2020) Puzzling out plant reproduction by haploid induction for innovations in plant breeding. Nat Plants **6**: 610–619

Jiao Y, Peluso P, Shi J, Liang T, Stitzer MC, Wang B, Campbell MS, Stein JC, Wei X, Chin C-S, et al. (2017) Improved maize reference genome with single-molecule technologies. Nature **546**: 524–527

Karimi M, Inzé D, Van Lijsebettens M, Hilson P (2013) Gateway vectors for transformation of cereals. Trends Plant Sci. **18**: 1–4

Knott GJ, Doudna JA (2018) CRISPR-Cas guides the future of genetic engineering. Science **361**: 866–869

Koyama T, Sato F, Ohme-Takagi M (2017) Roles of miR319 and TCP transcription factors in leaf development. Plant Physiol **175**: 874–885

Lampropoulos A, Sutikovic Z, Wenzl C, Maegele I, Lohmann JU, Forner J (2013). GreenGate—a novel, versatile, and efficient cloning system for plant transgenesis. PLoS ONE **8**: e83043

Lan J, Qin G. (2020) The regulation of CIN-like TCP transcription factors. Int J Mol Sci **21**: 4498

Lawit SJ, Wych HM, Xu D, Kundu S, Tomes DT (2010) Maize DELLA proteins dwarf plant8 and dwarf plant9 as modulators of plant development. Plant Cell Physiol **51**: 1854–1868

Lee K, Zhang Y, Kleinstiver BP, Guo JA, Aryee MJ, Miller J, Malzahn A, Zarecor S, Lawrence-Dill CJ, Joung JK, et al. (2019) Activities and specificities of CRISPR/Cas9 and Cas12a nucleases for targeted mutagenesis in maize. Plant Biotechnol J **17**: 362–372

Li C, Hao M, Wang W, Wang H, Chen F, Chu W, Zhang B, Mei D, Cheng H, Hu Q (2018) An efficient CRISPR/Cas9 platform for rapidly generating simultaneous mutagenesis of multiple gene homoeologs in allotetraploid oilseed rape. Front Plant Sci **9**: 442

Li J, Wang Z, He G, Ma L, Deng XW (2020) CRISPR/Cas9-mediated disruption of *TaNP1* genes results in complete male sterility in bread wheat. J Genet Genomics **47**: 263–272

Li N, Li Y (2016) Signaling pathways of seed size control in plants. Curr Opin Plant Biol **33**: 23–32

Li W, Herrera-Estrella L, Tran L-SP (2016) The yin–yang of cytokinin homeostasis and drought acclimation/adaptation. Trends Plant Sci **21**: 548–550

Li Y, Shan X, Jiang Z, Zhao L, Jin F (2021) Genome-wide identification and expression analysis of the *GA2ox* gene family in maize (*Zea mays* L.) under various abiotic stress conditions. Plant Physiol Biochem **166**: 621–633

Liebsch D, Palatnik JF (2020) MicroRNA miR396, GRF transcription factors and GIF co-regulators: a conserved plant growth regulatory module with potential for breeding and biotechnology. Curr Opin Plant Biol **53**: 31–42

Liu H-J, Jian L, Xu J, Zhang Q, Zhang M, Jin M, Peng Y, Yan J, Han B, Liu J, et al. (2020) High-throughput CRISPR/Cas9 mutagenesis streamlines trait gene identification in maize. Plant Cell **32**: 1397–1413

Liu L, Gallagher J, Arevalo ED, Chen R, Skopelitis T, Wu Q, Bartlett M, Jackson D (2021) Enhancing grain-yield-related traits by CRISPR–Cas9 promoter editing of maize *CLE* genes. Nat Plants **7**: 287–294

Long SP, Marshall-Colon A, Zhu X-G (2015) Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. Cell **161**: 56–66

Lu Y, Ye X, Guo R, Huang J, Wang W, Tang J, Tan L, Zhu J-K, Chu C, Qian Y (2017) Genome-wide targeted mutagenesis in rice using the CRISPR/Cas9 system. Mol Plant **10**: 1242–1245

McConnell JR, Emery J, Eshed Y, Bao N, Bowman J, Barton MK (2001) Role of *PHABULOSA* and *PHAVOLUTA* in determining radial patterning in shoots. Nature **411**: 709–713

Meng X, Yu H, Zhang Y, Zhuang F, Song X, Gao S, Gao C, Li J (2017) Construction of a genome-wide mutant library in rice using CRISPR/Cas9. Mol Plant **10**: 1238–1241

Mickelbart MV, Hasegawa PM, Bailey-Serres J (2015) Genetic mechanisms of abiotic stress tolerance that translate to crop yield stability. Nat Rev Genet **16**: 237–251

Miculan M, Nelissen H, Hassen MB, Marroni F, Inzé D, Pè ME, Dell'Acqua M (2021) A forward genetics approach integrating genome-wide association study and expression quantitative trait locus mapping to dissect leaf development in maize (*Zea mays*). Plant J **107**: 1056–1071

Mills A, Allsman L, Leon S, Rasmussen C (2020) Using seed chipping to genotype maize kernels. Bio-Protocol 101: e3553

Nelissen H, Sprenger H, Demuynck K, De Block J, Van Hautegem T, De Vliegher A, Inzé D (2020) From laboratory to field: yield stability and shade avoidance genes are massively differentially expressed in the field. Plant Biotechnol J **18**: 1112–1114

Nelissen H, Rymen B, Jikumaru Y, Demuynck K, Van Lijsebettens M, Kamiya Y, Inzé D, Beemster GTS (2012) A local maximum in gibberellin levels regulates maize leaf growth by spatial control of cell division. Curr Biol **22**: 1183–1187

Nelissen H, Eeckhout D, Demuynck K, Persiau G, Walton A, van Bel M, Vervoort M, Candaele J, De Block J, Aesaert S, et al. (2015) Dynamic changes in ANGUSTIFOLIA3 complex composition reveal a growth regulatory mechanism in the maize leaf. Plant Cell **27**: 1605–1619

Nuccio ML, Paul M, Bate NJ, Cohn J, Cutler SR (2018) Where are the drought tolerant crops? An assessment of more than two decades of plant biotechnology effort in crop improvement. Plant Sci **273**: 110–119

Paul BK, Frelat R, Birnholz C, Ebong C, Gahigi A, Groot JCJ, Herrero M, Kagabo DM, Notenbaert A, Vanlauwe B, et al. (2018) Agricultural intensification scenarios, household food availability and greenhouse gas emissions in Rwanda: Ex-ante impacts and trade-offs. Agric Syst **163**: 16–26

Poland J, Rutkoski J (2016) Advances and challenges in genomic selection for disease resistance. Annu Rev Phytopathol **54**: 79–98

Qin F, Kodaira K-S, Maruyama K, Mizoi J, Tran L-SP, Fujita Y, Morimoto K, Shinozaki K, Yamaguchi-Shinozaki K (2011) *SPINDLY*, a negative regulator of gibberellic acid signaling, is involved in the plant abiotic stress response. Plant Physiol **157**: 1900–1913

Ramadan M, Alariqi M, Ma Y, Li Y, Liu Z, Zhang R, Jin S, Min L, Zhang X (2021) Efficient CRISPR/Cas9 mediated pooled-sgRNAs assembly accelerates targeting multiple genes related to male sterility in cotton. Plant Methods **17**: 16

Rasheed A, Hao Y, Xia X, Khan A, Xu Y, Varshney RK, He Z (2017) Crop breeding chips and genotyping platforms: progress, challenges, and perspectives. Mol Plant **10**: 1047–1064

Rida S, Maafi O, López-Malvar A, Revilla P, Riache M, Djemel A (2021) Genetics of germination and seedling traits under drought stress in a MAGIC population of maize. Plants **10**: 1786

Rodríguez-Leal D, Lemmon ZH, Man J, Bartlett ME, Lippman ZB (2017) Engineering quantitative trait variation for crop improvement by genome editing. Cell **171**: 470–480.e478

Sarvepalli K, Nath U (2018) CIN-TCP transcription factors: transiting cell proliferation in plants. IUBMB Life **70**: 718–731

Schaumont D, Veeckman E, Van der Jeugt F, Haegeman A, Glabeke SV, Bawin Y, Lukasiewicz J, Blugeon S, Barre P, Leyva-Pérez M de la O, et al. (2022) Stack mapping anchor points (SMAP): a versatile suite of tools for read-backed haplotyping. bioRxiv 2022.03.10.483555

Simmons CR, Lafitte HR, Reimann KS, Brugière N, Roesler K, Albertsen MC, Greene TW, Habben JE (2021) Successes and insights of an industry biotech program to enhance maize agronomic traits. Plant Sci **307**: 110899

Snowdon RJ, Wittkop B, Chen T-W, Stahl A (2021) Crop adaptation to climate change as a consequence of long-term breeding. Theor Appl Genet **134**: 1613–1623

Sun X, Cahill J, Van Hautegem T, Feys K, Whipple C, Novák O, Delbare S, Versteele C, Demuynck K, De Block J, et al. (2017) Altered expression of maize *PLASTOCHRON1* enhances biomass and seed yield by extending cell division duration. Nat Commun **8**: 14752

Teixeira FF, Guimarães CT (2021) Chapter 5—Maize genetic resources and pre-breeding. *In* MT Azhar and SH Wani, eds, Wild Germplasm for Genetic Improvement in Crop Plants. Academic Press, Elsevier, London, pp 81–99

Torella JP, Lienert F, Boehm CR, Chen J-H, Way JC, Silver PA (2014) Unique nucleotide sequence-guided assembly of repetitive DNA parts for synthetic biology applications. Nat Protoc **9**: 2075–2089

Van Bel M, Silvestri F, Weitz EM, Kreft L, Botzki A, Coppens F, Vandepoele K (2022) PLAZA 5.0: extending the scope and power of comparative and functional genomics in plants. Nucleic Acids Res **50**: D1468–D1474

Vanhaeren H, Nam Y-J, De Milde L, Chae E, Storme V, Weigel D, Gonzalez N, Inzé D (2017) Forever young: the role of ubiquitin receptor DA1 and E3 ligase Big Brother in controlling leaf growth and development. Plant Physiol **173**: 1269–1282

Vanhaeren H, Gonzalez N, Coppens F, De Milde L, Van Daele T, Vermeersch M, Eloy NB, Storme V, Inzé D (2014). Combining growth-promoting genes leads to positive epistasis in *Arabidopsis thaliana*. eLife **3**: e02252

Vats S, Kumawat S, Kumar V, Patil GB, Joshi T, Sonah H, Sharma TR, Deshmukh R (2019) Genome editing in plants: exploration of technological advancements and challenges. Cells **8**: 1386

Verbraeken L, Wuyts N, Mertens S, Cannoot B, Maleux K, Demuynck K, De Block J, Merchie J, Dhondt S, Bonaventure G, et al. (2021) Drought affects the rate and duration of organ growth but not inter-organ growth coordination. Plant Physiol **186**: 1336–1353

Vercruysse J, Baekelandt A, Gonzalez N, Inzé D (2020) Molecular networks regulating cell division during Arabidopsis leaf growth. J Exp Bot **71**: 2365–2378

Voorend W, Nelissen H, Vanholme R, De Vliegher A, Van Breusegem F, Boerjan W, Roldán-Ruiz I, Muylle H, Inzé D (2016) Overexpression of *GA20-OXIDASE1* impacts plant height, biomass allocation and saccharification efficiency in maize. Plant Biotechnol J **14**: 997–1007

Voss-Fels K, Snowdon RJ (2016) Understanding and utilizing crop genome diversity via high-resolution genotyping. Plant Biotechnol J **14**: 1086–1094

Wang B, Li N, Huang S, Hu J, Wang Q, Tang Y, Yang T, Asmutola P, Wang J, Yu Q (2021) Enhanced soluble sugar content in tomato fruit using CRISPR/Cas9-mediated *SlINVINH1* and *SlVPE5* gene editing. PeerJ **9**: e12478

Wang H-Q, Liu P, Zhang J-W, Zhao B, Ren B-Z (2020) Endogenous hormones inhibit differentiation of young ears in maize (*Zea mays* L.) under heat stress. Front Plant Sci **11**: 533046

Wang H, Qin F (2017) Genome-wide association study reveals natural variations contributing to drought resistance in crops. Front Plant Sci **8**: 1110

Winkler RG, Freeling M (1994) Physiological genetics of the dominant gibberellin-nonresponsive maize dwarfs, *Dwarf8* and *Dwarf9*. Planta **193**: 341–348

Wu L, Zhang D, Xue M, Qian J, He Y, Wang S (2014) Overexpression of the maize *GRF10*, an endogenous truncated growth-regulating factor protein, leads to reduction in leaf size and plant height. J Integr Plant Biol **56**: 1053–1063

Xiao Y, Tong H, Yang X, Xu S, Pan Q, Qiao F, Raihan MS, Luo Y, Liu H, Zhang X, et al. (2016) Genome-wide dissection of the maize ear genetic architecture using multiple populations. New Phytol **210**: 1095–1106

Xing H-L, Dong L, Wang Z-P, Zhang H-Y, Han C-Y, Liu B, Wang X-C, Chen Q-J (2014) A CRISPR/Cas9 toolkit for multiplex genome editing in plants. BMC Plant Biol **14**: 327

Zhang X, Cai X (2011) Climate change impacts on global agricultural land availability. Environ Res Lett **6**: 014014

Zhang Y, Malzahn AA, Sretenovic S, Qi Y (2019) The emerging and uncultivated potential of CRISPR technology in plant science. Nat Plants **5**: 778–794