

## PLANT SCIENCES

# FIND-IT: Accelerated trait development for a green evolution

Søren Knudsen<sup>1†</sup>, Toni Wendt<sup>1†</sup>, Christoph Dockter<sup>1\*†</sup>, Hanne Cecilie Thomsen<sup>1†</sup>, Magnus Rasmussen<sup>1</sup>, Morten Egevang Jørgensen<sup>1</sup>, Qiongxin Lu<sup>1</sup>, Cynthia Voss<sup>1</sup>, Emiko Murozuka<sup>1</sup>, Jeppe Thulin Østerberg<sup>1</sup>, Jesper Harholt<sup>1</sup>, Ilka Braumann<sup>1</sup>, Jose A. Cuesta-Seijo<sup>1</sup>, Sandip M. Kale<sup>1</sup>, Sabrina Bodevin<sup>1</sup>, Lise Tang Petersen<sup>1</sup>, Massimiliano Carciofi<sup>1</sup>, Pai Rosager Pedas<sup>1</sup>, Jeppe Opstrup Husum<sup>1</sup>, Martin Toft Simmelsgaard Nielsen<sup>1</sup>, Kasper Nielsen<sup>1</sup>, Mikkel K. Jensen<sup>1</sup>, Lillian Ambus Møller<sup>1</sup>, Zoran Gojkovic<sup>1</sup>, Alexander Striebeck<sup>1</sup>, Klaus Lengeler<sup>1</sup>, Ross T. Fennesy<sup>1</sup>, Michael Katz<sup>1</sup>, Rosa Garcia Sanchez<sup>1</sup>, Natalia Solodovnikova<sup>1‡</sup>, Jochen Förster<sup>1</sup>, Ole Olsen<sup>1</sup>, Birger Lindberg Møller<sup>2</sup>, Geoffrey B. Fincher<sup>3</sup>, Birgitte Skadhauge<sup>1\*</sup>

Improved agricultural and industrial production organisms are required to meet the future global food demands and minimize the effects of climate change. A new resource for crop and microbe improvement, designated FIND-IT (Fast Identification of Nucleotide variants by droplet DigITal PCR), provides ultrafast identification and isolation of predetermined, targeted genetic variants in a screening cycle of less than 10 days. Using large-scale sample pooling in combination with droplet digital PCR (ddPCR) greatly increases the size of low-mutation density and screenable variant libraries and the probability of identifying the variant of interest. The method is validated by screening variant libraries totaling 500,000 barley (*Hordeum vulgare*) individuals and isolating more than 125 targeted barley gene knockout lines and miRNA or promoter variants enabling functional gene analysis. FIND-IT variants are directly applicable to elite breeding pipelines and minimize time-consuming technical steps to accelerate the evolution of germplasm.

## INTRODUCTION

Environmentally benign agricultural production systems are required to meet the future global food demands and minimize the effect on climate change. During the green revolution of the 1960s and 1970s, the global problem with food shortfalls was solved by conventional and variation breeding of high-yielding varieties suitable for intensified crop farming practices, but dependent on heavy use of non-renewable resources, with severe negative long-term environmental impacts (1). Globally, green revolution varieties in maize, wheat, and rice were integrated quickly in farming practices and these high-yielding variants remain in the pedigree of current elite germplasm (2). This stands in stark contrast to the restricted adoption of the subsequently developed crops obtained using the tools of genetic modification (GM; here defined as crops developed using transformation and introduction of foreign DNA). Today, in the 26th year of commercialization, the growth of GM crops is restricted to ~30 countries, with only ~5 of these being classified as industrialized countries (3). The total land area covered with GM crops is ~200 million hectares, equaling ~15% of the world's arable land (4). Furthermore, emerging gene editing technologies have technical

drawbacks and uncertainties that have so far precluded their widespread utilization in commercial plant breeding (5, 6). Accordingly, no generally accepted, fast, and reliable GM technology is available to breed new resilient high-yielding crop varieties with improved characteristics to meet the global food demand and counteract the negative effects of climate change, loss of biodiversity, and soil nutrient depletion (7, 8).

Traditionally, large, natural variant populations (e.g., the natural germplasm diversity in crop fields) with low spontaneous mutation densities have been the primary source of natural gene variation used to domesticate and adapt crops for superior performance (9). Induced mutagenesis increases genetic variation in populations and has been used for crop development for nearly 100 years (10), notably in the genetic improvements underlying the green revolution high-yield crop traits (11). More recently developed chemical mutagenesis treatments, which primarily induce single-nucleotide substitutions (12), can be used to obtain large populations with many thousands of useful and previously unidentified variants. Phenotype-based methods (13) or genome-targeted screening methods like TILLING (Targeting Induced Local Lesions IN Genomes) (14–16) have been used to screen the natural or induced variant populations for desirable traits, but the number of available gene variants successfully targeted, isolated, and adopted by breeders is limited.

Here, we present and validate FIND-IT [Fast Identification of Nucleotide variants by droplet DigITal PCR (polymerase chain reaction)], a new non-GM approach for screening variant populations, which meets the demands for rapid development of resilient and environmentally benign crop plants and for improving the industrial performance of microorganisms. This simple, agile, and high-throughput approach combines systematic sample pooling and

Copyright © 2022  
The Authors, some  
rights reserved;  
exclusive licensee  
American Association  
for the Advancement  
of Science. No claim to  
original U.S. Government  
Works. Distributed  
under a Creative  
Commons Attribution  
NonCommercial  
License 4.0 (CC BY-NC).

<sup>1</sup>Carlsberg Research Laboratory, J.C. Jacobsens Gade 4, DK-1799 Copenhagen V, Denmark. <sup>2</sup>Plant Biochemistry Laboratory, Centre for Synthetic Biology, Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, DK-1871 Frederiksberg C, Denmark. <sup>3</sup>Australian Research Council Centre of Excellence in Plant Cell Walls School of Agriculture, Food and Wine, University of Adelaide, Waite Campus, Glen Osmond, SA 5064, Australia.

\*Corresponding author. Email: birgitte.skadhauge@carlsberg.com (B.S.); christoph.dockter@carlsberg.com (C.D.)

†These authors contributed equally to this work.

‡Present address: European Innovation Center, FMC Corporation, Genvej 2, 2970 Hørsholm, Denmark.

splitting with high-sensitivity, droplet digital PCR (ddPCR)-based genotyping to screen large, low-mutation density variant populations (e.g., variant libraries derived from induced mutagenesis or natural germplasm collections) for targeted identification of desired traits at single-nucleotide resolution. The FIND-IT approach is independent of genetic engineering technologies and is applicable to any living organism that can be grown in the field or in culture (i.e., elite, wild, or orphan plant crop varieties, bacteria, and yeasts). The generation of induced variant libraries is simple and cost-competitive. The libraries can be updated or exchanged regularly, thus adapting the variant identification pipeline to current species and newly available breeding lines. FIND-IT integrates seamlessly with existing crossing protocols in breeding pipelines and in food or biotechnological production systems. Identified variants can be immediately used in crop or microbial strain improvement strategies while simultaneously alleviating the need for time-consuming backcrosses. Using the cereal crop barley as an example, we show the design, efficient screening, and validation of FIND-IT by targeted isolation of more than 125 knockout and missense variants, as well as miRNA and promoter gain-of-function mutations. By targeting specified single-nucleotide changes in genes that modify barley crop performance, grain morphology, and grain quality properties, we demonstrate how FIND-IT circumvents limitations of comparable techniques and provides a novel resource for fast and sustainable evolution of today's germplasm.

## RESULTS

### General description of the FIND-IT screening method

The mutational load and the type of genetic variance induced by mutagenesis form the basis for successful identification of the desired variant in FIND-IT. Chemical mutagenesis has a major advantage over radiation-induced mutagenesis by predominantly causing single-nucleotide changes. Depending on the species, the total population size, and the mutagen and mutagenesis protocol used (17, 18), a substitution at many genome nucleotides can be facilitated without additional changes associated with base/nucleotide excision repair. If achieved, random mutagenesis can be used in a targeted variant isolation approach (6), given that an efficient, high-sensitivity screening method is available.

The ddPCR technology is 1000-fold more sensitive than conventional PCR (19). In FIND-IT, this high sensitivity enables the screening of DNA pools representing a very high number of individuals (organism specific) from large variant libraries in a 96-well plate. After the addition of PCR reactants, which include two competing fluorescent TaqMan oligonucleotide probes that distinguish between the wild-type and variant alleles of interest [design principles described in (20)], the DNA and PCR mixture is compartmentalized by a film of oil into approximately 20,000 droplets per well, before PCR amplification. The PCR displaces the allele-specific fluorophores, and the released fluorochromes can be detected through ultraviolet scanning of the plates. Scanning all droplets reveals the well and DNA pool containing the variant of interest. In this first phase of the screening process, the population containing the targeted mutant of interest is narrowed down from all individuals of the initial library to a single pool of candidates. In the second phase, DNA from the selected individuals in the pool of interest is subsampled in a new 96-well plate and the ddPCR procedure is repeated, further reducing the number of candidate individuals in the subpool carrying the

targeted gene variant. In the third phase, DNA is isolated from the few remaining individuals in the phase 2-detected subpool and the variant of interest is confirmed by PCR and gene sequencing. Using this pool-and-split strategy, coupled with high-sensitivity ddPCR, facilitates the targeted isolation of genetic variants from large variant populations as described in detail for barley below and in Fig. 1.

### Designing SNP libraries for targeted variant identification using the FIND-IT approach

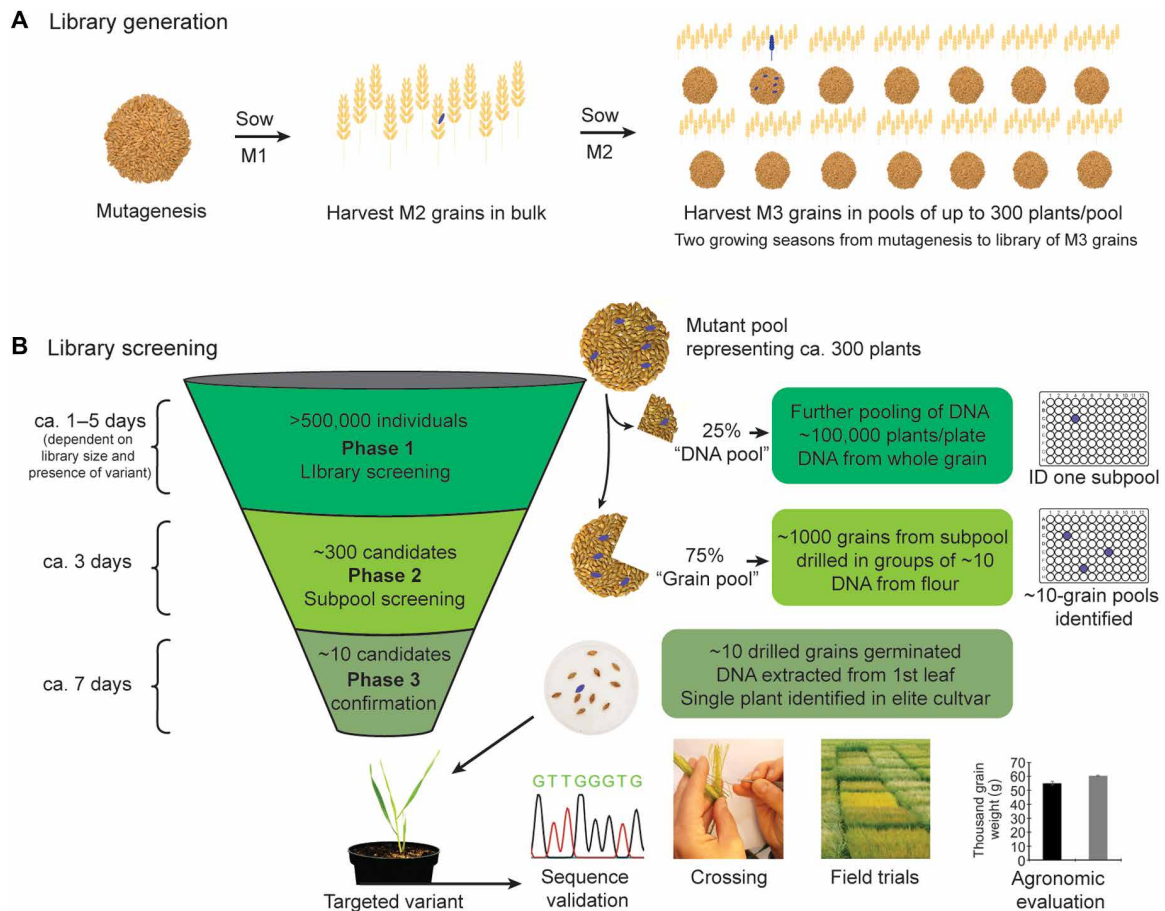
#### Mutagenesis

To make FIND-IT applicable for downstream applications and breeding, and to minimize extensive off-target loads per individual, we generated large variant libraries in barley (Fig. 1A), wheat, rapeseed, oat, yeast, and bacteria using mild treatments of inorganic azides or alkylating mutagens (tables S1 and S2). The barley variant libraries used here to demonstrate and validate the FIND-IT methodology were generated with sodium azide ( $\text{NaN}_3$ ), which is shown to primarily induce transitions (G:C-A:T) in barley (21). We aimed to achieve one to two induced nucleotide substitutions per genome megabase per plant, which is low in the context of existing barley TILLING resources (22, 23), barley natural genome variation, and the genome variation resulting from crossing events in traditional breeding [fig. S1 and note S1; (9)]. This reduces the burden of background mutations in downstream breeding applications (figs. S2 to S6 and note S2).

#### Estimation of mutation density

We documented mutation densities in our barley variant libraries (note S3) at the whole-genome level in five RGT Planet variants (from three individually designed libraries; see table S3) and in two elite line RGT Planet reference pools. The sequencing output ranged from 2.6 billion to 3.8 billion reads per sample, which equated to an average sequencing depth of more than 60-fold for each line. In our three low mutagen-treated plants (0.3 mM  $\text{NaN}_3$ ), the numbers of mutations were 4229, 7955, and 4510 (on average 5565) per line. Induced mutation rates per individual can be variable within cereal variant populations (24). Stronger mutagen treatment (1.67 mM  $\text{NaN}_3$ ) increased the rate to 14,770 mutations in average. On the basis of the RGT Planet barley genome size of 4.35 gigabases (Gb), it can be estimated that the lines generated with 0.3 and 1.67 mM  $\text{NaN}_3$  carried mutation loads of 0.8 and 2.4 induced variants/Mb, respectively (table S3). Using normalized genotyping-by-sequencing/ddRAD [(n)GBS/ddRAD] and amplicon Sanger sequencing on dozens and thousands of barley individuals, respectively, resulted in similar mutation loads (tables S4 and S5).

These analyses indicate that  $\text{NaN}_3$  preferentially induces G:C-A:T transitions (81.3%; table S6) and that mutagen-induced single-nucleotide polymorphisms (SNPs) are distributed evenly across genomes, except for the highly repetitive centromeric regions (fig. S7). The results show that to generate mutations at all transition sites in RGT Planet barley, minimum library sizes of about 864,000 and 284,000 are required for low and high mutagenized libraries, respectively (table S7). Accordingly, to reach high genome-wide mutation rates of G:C-A:T transitions, we designed our library to provide population sizes with more than 500,000 individual plants. Five G:C-A:T transitions induce premature stop codons (Trp>stop via TGG>TGA or TGG>TAG; Gln>stop via CAA>TAA or CAG>TAG; Arg>stop via CGA>TGA), and since an average barley gene encodes these amino acids multiple times [table S8, (25)], the probability to identify barley gene knockouts in such a population is high.



**Fig. 1. FIND-IT screening procedure.** (A) Barley M3 library generation: Mutagenized barley plants are grown densely to maturity in the field. M3 grains are harvested in pools of approximately 300 plants. (B) Library screening consists of three phases that narrow down the search of a target variant from >500,000 to eventually a single plant within days. See detailed description of the method in Materials and Methods and note S6.1. Screening phase 1: The harvested grain pools are each split into two fractions: One fraction (25%, DNA pool) is used for total DNA extraction, and the other (75%, grain pool) is stored for further analysis. Total DNA from DNA pools is further combined in each well, so one 96-well library plate can contain DNA from up to 100,000 individuals. The following ddPCR screening (fig. S11A), using competing fluorescently labeled TaqMan probes to target the nucleotide variant of interest, identifies one well for further analysis. Screening phase 2: The corresponding 75% grain pool of the identified DNA pool in phase 1 is selected for further analysis. DNA from approximately 1000 grains is sampled nondestructively in subpools of 10 and rescreened using ddPCR (fig. S11B), and several wells are potentially identified to contain the targeted variant. Screening phase 3: The grains from the identified wells in phase 2 are germinated, and total DNA is extracted from the first leaf of each individual plant and rescreened by ddPCR (fig. S11C). From this, a single plant is identified containing the targeted variant, which can be sequence-validated (fig. S11D). Identified variants of interest are immediately available for field trials and agronomic analyses, and incorporation into routine breeding crossing programs.

### Assessment of ddPCR sensitivity for barley library pooling

We confirmed the high sensitivity and usefulness of ddPCR in the targeted identification of gene variants by spiking a pool of 1000 wild-type barley grains with 4, 2, 1, or a half grain homozygous for a specific nucleotide exchange in the *LIPOXYGENASE-1* (*Lox-1*, HORVU.MOREX.r2.4HG0280010) gene. When targeted in the total DNA of such spiked pools, the *Lox-1* variant signal achieved by ddPCR was clearly distinguishable from the background signal in all cases (note S4 and fig. S8). On the basis of these results, barley libraries were harvested in pools containing up to 300 individual barley plants that were used for combined DNA extraction and library construction (Fig. 1A). In each pool, multiple grains will originate from the same plant since an elite barley plant develops several inflorescences with up to 30 grains each. A given induced variant will therefore be represented by many grains within a pool. This enables the pool to be split into a 25% sample (“DNA pool”) for

milling, DNA extraction, and initial large-scale genotyping, while the remaining 75% sample (“grain pool”) is stored as whole grain for subsequent specific variant identification and isolation (Fig. 1B).

As a result of the chimeric nature of the M1 plants obtained, an inflorescence with M2 grains will carry grains with different mutation profiles. This is reduced in the following generations due to segregation. Therefore, the signal strength of a given mutation will differ when working with M2 or M3 libraries (10). Accordingly, when working with M3 libraries, the ddPCR method was found to be sufficiently sensitive to enable detection of variants when DNA extracts of four pools of each 300 individual plants were combined. This enables the analysis of DNA from  $4 \times 300$  (1200) pooled plant lines in one well of the 96-well plate (Fig. 1A; see Materials and Methods). The initial pooling strategy of up to 300 plants reduced the time required for DNA isolation by up to 300-fold, compared to current screening strategies that use DNA isolates from individual plants in their libraries.

## FIND-IT uses a pool-and-split strategy coupled with high-sensitivity ddPCR

### FIND-IT screening phase 1

In phase 1, DNA representing up to 100,000 barley plants is screened on a single 96-well PCR plate. Here, 94 DNA pools (representing up to 1200 plants each), plus two negative controls, are genotyped by ddPCR for an induced nucleotide target, using a single set of primers and two competitive hydrolysis TaqMan probes that detect either wild-type or variant alleles (Fig. 1B). In our barley NaN<sub>3</sub> libraries, for example, pools harboring a premature stop variant in a gene of interest can be identified using a single set of gene-specific primers and two hydrolysis probes competing for the in-frame G to A substitution in TGG (Trp) and TGA (stop) (see table S9 with relevant gene sequences, gene-specific primers, and TaqMan assays used for identification of >150 barley variants).

### FIND-IT screening phase 2

Positive DNA pools containing the targeted variant are identified and further analyzed in phase 2 (Fig. 1B). Here, a new 96-well plate, where each well contains DNA from 10 single grains from the respective phase 1-identified grain pool, is ddPCR-genotyped to identify a positive subpool with the targeted variant. The DNA from single grains is extracted from flour sampled by nondestructively drilling into the endosperm of individual seeds (26). A single person can drill and tap out the flour from the endosperm and the drill bit from 1000 grains in about 8 hours. Thus, by combining a pool-and-split strategy with ddPCR genotyping, we can quickly narrow down the number of candidate grains carrying the targeted gene variant from 500,000 plants (in five 96-well plates) to 10 grains. Total DNA from drilled grains or drilled barley grains from phase 2 subpools can be stored safely at -20°C for several years without apparent loss in quality or germination capacity, respectively. Therefore, DNA from drilled grain subpools can be rescreened multiple times for other desired gene variants if identified in phase 1, substantially reducing subpool-generating workload.

### FIND-IT screening phase 3

The 10 grains from the positive subpool are germinated and genotyped by ddPCR and Sanger sequencing for identification and isolation of the plant carrying the desired induced gene variant. If present in the library, the FIND-IT workflow enables the identification and subsequent isolation of a desired variant within days and weeks, respectively. During phases 2 and 3, phase 1 can be reinitiated for screening of new targets and phase 3 targets can be accumulated for germination and validation (fig. S9 and note S5), further accelerating the FIND-IT workflow. For cost optimization, phase 2 and phase 3 screenings can be performed by quantitative PCR as these steps are not reliant on the high sensitivity of ddPCR.

### Validation of the method

To demonstrate FIND-IT's high capacity to target specific induced variants, we identified 100 induced premature stop codons in barley genes of our choice (Fig. 2A, table S9, and notes S6 and S7). Since our barley populations were generated with NaN<sub>3</sub>, we focused our targeted screenings on transitions that substitute tryptophan residue-encoding codons with stop codons. In all cases, we identified weak to strong variant signals in phase 1 pools and strong variant signals in the respective phase 2 subpools. Screening for other stop-inducing transitions (Gln>stop or Arg>stop) was not necessary for the genes chosen. Using gene HORVU.MOREX.r2.2HG0107310 (*Effector Of Transcription Protein*) as an example, we screened for three individual

Trp-nonsense mutations (W170\*, W371\*, and W431\*) and found all three in our very large RGT Planet library, and additionally one of three (W431\*) in our smaller cv. Paustian library (table S9 and note S7.1).

Single seeds of all 100 knockout variants (Trp>stop) were isolated, genotyped, and grown to maturity (phase 3). Seeds of these variants were stored at -20°C and are ready for further investigation (see the "Data and materials availability" statement). For method validation, we propagated and field characterized five knockout variants of genes that mediate well-documented changes in barley grain morphology (Fig. 2, B to H). See note S6.1 for additional information and characterization of the five knockout variants.

### Hull-less grain

The *Nud* gene (note S6.1) encodes a transcription factor of the ethylene response factor family that controls the covered/naked caryopsis trait in barley. Knockout of this gene (Fig. 2B) converts hulled to hull-less grain (27), making the grain suitable for human consumption (Fig. 2C). Here, we isolated a previously unknown premature stop allele of the *Nud* gene in an elite spring barley cultivar (RGT Planet) background (*nud*<sup>W16\*</sup>; Fig. 2B). Field trial evaluation of *nud*<sup>W16\*</sup> showed the expected loosely attached hull and accompanying reduced grain length, width, and thousand grain weight (TGW) after threshing and hull removal (Fig. 2, B to E). Using FIND-IT, this known domestication trait renders any wild, locally adapted or climate-resilient barley, or any elite hulled barley cultivar suitable for human consumption.

### Inflorescence row type

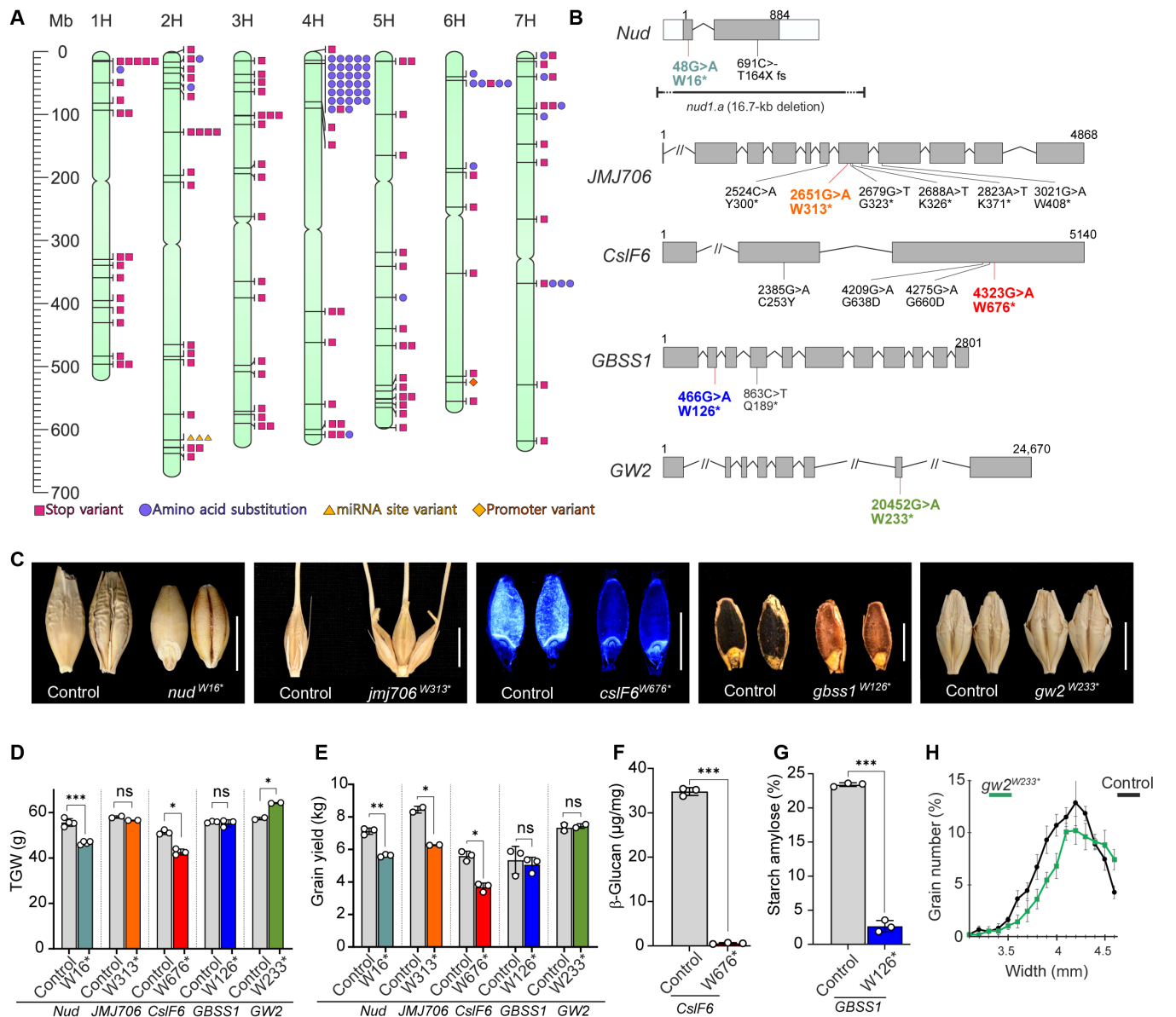
The *JMJ706/VRS3* locus (note S6.1) encodes a putative Jumonji C-type H3K9me2/me3 histone demethylase involved in defining the barley inflorescence row type, a yield-related domestication trait controlled by several genes (28). Here, we isolated a homozygous loss-of-function variant in the *JMJ706* gene in the two-row RGT Planet cultivar (*jmj706*<sup>W313\*</sup>; Fig. 2B) that has the reported barley inflorescence phenotype with enlarged lateral spikelets, promoting uniform grain size and increased total yield in six-rowed barleys (Fig. 2C) (28).

### Grain (1,3;1,4)-β-glucan

The *Cellulose synthase-like* gene *CsIF6* (note S6.2) is a key determinant controlling the biosynthesis and structure of (1,3;1,4)-β-glucan in the barley grain, making it a prime target for lowering the content of (1,3;1,4)-β-glucan (29), a preferred grain trait for the distilling and brewing industries (30). We identified and isolated a *CsIF6* knockout variant (*csIF6*<sup>W676\*</sup>; Fig. 2B) in RGT Planet with undetectable levels of (1,3;1,4)-β-glucan content in the grain (Fig. 2, C and F). Field testing for agronomic performance of *csIF6*<sup>W676\*</sup> confirmed previously published greenhouse experiments on *CsIF6* knockouts, namely, that TGW and grain width were reduced and changed (Fig. 2, D to F, and note S6.2) (31). Here, we show that total grain yield of *CsIF6* knockout variants grown in the field is reduced by more than 30% (Fig. 2E), limiting the agronomic and industrial applicability of *CsIF6* knockout varieties.

### Waxy grain

Granule-bound starch synthase I (GBSSI) (note S6.1) is the key enzyme for amylose synthesis in barley (32). We identified a variant carrying a loss-of-function allele (*gbss1*<sup>W126\*</sup>; Fig. 2B) that features the characteristic waxy barley grain phenotype and the underlying reduction of amylose content in the grain starch (Fig. 2, C and G), with no significant yield and TGW penalties (Fig. 2, D and E). Waxy starches are of high value in many processed foods, as the amylose content directly influences the texture of cooked starches (33).



**Fig. 2. Barley knockout library with selected variants isolated and agronomically evaluated.** (A) Barley physical reference genome map (chromosomes 1H to 7H) with location of more than 150 identified FIND-IT variants (variant type indicated by symbol; table S9). (B) Gene models of selected barley reference targets with known loss-of-function mutations and previously unknown, isolated stop variants (in color). (C) Loss-of-function phenotypes of isolated variants [hull-less grain, *nud*<sup>W16\*</sup>; six-rowed spikelets, *jmj706*<sup>W313\*</sup>; (1,3;1,4)-β-glucan Calcofluor staining, *csIF6*<sup>W676\*</sup>; starch amylose Lugol staining, *gbss1*<sup>W126\*</sup>; grain width, *gw2*<sup>W233\*</sup>]. (D) TGW and (E) grain yield of field-grown barley variants. (F) Grain (1,3;1,4)-β-glucan content of *csIF6*<sup>W676\*</sup> grain. (G) Starch amylose content of *gbss1*<sup>W126\*</sup> grain. (H) Grain width distribution of *gw2*<sup>W233\*</sup> grain. Scale bars, 5 mm (C). Error bars are SD; two-tailed *t* test was performed to obtain *P* values (\**P* < 0.05, \*\**P* < 0.01, and \*\*\**P* < 0.001; ns, not statistically significant). See table S12 for statistical test details and number of replicates.

Using the FIND-IT approach, it will be possible to isolate waxy variants of all common cereals.

### Grain width

*Grain Width 2* (*GW2*) is a modulator of grain width and weight in rice (note S6.1) (34). To illustrate the option of translating agronomic traits from one cereal to another, we introduced the wide grain trait into barley (34) by isolating a loss-of-function allele in the elite malting barley cultivar Paustian (*gw2*<sup>W233\*</sup>; Fig. 2B). To date, no knockout of barley *GW2* has been identified or studied (35). Initial field yield

trials showed that *gw2*<sup>W233\*</sup> has increased TGW, caused by wider grain, as previously shown in rice. However, the total yield per area was not increased (Fig. 2, C to E and H). Knockout variants of *OsGW2* improve grain yield in rice via acceleration of the grain milk filling rate, a trait of potential importance for cereals grown in suboptimal growth conditions such as drought and heat stress associated with shortened grain filling periods. The barley *gw2*<sup>W233\*</sup> variant can now be field-tested in diverse climatic environments to identify its potential for yield improvement and climate adaptation in barley and other cereals.

### Beyond barley knockout variants

In addition to knockout lines, FIND-IT enables the identification and isolation of missense mutations and mutations in noncoding and intragenic regions that fine-tune transcription and hence gene activity. To demonstrate this, we targeted more than 50 genetic variants with single-base changes, either for changing transcript levels or for altering the amino acid sequences of encoded proteins (Fig. 2A, table S9, and note S7.1). Here, single seeds of 25 variants were isolated, propagated, and stored at  $-20^{\circ}\text{C}$  for further investigation (see the “Data and materials availability” statement).

### Nitrogen-use efficiency

In one example, we targeted the GROWTH-REGULATING FACTOR 4 (GRF4) transcription factor and the DELLA protein (SLENDER1/SLN1) that promote and repress growth, respectively, as well as influencing nitrogen assimilation and carbon fixation (note S6.2) (36). In rice green revolution varieties, enhanced nitrogen use efficiency and grain yield are achieved by increasing *GRF4* transcript abundance through miRNA396 interference and by reducing DELLA activity (36). In barley green revolution varieties RGT Planet and Paustian, we identified 31 variants of *SLN1/DELLA* with specific amino acid exchanges (Fig. 3A and table S9). One variant, *sln1*<sup>A277T</sup> (Fig. 3, B to D and note S6.2), showed normal plant height, increased seed length, and increased TGW. Next, we identified three variants targeting the miRNA396-binding site of *GRF4* (Fig. 3E) with the aim of increasing GRF4 activity by reducing miRNA-mediated transcript repression. Such genetic variants of *GRF4* and *SLN1/DELLA* identified in modern green revolution varieties enable crosses and field trials aimed at generating new elite barley varieties with enhanced nitrogen use efficiency and hence reduced requirements for nitrogenous fertilizers (36). MiRNA-resistant alleles of the *HOMEODOMAIN-2 (HB-2)* gene in wheat have recently been shown to result in increased *HB-2* transcript levels, further highlighting the usefulness of targeting single-base changes in miRNA-binding sites to fine-regulate gene transcript abundance (37).

### Germination vigor

In another example, we targeted the promoter of the barley  $\alpha$ -amylase gene to enhance gene expression and germination vigor (Fig. 3F). In germinated RGT Planet grain, expression of the *Amy1\_2* gene exceeded the expression of genes in the *Amy1\_1* or *Amy2* clusters at 72 hours after germination (Fig. 3G). Inspired by the observed natural variation in the tandem repeat W-box/O2S binding domain in  $\alpha$ -amylase promoters regulating the repression via WRKY38 (38), we used FIND-IT to identify and isolate a previously unknown allele of *Amy1\_1* (*Amy1\_1-var*). This variant has a single-nucleotide substitution in its promoter that changes one TGAC core motif to TGAT (Fig. 3F). We then demonstrated that the overall *Amy1\_1* transcript level in this RGT Planet variant *Amy1\_1-var* increased 72 hours after initiation of germination by close to 100% relative to the control (Fig. 3H). This demonstrates that changing the TGAC core motif to TGAT releases the repression of *Amy1\_1* expression by WRKY38. This result specifically demonstrates the function of the tandem repeat W-box/O2S binding domain in  $\alpha$ -amylase promoters in vivo and suggests a direct way to fine-regulate the levels of this industrially important enzyme. Our example highlights the possibilities of modifying gene transcript levels by targeting promoter domains, both for gene function analyses and for specific breeding purposes. See note S6.2 for additional information and characterization of distinct nonsynonymous nucleotide variants and variants that modulate transcript abundance.

### Malting

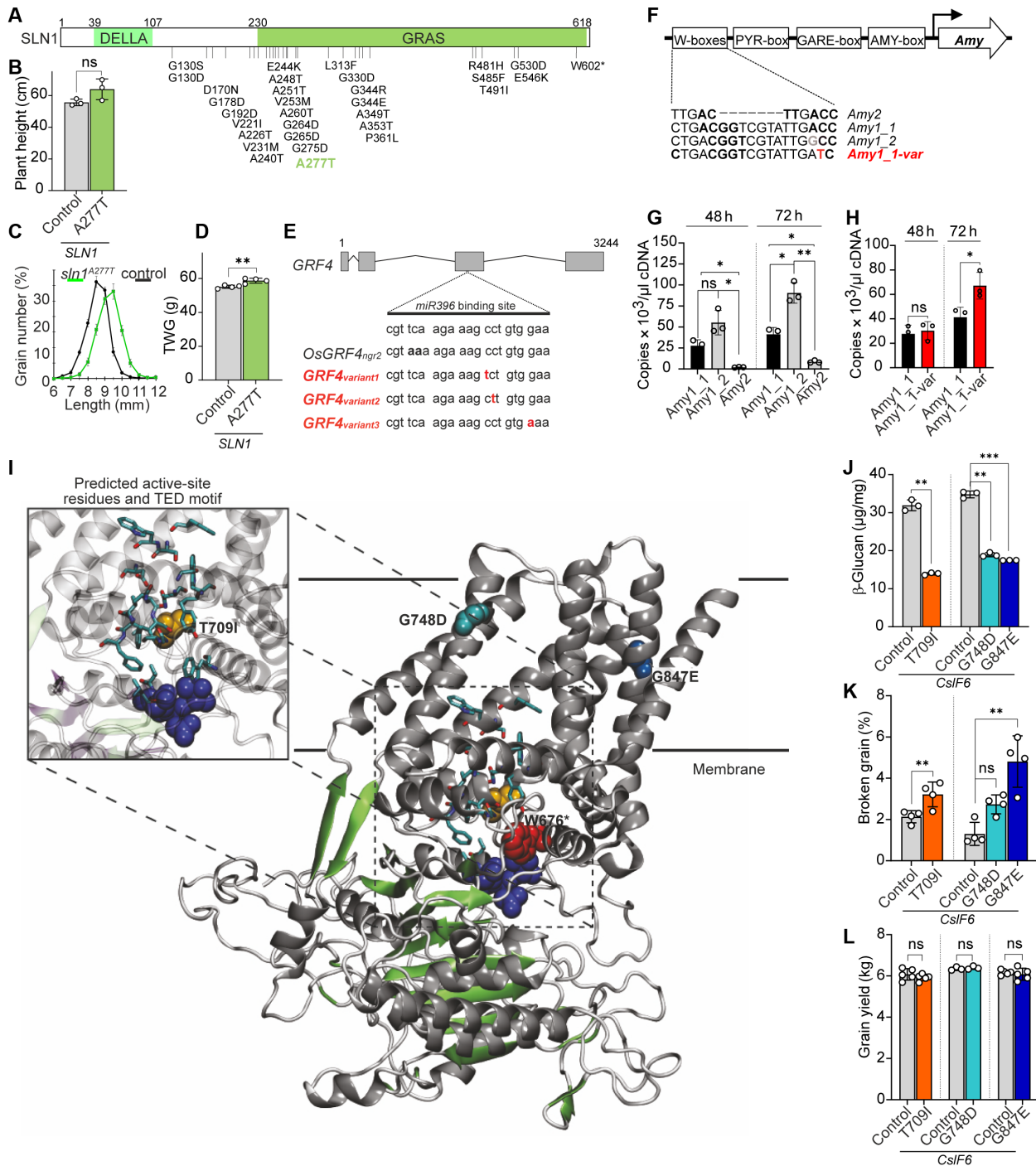
To isolate variants with an intermediate (1,3;1,4)- $\beta$ -glucan content phenotype and to overcome the adverse yield and grain morphology phenotypes of the *CsIF6* knockout line (Fig. 2, C to E), we used FIND-IT to identify and isolate three additional *CsIF6* variants with amino acid exchanges close to or within *CsIF6* transmembrane helices (T709I, G748D, and G847E) (Fig. 3I and note S6.2). The *CsIF6* variant plants obtained have a 50% reduction in grain (1,3;1,4)- $\beta$ -glucan levels but at the same time maintain yields; for *csIF6*<sup>G748D</sup> and *csIF6*<sup>G847E</sup>, TGW is also maintained (Fig. 3, J and L, and note S6.2). While grains of variants *csIF6*<sup>G847E</sup> and *csIF6*<sup>T709I</sup> are susceptible to breakage in hard mechanical threshing (Fig. 3K), variant *csIF6*<sup>G748D</sup> is proving valuable in malting and brewing processes, where high (1,3;1,4)- $\beta$ -glucan causes filtration difficulties and undesirable hazes in the final product (30). In less than 4 years, the *csIF6*<sup>G748D</sup> variant has been crossed with elite varieties, registered as a new barley cultivar (CB Celina) and propagated to approximately 100 metric tons of grain for commercial pilot trials, showing the speed at which FIND-IT variants can be deployed in the commercial marketplace.

### Designing and screening libraries of industrially important microorganisms and other crops

The broad applicability of FIND-IT makes it possible to screen large variant libraries of any living organism that can be grown in the field or in culture, as highlighted here by additional examples in wheat, yeast (*Saccharomyces cerevisiae*), and bacteria (*Lactobacillus pasteurii*) (note S6.3). Diverse libraries of rapeseed and oat were also designed (table S1). For microbial applications, the procedure is performed entirely in the liquid phase. For example, with the identification of valuable microbial variants, mutagenized cell suspensions can be aliquoted directly into 96-well plates, ddPCR can be used to identify wells with the targeted mutation, and the cell of interest can be isolated through progressive dilution and ddPCR detection.

### DISCUSSION

Our studies document that FIND-IT represents a methodology facilitating the rapid development of resilient crop plants and improved microbial production systems (Fig. 1). FIND-IT is a new species- and variety-agnostic approach for targeted isolation of single-nucleotide variants in large genome pools and relies on three crucial components: (i) the generation or the existing availability of a variant library, such as libraries induced by chemical mutagenesis, and (ii) high-sensitivity ddPCR technology that allows (iii) a sample pool-and-split strategy, which enables the pooling of 300 or more individuals before DNA isolation. On the basis of the integrated use of these three components, FIND-IT allows the size of routinely screenable variant libraries to be expanded beyond 500,000 individuals. This enables variant pool generation toward single-nucleotide resolution. The size of the library and its mutation load and type are inversely correlated with respect to the probability of finding the variant of interest. Thus, for functional analysis of individual genes, a relatively small library with a higher mutation load can be screened and, where direct incorporation of variants into a breeding pipeline is required, a much larger library with a relatively low mutation load can be generated. FIND-IT therefore allows these library size and mutation load parameters to be independently adjusted to meet the intended application of the library. Although variants generated by FIND-IT are partially limited by the chosen mutagen used, we



**Fig. 3. FIND-IT libraries with distinct nonsynonymous nucleotide variants and variants to modulate transcript abundance for precision breeding.** (A) Identified and isolated variants (in gray and green, respectively) of the SLN1/DELLA protein. (B) Plant height, (C) grain length distribution, and (D) TWG of *sln1*<sup>A277T</sup>. (E) Gene model of barley *GRF4* showing the *miR396*-binding site including a known double substitution in rice (*OsGRF4*<sup>ngt2</sup>, bold) and FIND-IT barley variants (in red). (F) Known cis-acting elements in  $\alpha$ -amylase promoters with natural and induced W-box sequence variation: W-boxes with two TGAC core motifs, a natural variation in the *Amy1\_2* gene in gray and an induced variation in an *Amy1\_1* gene in red. (G) Expression of  $\alpha$ -amylase genes in RGT Planet grain 48 and 72 hours after initiation of germination. (H) Expression of *Amy1\_1* genes in RGT Planet grain (black) and RGT Planet variant *Amy1\_1-var* (red) 48 and 72 hours after initiation of germination. (I) Protein homology model of CslF6 based on poplar cellulose synthase isoform 8 (Protein Data Bank ID: 6wlv.1.A). Amino acid positions at residue W676, T709, G748, and G847 are shown as spheres. Inset shows uridine diphosphate (UDP)–glucose binding sites as light blue sticks and the TED motif as dark blue sticks. The inset is shown at a slightly tilted angle to visually optimize the position of the residues discussed. (J) Grain (1,3;1,4)- $\beta$ -glucan content of controls (gray bars) and *CslF6* variants. (K) Percentage of broken grain after threshing in controls (gray bars) and *CslF6* variants. (L) Field grain yield of controls (gray bars) and *CslF6* variants. Error bars are SD; two-tailed *t* test was performed to obtain *P* values (\**P* < 0.05, \*\**P* < 0.01, and \*\*\**P* < 0.001). See table S12 for statistical test details and number of replicates.

found that, in barley, about 80% of mutations were of the transition class predicted for  $\text{NaN}_3$  mutagenesis. Alternative mutagens, such as methyl methanesulfonate, are known to primarily induce purine-pyrimidine transversions in barley (18) and could be used to enhance the flexibility of FIND-IT for tailor-made library design.

In our current study, we have successfully initiated and validated the FIND-IT approach using the cereal crop barley and selected industrial microorganisms. We silenced targeted genes that are linked with readily observable phenotypic characteristics, such as hulled versus hull-less barley, two-row versus six-row barley, and the waxy, low-amylose phenotype. It was shown that a grain width phenotype found in rice could be transferred to barley. Targeting individual amino acids in the barley (1,3;1,4)- $\beta$ -glucan CslF6 enzyme resulted in reduced levels of the polysaccharide in grain, while altered amino acids in proteins and miRNA-resistant gene alleles associated with nitrogen use efficiency provide a potential route to the reduction of nitrogenous fertilizers in global agriculture. Last, we targeted a promoter element of  $\alpha$ -amylase genes and showed that expression levels could be manipulated (Figs. 2 and 3). These examples demonstrate the species-independent potential of FIND-IT for gene knockout experiments and gene up- and down-regulation. Its speed and flexibility make the FIND-IT approach applicable across most species. As an example of the speed with which FIND-IT variants can be incorporated into commercial varieties, we have developed a new barley cultivar and propagated it to approximately 100 metric tons of grain for commercial pilot trials within less than 4 years of the FIND-IT identification of the variant.

FIND-IT has advantages and disadvantages when compared with existing methods for the generation and screening of variants in plants and other species. In contrast to transgenesis and CRISPR-Cas9 technologies, the insertion of complete genes, gene replacement, or the alteration of sections of genomes is not possible using FIND-IT. Thus, plant genome editing by CRISPR-Cas9 and transgenesis technologies remains of central importance for the long-term future of crop adaptation (39). However, CRISPR-Cas9 technologies currently have technical drawbacks and uncertainties that have so far precluded their widespread application in commercial plant breeding (5, 6). These include a requirement for transformation and tissue culture steps that can adversely affect genome integrity and cause undesired somaclonal variations (40). This leads to the noncompatibility with most breeding pipelines because existing elite, high-yielding, climate-resilient, and regionally adapted breeding lines are often not amenable to transformation. The use of nonelite varieties that are transformable requires lengthy prebreeding efforts and comes with a risk of large natural off-target variation (note S2). Emerging methods of de novo meristem induction may overcome these problems in the future (41). The generation and identification of single-nucleotide changes that form the basis of FIND-IT are still challenging for CRISPR in plants, where deletion mutations predominate (5). Persistent problems with off-target edits (42) have also proved to be challenging, and in several jurisdictions, lines generated by CRISPR-Cas9 currently fall under GM guidelines (43) and are therefore subject to the regulatory and financial barriers associated with releasing GM crops. The latter also applies to new crop varieties generated through transgenesis.

Other targeted genome screening methods like TILLING (14, 15, 44, 45) or recent TILLING-by-sequencing approaches that integrate next-generation sequencing with the latest capture methodologies (12, 16, 22, 23, 46) provide genome-wide scanning of induced and natural variant populations directly at the genotype level.

However, technical, practical, and financial hurdles have limited the widespread introduction of TILLING lines into commercial varieties. Current restrictions for application of TILLING include the (i) necessity for single-plant DNA extractions, (ii) progressively increasing costs in screening libraries with decreasing mutation density, (iii) species- or accession-specific exome capture design and exome or amplicon sequencing, including lengthy time gaps between bioinformatic processing and pipeline establishment, and (iv) screenable population sizes that are generally fewer than 10,000 individuals (16, 22, 46, 47). The limited number of available gene variants prevents single-nucleotide resolution and the discovery of desired variants. This issue can be offset, to some extent, by developing libraries with very high mutation rates, but this results in high off-target background mutation loads that are undesirable in downstream breeding applications (17).

In summary, FIND-IT releases the bottlenecks of traditional phenotypic selection and TILLING while standing together with evolving CRISPR-Cas9 technologies for the efficient improvement of today's germplasms. For breeding and fast commercial rollout, FIND-IT has an important advantage, insofar as it enables species and cultivar flexibility, low off-target mutation loads that circumvent lengthy crossing protocols, and potentially rapid translation of variants into commercial cultivars.

In the future, FIND-IT opens up new opportunities for crop improvement by providing a means to select and incorporate domestication traits into the approximately 7000 undomesticated or semidomesticated crop plants that have superior drought tolerance, water tolerance, disease resistance, and mineral use efficiency, including perennials or wild crop relatives (48, 49). It also has potential for the introduction of valuable gene alleles found in exome collections and pan-genomes (9, 50, 51) and improved variants designed and identified in vitro. Thereby, FIND-IT provides a long-sought tool that will foster a green evolution enabling us, e.g., to meet the food demands of the future.

## MATERIALS AND METHODS

### Generation of FIND-IT libraries

A range of FIND-IT libraries was generated by chemical mutagenesis of elite varieties of barley, wheat, rapeseed, oat, as well as yeast (*S. cerevisiae*) and a bacterium (*L. pasteurii*; DSM 23907). Here, we describe in detail the preparation of the barley libraries, and while the methods for preparing libraries from other species were similar, key differences are presented in table S1.

Barley grains (2 to 3 kg) were presoaked in water for 16 hours at 4°C, drained, and subjected to mutagenesis by soaking in 0.3 or 1.67 mM sodium azide ( $\text{NaN}_3$ ) at pH 3.0 for 2 hours (52). After thoroughly washing with water, the grains were air-dried in a fume hood for 16 hours. The mutagenized M1 grains were sown at high density (600 plants/m<sup>2</sup>) to reduce tillering and grown to maturity under field conditions. At maturity, M2 grains were harvested in pools of approximately 300 plants (0.5 m<sup>2</sup> per pool). Grains from each pool were threshed, cleaned, and used to generate a low-zygosity M2 library (containing M2 grains).

We initially built M2 libraries for developing the FIND-IT method. Working with M2 grains significantly speeds up library construction, and a fully screenable library can be available within one growing cycle. However, all cells are susceptible to mutagenesis in the seed embryo, but only a few cells in the embryo's shoot meristem will



contribute to seed heredity. These “germline cells” are independently mutagenized, giving rise to a chimerical plant, descending into several independent variant lines (53).

Therefore, bulk harvested M2 grains were used to also generate a high-zygosity M3 library the following season. The M2 grains were again sown at high density (600 plants/m<sup>2</sup>) and field-grown to maturity. At maturity, M3 grains were harvested in pools of approximately 300 plants (0.5 m<sup>2</sup> per pool). Grains from each pool were threshed and cleaned individually, avoiding carryover between pools, and used to generate an M3 library (containing M3 grains).

M2 or M3 library grains from each pool of up to 300 plants (200- to 600-g grain) were split using Perten Sequential Divider SPD4200 (Perkin Elmer, Waltham, MA, USA) to obtain a representative 25% sample of grains from the pool for DNA extraction (DNA pool). The representative 25% sample was milled (Retsch GM200, Haan, Germany), and DNA was extracted from 50 g of the flour by LGC Genomics GmbH (Berlin, Germany). DNA concentrations were adjusted to 150 ng/μl for each sample extract and distributed in 96-well library plates that each represented approximately 25,000 to 30,000 individual variant plants. After DNA extraction, all M3 libraries were further pooled fourfold to represent 1200 plants per well and approximately 100,000 individual plants per 96-well plate. The remaining 75% sample (grain pool) was stored at 4°C until positively identified in a DNA library screen for further screening. Currently, twenty 96-well plates in our library, representing a range of barley cultivars, generations, and mutagenesis treatments, are available for screening (tables S1 and S2).

### Determination of mutation load

Mutation load in the barley libraries was measured at three levels. In a first approach, mutation loads were evaluated at the whole-genome level using the Chromium linked-reads technology (10x Genomics, Pleasanton, CA, USA). Whole-genome sequencing was performed at BGI Tech Solutions Co. Ltd. (Hong Kong, China) and comprised two wild-type RGT Planet pools with around 10 plants each and five randomly selected RGT Planet variants derived from different generations (M2 and M3) and different mutagen treatments (0.3 and 1.67 mM NaN<sub>3</sub>; table S3). Briefly, high-molecular weight DNA was extracted from frozen leaf samples as described elsewhere (54) followed by 10x Genomics Chromium library construction and Illumina PE150 sequencing. Sequencing output ranged from 2.6 billion to 3.8 billion reads per sample. Reads were mapped to the chromosome scale assembly of RGT Planet (9) with Long Ranger 2.2.2 (<https://support.10xgenomics.com/genome-exome/software/pipelines/latest/what-is-long-ranger>), followed by variant calling with GATK v3.7 (55). For further analyses and mutation rate estimation, we focused on SNPs since NaN<sub>3</sub> primarily induces point mutations (21). For each RGT Planet variant, we removed SNPs detected in wild type and those found in two or more RGT Planet variants using BCFtools v1.9 (56). SNPs were further filtered as follows: (i) SNPs with variant confidence/quality by depth < 2.0 or Phred-scaled p-value using Fisher’s exact test to detect strand bias > 60 or root mean square of mapping quality < 40 or MQRankSum < -12.5 or ReadPosRankSum > -8.0 or Symmetric Odds Ratio of 2×2 contingency table to detect strand bias (SOR) > 3 were removed, (ii) SNPs with read depths in total less than 10 or supporting reads for alternate allele less than 3 or alternate fraction lower than 0.3 were removed, and (iii) only 1 SNP and 0 indel were allowed within 100 base pairs (bp), i.e., upstream and downstream 50 bp of the SNP. The last step was manually validated by randomly sampling 120 SNPs from

four variants and visual inspection in IGV 2.12.2 (57). In all cases, only false positives were removed, and the average accuracy was increased from 75 to 85%.

In a second approach, mutation load was evaluated at a sparse genome-wide level using (n)GBS/ddRAD. This experiment was performed using 48 plants of RGT Planet, including 8 wild-type plants, 20 M3 variant progeny plants derived from 0.3 mM NaN<sub>3</sub> treatment, and 20 M3 variant progeny plants derived from 1.67 mM NaN<sub>3</sub> treatment. The (n)GBS/ddRAD analysis, including DNA extraction, library construction, sequencing, and initial bioinformatic analysis, was performed at LGC Genomics GmbH (Germany). Briefly, genomic DNA (gDNA) was extracted from dried leaf material, and the restriction enzymes Pst I and ApeK I were used for complexity reduction. Samples were sequenced on an Illumina sequencer (Illumina NextSeq 500/550 v2), and approximately 3 million single-end reads with a length of 75 bp per sample were obtained. Read preprocessing was performed using the Illuminabcl2fastq 2.20 software and demultiplexed, and cleaned reads were aligned to the Morex\_V2 reference genome using the BWA-MEM software version 0.7.12 (58), followed by SNP calling using Freebayes v1.0.2-16 (59). To lower false-positive rates, the following filters were applied: (i) removal of loci with coverage depth lower than eight reads in any of the samples, (ii) removal of samples that were outliers based on Student’s *t* test ( $P < 0.01$ ) on genetic distance matrices (60) generated for all plants with similar NaN<sub>3</sub> treatment, (iii) removal of loci that had any neighboring loci within 50 bp, (iv) disregarding loci whose alternate fraction is lower than 25%, and (v) disregarding loci that were detected in two or more samples. Mutation rates for each fragment were estimated and are given in table S4.

In a third approach, 464- to 497-bp amplicons of four selected genes from a barley cv. Quench library (0.3 mM NaN<sub>3</sub>) were sequenced to identify polymorphisms and then resequenced to confirm the recorded differences. The barley variant population of cv. Quench was prepared as described above with the conditions shown in tables S1 and S2. The variant population was field-grown, and 12,000 single spikes of the M2 generation were harvested. A single grain (M3) per spike was isolated for further gDNA extraction and single-read Sanger sequencing of target gene sequences (Eurofins Genomics Germany GmbH, Ebersberg, Germany). The chosen sequencing targets were single fragments of the genes HORVU.MOREX.r2.7HG0553990 (*Amino Acid Permease3*, *APP3*) and HORVU.MOREX.r2.6HG0462530 (*No Apical Meristem1*, *NAM1*) as well as two individual fragments of HORVU.MOREX.r2.4HG0278150 (*Iron-Regulated Transporter 1*, *IRT1*) using the amplification primers listed below. In total, *APP3*, *NAM1*, and two *IRT1* fragments were sequenced in 6144, 6048, 5952, and 4704 individual variant grains, respectively. Here, 5763 (*APP3*), 5551 (*NAM1*), 5446 (*IRT1\_Ex1*), and 3731 (*IRT1\_Ex2*) sequences were of appropriate quality for detailed analysis (table S5).

A certain level of heterozygosity can be expected in gDNA derived from M3 generation barley seeds. Thus, for the sequence analysis of each fragment, the sangerseqR package (61) was used to capture heterozygous bases with a “makeBaseCalls” ratio set at 0.8. Consensus sequences from “Primary Basecalls” and “Secondary Basecalls” were built for each fragment at the target region using Biostrings (62), where potential heterozygous mutation bases were coded according to IUPAC (International Union of Pure and Applied Chemistry) codes. Consensus sequences shorter than median length were excluded. Multiple sequence alignment was performed using the

MAFFT v.7 software (63). Alignments were read with seqinr (64) and converted into a matrix from which 500 bp of sequence was extracted (300 bp from upstream and 200 bp from downstream of the mid-base). Positions with nonunique bases were reported as mutation sites for further confirmation by manually checking chromatograms. Manually confirmed mutations are listed in table S5.

In each confirmed case, four additional grains from the respective candidate spike were collected and germinated on petri dishes. gDNA from seedling leaves was extracted, and the respective gene fragment was PCR-amplified using the REExtract-N-Amp Plant PCR Kit (Sigma-Aldrich, St. Louis, MO, USA), according to the manufacturer's instructions. PCRs were performed for 35 cycles (initial denaturation at 94°C/3 min followed by 35 cycles of 94°C/30 s, 59°C/30 s, and 72°C/60 s for extension, with a final extension step of 72°C/10 min) using the following primers: HvAAP3\_Ex6\_F1 (CGIGTACCTGGAAATGCAGG) and HvAAP3\_Ex6\_R1 (GGATTTCCGGCC TTCCAAGTG), LU\_HvNAM1\_For2 (GTCTCATCGATCAGTTG-GACG) and LU\_HvNAM1\_Rev2 (GTGTCATTCGTTTCAGGGAT-TCC), HvIRT1\_Ex1\_F (AATCGCAACCAAAAGATCGAGC) and HvIRT1\_Ex1\_R (GCAACACAAGATTACCTGAACG), and HvIRT1\_Ex2\_F (GTCTAGTTCCCGTGTTCATG) and HvIRT1\_Ex2\_R (AGACACACCCTCATCAACCATA). PCR products were purified using the NucleoSpin Gel and PCR Clean-Up Kit (Macherey-Nagel GmbH & Co. AG, Düren, Germany) according to the manufacturer's instructions for Sanger sequencing (Eurofins Genomics Germany GmbH, Ebersberg, Germany). Mutation rates for each fragment were estimated and are given in table S5.

### Analysis of naturally occurring genome variation in domesticated and wild barley of the barley pan-genome

Data of RGT Planet presence/absence variations versus Morex\_V2 were extracted from the barley pan-genome (9) and illustrated with Circos 0.69.8 (65). The SNP matrix from 200 domesticated and 100 wild barley lines published as part of the barley pan-genome was retrieved from IPK (SNP\_matrix\_WGS\_300\_samples.vcf.gz, <https://doi.ipk-gatersleben.de/DOI/c4d433dc-bf7c-4ad9-9368-69bb-77837ca5/1b898e9a-11bf-4eb6-b165-e1fa2167f0ec/0>) (9). Summary statistics for each sample was calculated by bcftools stats (1.10.2+htslib-1.10.2-3) (54) with the command `bcftools stats -s - SNP_matrix_WGS_300_samples.vcf.gz`.

### Identification of structural variations in the chromosome 7H inversion region

The structural variations were detected using the SyRI pipeline (66) as described in (67).

### Method sensitivity validation

The FIND-IT approach is based on ddPCR technology, which is a highly sensitive technique for identifying rare variants in high-background samples (19). To validate the sensitivity of ddPCR for identification of single-nucleotide variation in grain samples, 1000 wild-type grains (cv. Barke) were milled with 0.5, 1, 2, 3, or 4 grains (cv. China) containing a single-nucleotide mutation in the *Lox-1* gene, HORVU.MOREX. r2.4HG0280010, followed by extraction of total DNA to be analyzed using ddPCR (note S4).

The ddPCR assay reactants include competing fluorescently labeled TaqMan probes. The wild-type probe is labeled with the hexachlorofluorescein (HEX; green) fluorophore, and the "variant" probe is labeled with 6-carboxyfluorescein (FAM; blue) fluorophore. The

TaqMan PCR probes and primers were designed using the Bio-Rad design tool ([www.bio-rad.com](http://www.bio-rad.com)), with the two probes distinguishing between two different nucleotides at the same location in a targeted gene of interest. The reaction mixture for the ddPCR assay consisted of 2× QX200 ddPCR Supermix for probes (no dUTP) (Bio-Rad, Hercules, CA, USA), 900 nM/250 nM of 20× target primers/probe (FAM for variant, HEX for wild type), 5 µl of diluted DNA, and molecular-grade water added up to a total reaction volume of 20 µl. *Lox-1* probes were designed to specifically distinguish between the cv. Barke and cv. China alleles (*lox1*-fwd: GCCATTGAGCAG-TACG; *lox1*-rev: CATGGGGCGTCCTTG; *lox1*\_G\_hex: AGGC-GTGGTGGAAAGG; *lox1*\_A\_fam: AGGCGTGATGGAAGGA). ddPCR was performed as described below (FIND-IT screening phase 1).

### Screening of FIND-IT barley libraries for specific nucleotide substitutions

In phase 1 of the FIND-IT procedure, hundreds of thousands of plants can be screened in 96-well library plates. Each well contained DNA extracted from the 25% DNA pools, where each pool represents up to 300 individual plants. However, the increased ddPCR signal strength of the M3 grain pools allowed us to further pool four DNA extracts, thus enabling the screening of DNA representing up to 1200 individual plants in a single well for M3 libraries. Library plates were screened until a certain well showed an increased abundance of the nucleotide of interest. This pool was advanced to phase 2 in the screening process (Fig. 1B).

For phase 1 screening, a master mixture of the TaqMan probe solution, distinguishing between specific nucleotides, and the ddPCR reagents was prepared and mixed with diluted DNA from a library DNA plate in a 96-well PCR plate. The reaction mixture was prepared on a QX200 AutoDG Droplet Digital PCR system (Bio-Rad, Hercules, CA, USA), where each 20 µl of reaction mixture is separated and encapsulated into approximately 20,000 droplets using an oil film. Each droplet encases, on average, one genome equivalent of DNA, and droplets are dispensed into 96-well plates for ddPCR analysis. The PCR plate containing the droplets was heat-sealed at 180°C for 5 s with pierceable foil, using a PX1 PCR plate sealer (Bio-Rad, Hercules, CA, USA) followed by PCR amplification (Uno96, VWR, Radnor, PA, USA) as follows: enzyme activation at 95°C/10 min followed by 40 cycles of denaturation at 94°C/30 s followed by annealing/extension at 55°C/1 min ending with enzyme deactivation at 98°C/10 min, all steps with a ramp of 2°C/s. During the PCR, the fluorophore is displaced from the oligonucleotide and quencher, and its fluorescence is unmasked. The droplet PCR plate was placed in a QX200 Droplet reader (Bio-Rad, Hercules, CA, USA) to count positive and negative droplets for each fluorophore, and data were subsequently analyzed using Bio-Rad QuantaSoft software (Bio-Rad, Hercules, CA, USA), based on the amplitude and concentration of fluorophore signal for each probe.

In phase 2, about 1000 grains were randomly selected from the stored 75% grain pool (approximately 5000 remaining grains) of the positively identified DNA pool to further identify and isolate the variant grain of interest. Flour from single grain was sampled non-destructively by drilling a 1-mm hole approximately 2 to 3 mm into the endosperm (Marathon-3 Champion, Saeyang Microtech, Daegu, Korea) and collecting the flour emerging from that hole on weighing paper [about 1.0 to 1.5 mg per grain; (26)]. The flour from 1000 grains, in pools of ~10 grains, was distributed into a single deep 96-well plate, and DNA was extracted (NucleoSpin Plant II, Mini kit, Macherey

Nagel, Düren, Germany). An identical plate containing the small sub-pools of ~10 drilled grains was stored at 4°C until further analysis or at –20°C for long-term storage. The ddPCR was repeated, as described above, using the same TaqMan assay as in the phase 1 library screen. From the “flour” plate, positives for the nucleotide substitution of interest were detected with a fractional abundance of 5 to 10%, depending on zygosity. Thus, the phase 2 screen narrowed down the potential candidates to ~10 grains (Fig. 1B).

In phase 3, all ~10 grains from the corresponding positive well were germinated on wet filter paper on a petri dish at 15°C. Germinated grains were transferred to soil after 4 days, and DNA was extracted (DNeasy, Plant Mini Kit, Qiagen, Hilden, Germany) from the first emerging leaf of individual plants after 5 to 7 days. DNA from the ~10 plants was then analyzed by repeating the initial ddPCR procedure using the same TaqMan assay as in the original phase 1 library screen. This phase 3 screen identified one variant plant that contained the nucleotide substitution of interest (Fig. 1B). The variant carrying the nucleotide substitution of interest was transferred to a large pot, and DNA was re-extracted from a random leaf for verification of the targeted nucleotide substitution using both the original ddPCR procedure and Sanger sequencing.

### Sequence validation of isolated barley variants

A single grain from each identified and isolated variant was germinated, and gDNA was extracted from the first emerging leaf (DNeasy, Plant Mini Kit, Qiagen, Hilden, Germany). PCRs were performed with the respective primers listed in table S10 for 35 cycles (initial denaturation at 94°C/3 min followed by 35 cycles of 94°C/30 s, 61°C/30 s, and 72°C/45 s for extension, with a final extension step of 72°C/10 min). Appropriate PCR products were purified using the NucleoSpin Gel and PCR Clean-Up Kit (Macherey-Nagel GmbH & Co. AG, Düren, Germany) according to the manufacturer’s instructions for Sanger sequencing (Eurofins Genomics Germany GmbH, Ebersberg, Germany).

### 50K SNP chip genotyping

50K SNP genotyping, including DNA extraction from freeze-dried leaf material, was conducted at TraitGenetics (SGS - TraitGenetics GmbH, Germany). Barley FIND-IT variants used for method validation, as well as their corresponding wild types, were genotyped for variant background confirmation. The genetic distance between individuals was calculated as the average of their per-locus distances using R package stringdist (68). Principal coordinate analysis (PCoA) was done with R base function cmdscale based on this genetic distance matrix. The first two PCs were illustrated by ggplot2.

To further determine the efficiency of off-target elimination in barley FIND-IT variants, original FIND-IT variants, crossing parents, and progeny plants (69), derived from two rounds of crosses with RGT Planet or CB Casper, were also 50K SNP-genotyped (table S11). Introgression intervals in progeny plants were defined based on the number of retained variant-specific SNPs.

### Plant material and growth conditions

Barley cultivars and isolated FIND-IT variants were grown in a greenhouse at 18°C under 16/8-hour light/dark cycles. At maturity, grains were harvested and further propagated, either in dozens or in hundreds in the greenhouse or in field row plots, respectively, followed by field trial plots for agronomic evaluation (7.5-m<sup>2</sup> plots). Grain from field plots was harvested and threshed using a Wintersteiger Elite plot combiner (Wintersteiger AG, Germany), and grains were

sorted by size (threshold, 2.5 mm) using a Pfeuffer SLN3 sample cleaner (Pfeuffer GmbH, Germany).

### Yield performance evaluation

TGW, grain width, and grain length of mature dry grains (sample size, approximately 500 grains per measurement) from field-grown plants were determined using the digital seed analyzer MARVIN (GTA Sensorik GmbH, Neubrandenburg, Germany).

### Starch amylose Lugol staining

Grains were embedded in a clay block and grinded with sandpaper. The remaining half-grains were dipped in diluted Lugol solution (0.1% I<sub>2</sub> and 0.2% KI) for 20 min, tapped dry with paper, and exposed to air for approximately 10 min before photographing.

### Starch amylose assay

Three replicates each of 50 grains of the variant *gbss1*<sup>W126\*</sup> and of RGT Planet control were ground to flour in IKA Tube Mill 100 (IKA-Werke GmbH & Co. KG, Germany) at 20,000 rpm, and the amylose:starch ratio was determined using a miniaturized version of the K-AMYL Kit (Megazyme, Wicklow, Ireland).

### Green revolution varieties genotyping—*Sdw1* sequencing

Barley cvs. Paustian and RGT Planet were grown in a greenhouse, and gDNA was extracted from green leaf material using the NucleoSpin 96 Plant II Kit (Macherey-Nagel GmbH & Co. AG, Düren, Germany) according to the manufacturer’s instructions. A fragment of exon 1, harboring the 7-bp deletion of the *sdw1.d* variant allele, was PCR-amplified using the TaKaRa LA Taq polymerase (0.25 µl) with 2× GC buffer I (12.5 µl) (TaKaRa Bio Europe), mixed with Milli-Q water (4.25 µl), 2.5 mM deoxynucleotide triphosphate (4 µl), 10 mM forward primer (1 µl), reverse primer (1 µl), and gDNA (10 ng/µl, 2 µl). PCR with primers CD348\_sdw1d\_fw (5'-GGTGCTCCAGACCGCTCAGC-3') and CD347\_sdw1d\_rc (5'-CCTCCGGAGGTCGTACACC-3') was performed for 35 cycles (initial denaturation at 94°C/3 min followed by 35 cycles of 94°C/30 s, 60°C/30 s, and 72°C/45 s for extension, with a final extension step of 72°C/10 min). PCR products were purified using the NucleoSpin Gel and PCR Clean-Up Kit (Macherey-Nagel GmbH & Co. AG, Düren, Germany) according to the manufacturer’s instructions. Purified PCR products were sequenced by Eurofins (Eurofins Genomics Germany GmbH, Ebersberg, Germany).

### Broken kernel test

Field-propagated mature grains of *CslF6* variant plants *cslF6*<sup>G748D</sup>, *cslF6*<sup>G847E</sup>, and *cslF6*<sup>T709I</sup>, as well as respective reference lines cvs. Paustian and Quench, were harvested and threshed using a trial Wintersteiger Classic (Wintersteiger AG, Austria). Further cleaning of grains was performed on a Pfeuffer 20 sample cleaner, model SLN4 (Pfeuffer GmbH, Germany) using a 2.5-mm screen. Broken grains were counted in four randomly selected 50-g samples of each line and respective wild type, and the number of broken grains for each line was calculated on weight basis.

### α-Amylase gene expression analysis

Triplicates of 100 grains of *Amyl1\_1-var* and wild-type RGT Planet were germinated on filter paper in a petri dish (ø 85 mm, Frisnette, Knebel, Denmark) in 4 ml of water at 16°C. At 48 or 72 hours after the initiation of germination, the grains were snap-frozen in liquid nitrogen and freeze-dried for 48 hours before being finely ground

(MM300, Retsch, Haan, Germany). Primers and probes for *Amy1\_1*, *Amy1\_2*, and *Amy2* were designed based on Morex\_V2 (70). Primers and fluorescent probes for *Amy1\_1* and *Amy2* were designed to specifically amplify and bind to the *Amy1\_1* (*a,b,c,d*) and *Amy2* (*\_1,\_2,\_3*) gene clusters, respectively. *Amy1\_2* is present as a single gene. The amplicon location was specifically targeted to the 3' untranslated region of all targets, and the probe was labeled with FAM for fluorescence detection of transcript. *Amy1\_1*: Fwd\_GACTGG-GGCCTGAAG, rev\_GTGCCGGTCTCCTGAC, FAM\_AGATC-GATCGCCTGGTGTG; *Amy1\_2*: Fwd\_AGATCGATCGTCTGGTG, rev\_TCCATGATCTGCAGCTTG, FAM\_TCAATCAGGACCCG-ACAGG; *Amy2*: Fwd\_CGAGCTCAAGGAGTGG, rev\_CGTC-GATGTACACCTTG, FAM\_AAGAGCGACCTCGGCTTC.

#### RNA extraction and cDNA synthesis

Total RNA was extracted according to Betts *et al.* (71) from the equivalent of two grains (~100 mg of freeze-dried and milled grain material) by a modified Spectrum Plant Total RNA Kit (Sigma-Aldrich, St. Louis, MO, USA) protocol. DNA was removed using deoxyribonuclease I (Sigma-Aldrich, St. Louis, MO, USA), and ZYMO Clean & Concentrator Kit (ZymoResearch, Irvine, CA, USA) was used to clean and concentrate the RNA sample for further downstream applications. RNA quality was validated by gel electrophoresis (1% agarose) and quantified by NanoDrop (Thermo Fisher Scientific, Waltham, MA, USA). cDNA was synthesized from 200 ng of total RNA using a Bio-Rad iScript Select (oligo-dT) kit (Bio-Rad, Hercules, CA, USA) according to the manufacturer's protocol and diluted 10× with Milli-Q water before ddPCR analysis.

#### ddPCR analysis

*Amy1\_1*, *Amy1\_2*, and *Amy2* transcripts were quantified using the Bio-Rad ddPCR Supermix for probes (no dUTP) system (Bio-Rad, Hercules, CA, USA). A master mix of 2× QX200TM ddPCR Supermix for probes (no dUTP) (Bio-Rad, Hercules, CA, USA) and 900/250 nM of 20× target primers and probe, 5 µl of diluted cDNA, and molecular-grade water was made up to a total reaction volume of 20 µl and placed into the QX200 Droplet Generator (Bio-Rad, Hercules, CA, USA). Following droplet generation, the plate was heat-sealed at 180°C for 6 s with pierceable foil using a PX1 PCR plate sealer (Bio-Rad, Hercules, CA, USA). Subsequent PCR amplification (Uno96, VWR, Radnor, PA, USA) was as follows: enzyme activation at 95°C/10 min followed by 40 cycles of denaturation at 94°C/30 s followed by annealing/extension at 55°C/1 min ending with enzyme deactivation at 98°C/10 min, all steps with a ramp of 2°C/s. The PCR plate was then placed in the QX200 Droplet Reader (Bio-Rad, Hercules, CA, USA), and positive and negative droplets were recorded and subsequently analyzed using the QuantaSoft software (Bio-Rad, Hercules, CA, USA) to evaluate amplitude and concentration (copies per microliter of cDNA) of the different transcripts.

#### Analysis of grain (1,3;1,4)-β-glucan content

Barley grains were milled on a Retch cyclone mill (Retsch, Haan, Germany), and replicates of 20 mg of flour were heated for 2 hours at 100°C. After cooling to room temperature, 500 µl of 50% (v/v) aqueous methanol was added, and the sample was shaken gently for 1 hour. Following centrifugation at 16,000g for 10 min, the supernatant was discarded, and the flour was dried overnight. Sodium phosphate buffer [400 µl and 20 mM (pH 6.5)] was added, together with lichenase (1 U/ml) (Megazyme, Wicklow, Ireland) per 10 mg of flour. After incubation at 50°C for 2.5 hours, the sample was centrifuged at 16,000g for 10 min and the supernatant was filtered through 0.45-µm filters.

Released oligosaccharides were quantified by high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) using a Dionex ICS-5000+ DC system (Thermo Fisher Scientific, Waltham MA USA) equipped with a 4-µm SA-10 column at 0.4 ml/min and a column temperature 40°C, with an isocratic 100 mM NaOH eluent for 15 min. Standards were produced by lichenase (1 U/ml) (Megazyme, Wicklow, Ireland) digestion of known quantities of medium viscosity (1,3;1,4)-β-glucan (Megazyme, Wicklow, Ireland) in 50 mM MES buffer, pH 6.5. Visualization of (1,3;1,4)-β-glucan in longitudinal sections of *CslF6* knockout variants and wild-type grains was carried out according to Aastrup *et al.* (72) using 0.05% Calcofluor for staining and 0.05% Fast Green for counterstaining.

#### Homology modeling of protein structures

Homology models were built using the automated SWISS-MODEL homology modeling pipeline (73). The homology model of the barley cellulose synthase-like *CslF6* enzyme (encoded by HORVU.MOREX.r2.7HG0579230) was based on the poplar cellulose synthase isoform 8 (Protein Data Bank ID: 6wlb; 43.18% identity). All models were visualized using VMD (74), and final figures were arranged using Adobe Illustrator and CorelDRAW.

#### Yeast library generation

A haploid yeast strain was treated with ethyl methanesulfonate (EMS) using a standard protocol (75). Briefly, the strain was first grown in standard yeast extract/peptone/dextrose (YPD) medium to saturation. The cells were spun down and washed once with sterile water and one more time with 0.1 M sodium phosphate buffer (pH 7). Last, a cell suspension was generated in the same buffer with a cell titer of approximately  $2 \times 10^8$  cells/ml. For each mutagenesis reaction, 1 ml of this cell suspension was pipetted into a 2-ml safe-lock reaction tube and 30 µl of EMS was added to the cells. The reactions were placed on the Eppendorf Thermomixer (1.5 ml) and incubated at 1000 rpm and 30°C for 75 min. These conditions resulted in a survival rate of approximately 30%. The mutagenesis was stopped by washing the cells three times with 1 ml of freshly prepared sterile sodium thiosulfate solution (5%) and one time with sterile water. Cells were finally revived in 1 ml of YPD for 1 hour at beforementioned conditions. Cells were either processed right away or stored at 4°C until use.

It has been suggested that growing yeast cells under selective pressure immediately after mutagenesis would select for cells with mutator-like phenotype accumulating significantly more mutations per single cell than under nonselective conditions (76, 77). Hence, one could in theory reduce the initial size of our variant library significantly due to the cells' increased mutation density. We therefore plated the mutagenized yeast cells on solid synthetic dextrose medium containing the herbicide metsulfuron methyl (2 µg/ml) and incubated the plates at 30°C until resistant colonies appeared. The cells of approximately 50,000 colonies were collected from the plates, washed with sterile phosphate buffer (see above), and used for another round of EMS mutagenesis and selection as described. The described mutagenesis/selection cycle was done four times overall to increase the mutation density in the yeast population even further.

After the final cycle, the cell titer of the yeast suspension was determined, and cells were spread onto solid YPD medium to create plates with approximately 1200 to 1500 colonies per plate. After 3 days at 30°C, cells from 96 individual plates were washed off with phosphate buffer and collected separately to form 96 library pools.

These cells were subsequently used for DNA isolation using the PureLink Pro 96 gDNA Kit (Thermo Fisher Scientific, Waltham, MA, USA) with EveryPrep Universal Vacuum Manifold (Thermo Fisher Scientific, Waltham, MA, USA) and for generating glycerol stocks for long-term storage. The final “total random library” created with the described procedure should contain approximately 120,000 to 150,000 variant clones overall.

### Yeast library screening

We used the FIND-IT approach to screen for mutations in the *Ferulic Acid Decarboxylase* (*FDC1*) gene. This enzyme is responsible for converting aromatic carboxylic acids such as ferulic acid derived from malt to their corresponding vinyl derivatives, which often impart undesirable clove-like off-flavors in beer (78). Two Trp residues in the enzyme were targeted for the formation of premature stop codons (W159\* and W171\*) using the S288C *ScFDC1* sequence as a reference (National Center for Biotechnology Information ID: NM\_001180847). The mutations were confirmed by sequence analysis. The assay of variant lines was based on the inactivation of the *FDC1* gene, which resulted in increased sensitivity of the yeast to growth inhibition by cinnamic acid (78); this effect was clearly visible in the variant lines (note S6.3).

### Bacterial library generation and screening

*A. L. pasteurii* (DSM 23907) library of about 960,000 genetic variants was generated by treating cells with 1.4% EMS for 30 min. After mutagenesis and recovery, cells were counted using a NovoCyte flow cytometer (Acea Biosciences) and about 1000 mutagenized bacterial cells were dispensed into each well of ten 96-well plates. The library was screened for variants of the *Galactose-1-Phosphate Uridyltransferase* (*galT*) gene, which mediates the second step of the Leloir pathway in which D-galactose is metabolized (79, 80). Mutations in this gene should show no phenotype if the bacteria are grown on glucose. After screening about 10% of the library, two variants with stop codons in this gene were identified. The variants were *Lac-GalMut\_W89\** and *Lac-GalMut\_Q148\** (note S7.2).

### Nucleotide and amino acid numbering

Nucleotide and amino acid numbering are based on publicly available gene and protein sequences. The exact amino acid positions described are dependent on the template haplotype used for generating the probe. Where possible, we have aligned the numbering in barley sequences with the Morex\_V2 genome sequence (9), but we also provide full-length Morex\_V2 probe sequences that will guide readers to the correct nucleotide positions (table S9 and note S7.1).

### Statistical analysis

All field-grown materials used for agronomical, enzymatic, or gene expression characterization were grown in trials in either New Zealand (September to February) or Denmark (April to August) and in distinct replicates of  $n = 2$  to  $n = 5$ . Microsoft Excel data analysis package was used for comparison of two groups by paired Student's two-tailed *t* test to test whether the hypothesized mean difference is zero. Values of *P* were considered significant different if  $***P \leq 0.001$ ,  $**P \leq 0.01$ , or  $*P \leq 0.05$ . All statistical parameters and outputs are presented in table S12.

### SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <https://science.org/doi/10.1126/sciadv.abq2266>

### REFERENCES AND NOTES

1. D. Pimentel, Green revolution agriculture and chemical hazards. *Sci. Total Environ.* **188**, 86–98 (1996).
2. R. E. Evenson, D. Gollin, Assessing the impact of the green revolution, 1960 to 2000. *Science* **300**, 758–762 (2003).
3. ISAAA. Global status of commercialized biotech/GM crops in 2019: biotech crops drive socio-economic development and sustainable environment in the new frontier. *ISAAA Brief No. 55* (Ithaca, ISAAA, 2019).
4. Wikipedia, Arable land (2022); [https://en.wikipedia.org/wiki/Arable\\_land](https://en.wikipedia.org/wiki/Arable_land).
5. M. Eisenstein, Base edit your way to better crops. *Nature* **604**, 790–792 (2022).
6. C. Jung, B. Till, Mutagenesis and genome editing in crop improvement: Perspectives for the global regulatory landscape. *Trends Plant Sci.* **26**, 1258–1269 (2021).
7. D. K. Ray, N. Ramankutty, N. D. Mueller, P. C. West, J. A. Foley, Recent patterns of crop yield growth and stagnation. *Nat. Commun.* **3**, 1293 (2012).
8. K. P. Voss-Fels, A. Stahl, L. T. Hickey, Q&A: Modern crop breeding for future food security. *BMC Biol.* **17**, 18 (2019).
9. M. Jayakodi, S. Padmarasu, G. Haberer, V. S. Bonthala, H. Gundlach, C. Monat, T. Lux, N. Kamal, D. Lang, A. Himmelbach, J. Ens, X. Q. Zhang, T. T. Angessa, G. Zhou, C. Tan, C. Hill, P. Wang, M. Schreiber, L. B. Boston, C. Plott, J. Jenkins, Y. Guo, A. Fiebig, H. Budak, D. Xu, J. Zhang, C. Wang, J. Grimwood, J. Schmutz, G. Guo, G. Zhang, K. Mochida, T. Hirayama, K. Sato, K. J. Chalmers, P. Langridge, R. Waugh, C. J. Pozniak, U. Scholz, K. F. X. Mayer, M. Spannagl, C. Li, M. Mascher, N. Stein, The barley pan-genome reveals the hidden legacy of mutation breeding. *Nature* **588**, 284–289 (2020).
10. J. E. Bradshaw, *Plant Mutation Breeding and Biotechnology*, Q. Y. Shu, B. P. Forster, H. Nakagawa, Eds. (CABI Publishing, 2012).
11. J. Peng, D. E. Richards, N. M. Hartley, G. P. Murphy, K. M. Devos, J. E. Flintham, J. Beales, L. J. Fish, A. J. Worland, F. Pelica, D. Sudhakar, P. Christou, J. W. Snape, M. D. Gale, N. P. Harberd, ‘Green revolution’ genes encode mutant gibberellin response modulators. *Nature* **400**, 256–261 (1999).
12. I. M. Henry, U. Nagalakshmi, M. C. Lieberman, K. J. Ngo, K. V. Krasileva, H. Vasquez-Gross, A. Akhunova, E. Akhunov, J. Dubcovsky, T. H. Tai, Efficient genome-wide detection and cataloging of EMS-induced mutations using exome capture and next-generation sequencing. *Plant Cell* **26**, 1382–1397 (2014).
13. W. Yang, H. Feng, X. Zhang, J. Zhang, J. H. Doonan, W. D. Batchelor, L. Xiong, J. Yan, Crop phenomics and high-throughput phenotyping: Past decades, current challenges, and future perspectives. *Mol. Plant* **13**, 187–214 (2020).
14. C. M. McCallum, L. Comai, E. A. Greene, S. Henikoff, Targeted screening for induced mutations. *Nat. Biotechnol.* **18**, 455–457 (2000).
15. L. Comai, K. Young, B. J. Till, S. H. Reynolds, E. A. Greene, C. A. Codomo, L. C. Enns, J. E. Johnson, C. Burtner, A. R. Odden, S. Henikoff, Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. *Plant J.* **37**, 778–786 (2004).
16. K. V. Krasileva, H. A. Vasquez-Gross, T. Howell, P. Bailey, F. Paraiso, L. Clissold, J. Simmonds, R. H. Ramirez-Gonzalez, X. Wang, P. Borrill, C. Fosker, S. Ayling, A. L. Phillips, C. Uauy, J. Dubcovsky, Uncovering hidden variation in polyploid wheat. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E913–E921 (2017).
17. G. Li, R. Jain, M. Chern, N. T. Pham, J. A. Martin, T. Wei, W. S. Schackwitz, A. M. Lipzen, P. Q. Duong, K. C. Jones, L. Jiang, D. Ruan, D. Bauer, Y. Peng, K. W. Barry, J. Schmutz, P. C. Ronald, The sequences of 1504 mutants in the model rice variety kitaake facilitate rapid functional genomic studies. *Plant Cell* **29**, 1218–1231 (2017).
18. J. Leita, Q. Shu, B. Forster, H. Nakagawa, Chemical mutagenesis, in *Plant Mutation Breeding and Biotechnology* (CABI Publishing, 2012), pp. 135–158.
19. B. J. Hindson, K. D. Ness, D. A. Masquelier, P. Belgrader, N. J. Heredia, A. J. Makarewicz, I. J. Bright, M. Y. Lucero, A. L. Hiddessen, T. C. Legler, T. K. Kitano, M. R. Hodel, J. F. Petersen, P. W. Wyatt, E. R. Steenblock, P. H. Shah, L. J. Bousse, C. B. Troup, J. C. Mellen, D. K. Wittmann, N. G. Erndt, T. H. Cauley, R. T. Koehler, A. P. So, S. Dube, K. A. Rose, L. Montesclaros, S. Wang, D. P. Stumbo, S. P. Hodges, S. Romine, F. P. Milanovich, H. E. White, J. F. Regan, G. A. Karlin-Neumann, C. M. Hindson, S. Saxonov, B. W. Colston, High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal. Chem.* **83**, 8604–8610 (2011).
20. Bio-Rad, qPCR assay design and optimization (2022); [www.bio-rad.com/en-dk/applications-technologies/qpcr-assay-design-optimization?ID=LUSO7R1VK](http://www.bio-rad.com/en-dk/applications-technologies/qpcr-assay-design-optimization?ID=LUSO7R1VK).
21. V. Talamè, R. Bovina, M. C. Sanguineti, R. Tuberosa, U. Lundqvist, S. Salvi, TILLMore, a resource for the discovery of chemically induced mutants in barley. *Plant Biotechnol. J.* **6**, 477–485 (2008).
22. C. Jiang, M. Lei, Y. Guo, G. Gao, L. Shi, Y. Jin, Y. Cai, A. Himmelbach, S. Zhou, Q. He, A reference-guided TILLING by amplicon-seq platform supports forward and reverse genetics in barley. *Plant Commun.* 100317 (2022).
23. M. Schreiber, A. Barakate, N. Uzrek, M. Macaulay, A. Sourdille, J. Morris, P. E. Hedley, L. Ramsay, R. Waugh, A highly mutagenised barley (cv. Golden Promise) TILLING population coupled with strategies for screening-by-sequencing. *Plant Methods* **15**, 99 (2019).

24. A. Sevanthi, P. Kandwal, P. B. Kale, C. Prakash, M. Ramkumar, N. Yadav, A. K. Mahato, V. Sureshkumar, M. Behera, R. K. Deshmukh, Whole genome characterization of a few EMS-induced mutants of upland rice variety Nagina 22 reveals a staggeringly high frequency of SNPs which show high phenotypic plasticity towards the wild-type. *Front. Plant Sci.* **9**, 1179 (2018).
25. M. Mascher, T. Wicker, J. Jenkins, C. Plott, T. Lux, C. S. Koh, J. Ens, H. Gundlach, L. B. Boston, Z. Tulpová, Long-read sequence assembly: A technical evaluation in barley. *Plant Cell* **33**, 1888–1906 (2021).
26. R. Von Post, L. Von Post, C. Dayteg, M. Nilsson, B. P. Forster, S. Tuveesson, A high-throughput DNA extraction method for barley seed. *Euphytica* **130**, 255–260 (2003).
27. S. Taketa, S. Amano, Y. Tsujino, T. Sato, D. Saisho, K. Kakeda, M. Nomura, T. Suzuki, T. Matsumoto, K. Sato, Barley grain with adhering hulls is controlled by an ERF family transcription factor gene regulating a lipid biosynthesis pathway. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 4062–4067 (2008).
28. H. Bull, M. C. Casao, M. Zwirek, A. J. Flavell, W. T. B. Thomas, W. Guo, R. Zhang, P. Papazote-Flores, S. Kyriakidis, J. Russell, A. Druka, S. M. McKim, R. Waugh, Barley SIX-ROWED SPIKE3 encodes a putative Jumonji C-type H3K9me2/me3 demethylase that represses lateral spikelet fertility. *Nat. Commun.* **8**, 936 (2017).
29. G. Hu, C. Burton, Z. Hong, E. Jackson, A mutation of the cellulose-synthase-like (CslF6) gene in barley (*Hordeum vulgare* L.) partially affects the  $\beta$ -glucan content in grains. *J. Cereal Sci.* **59**, 189–195 (2014).
30. M. Gupta, N. Abu-Ghannam, E. Gallagher, Barley for brewing: Characteristic changes during malting, brewing and applications of its by-products. *Compr. Rev. Food Sci.* **9**, 318–328 (2010).
31. G. Garcia-Gimenez, A. Barakate, P. Smith, J. Stephens, S. F. Khor, M. S. Doblin, P. Hao, A. Bacic, G. B. Fincher, R. A. Burton, R. Waugh, M. R. Tucker, K. Houston, Targeted mutation of barley (1,3;1,4)- $\beta$ -glucan synthases reveals complex relationships between the storage and cell wall polysaccharide content. *Plant J.* **104**, 1009–1022 (2020).
32. N. J. Patron, A. M. Smith, B. F. Fahy, C. M. Hylton, M. J. Naldrett, B. G. Rossnagel, K. Denyer, The altered pattern of amylose accumulation in the endosperm of low-amylose barley cultivars is attributable to a single mutant allele of granule-bound starch synthase I with a deletion in the 5'-non-coding region. *Plant Physiol.* **130**, 190–198 (2002).
33. K. a. y. Denyer, P. Johnson, S. Zeeman, A. M. Smith, The control of amylose synthesis. *J. Plant Physiol.* **158**, 479–487 (2001).
34. X.-J. Song, W. Huang, M. Shi, M.-Z. Zhu, H.-X. Lin, A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat. Genet.* **39**, 623–630 (2007).
35. X. Xu, R. Sharma, A. Tondelli, J. Russell, J. Comadran, F. Schnaithmann, K. Pillen, B. Kilian, L. Cattivelli, W. T. B. Thomas, A. J. Flavell, Genome-wide association analysis of grain yield-associated traits in a Pan-European Barley cultivar collection. *Plant Genome* **11**, 170073 (2018).
36. S. Li, Y. Tian, K. Wu, Y. Ye, J. Yu, J. Zhang, Q. Liu, M. Hu, H. Li, Y. Tong, N. P. Harber, X. Fu, Modulating plant growth-metabolism coordination for sustainable agriculture. *Nature* **560**, 595–600 (2018).
37. L. E. Dixon, M. Pasquariello, R. Badgami, K. A. Levin, G. Poschet, P. Q. Ng, S. Orford, N. Chayut, N. M. Adamski, J. Brinton, J. Simmonds, B. Steuernagel, I. R. Searle, C. Uauy, S. A. Boden, MicroRNA-resistant alleles of *HOMEODOMAIN-2* modify inflorescence branching and increase grain protein content of wheat. *Sci. Adv.* **8**, eabn5907 (2022).
38. X. Zou, D. Neuman, Q. J. Shen, Interactions of two transcriptional repressors and two transcriptional activators in modulating gibberellin signaling in aleurone cells. *Plant Physiol.* **148**, 176–186 (2008).
39. Y. Zhang, M. Pribil, M. Palmgren, C. Gao, A CRISPR way for accelerating improvement of food crops. *Nat. Food* **1**, 200–205 (2020).
40. D. Zhang, Z. Wang, N. Wang, Y. Gao, Y. Liu, Y. Wu, Y. Bai, Z. Zhang, X. Lin, Y. Dong, X. Ou, C. Xu, B. Liu, Tissue culture-induced heritable genomic variation in rice, and their phenotypic implications. *PLOS ONE* **9**, e96879 (2014).
41. M. F. Maher, R. A. Nasti, M. Vollbrecht, C. G. Starker, M. D. Clark, D. F. Voytas, Plant gene editing through de novo induction of germlines. *Nat. Biotechnol.* **38**, 84–89 (2020).
42. T. Blondal, C. Gamba, L. Møller Jagd, L. S. Demirov, S. Guo, C. M. Johnston, E. M. Riising, X. Wu, M. J. Mikkelsen, L. Szabova, P. Mouritzen, Verification of CRISPR editing and finding transgenic inserts by Xdrop indirect sequence capture followed by short- and long-read sequencing. *Methods* **191**, 68–77 (2021).
43. E. Callaway, CRISPR plants now subject to tough GM laws in European Union. *Nature* **560**, 16 (2018).
44. B. J. Till, T. Colbert, R. Tompa, L. C. Enns, C. A. Codomo, J. E. Johnson, S. H. Reynolds, J. G. Henikoff, E. A. Greene, M. N. Steine, High-throughput TILLING for functional genomics, in *Plant Functional Genomics* (Springer, 2003), pp. 205–220.
45. A. J. Slade, S. I. Fuerstenberg, D. Loeffler, M. N. Steine, D. Facciotti, A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nat. Biotechnol.* **23**, 75–81 (2005).
46. M. Fruzangohar, E. Kalashyan, P. Kalambettu, J. Ens, K. Wiebe, C. J. Pozniak, P. J. Tricker, U. Baumann, Novel informatic tools to support functional annotation of the durum wheat genome. *Front. Plant Sci.* **10**, 1244 (2019).
47. P. Borrill, S. A. Harrington, C. Uauy, Applying the latest advances in genomics and phenomics for trait discovery in polyploid wheat. *Plant J.* **97**, 56–72 (2019).
48. G. Luo, J. Najafi, P. M. P. Correia, M. D. L. Trinh, E. A. Chapman, J. T. Østerberg, H. C. Thomsen, P. R. Pedas, S. Larson, C. Gao, J. Poland, S. Knudsen, L. DeHaan, M. Palmgren, Accelerated domestication of new crops: Yield is key. *Plant Cell Physiol.* pcc065 (2022).
49. R. L. López-Marqués, A. F. Nørrevang, P. Ache, M. Moog, D. Visintainer, T. Wendt, J. T. Østerberg, C. Dockter, M. E. Jørgensen, A. T. Salvador, R. Hedrich, C. Gao, S.-E. Jacobsen, S. Shabala, M. Palmgren, Prospects for the accelerated improvement of the resilient crop quinoa. *J. Exp. Bot.* **71**, 5333–5347 (2020).
50. P. Langridge, R. Waugh, Harnessing the potential of germplasm collections. *Nat. Genet.* **51**, 200–201 (2019).
51. D. Bustos-Korts, I. K. Dawson, J. Russell, A. Tondelli, D. Guerra, C. Ferrandi, F. Strozzi, E. L. Nicolazzi, M. Molnar-Lang, H. Ozkan, M. Megyeri, P. Miko, E. Cakir, E. Yakisir, N. Trabanco, S. Delbono, S. Kyriakidis, A. Booth, D. Cammarano, M. Mascher, P. Werner, L. Cattivelli, L. Rossini, N. Stein, B. Kilian, R. Waugh, F. A. van Eeuwijk, Exome sequences and multi-environment field trials elucidate the genetic basis of adaptation in barley. *Plant J.* **99**, 1172–1191 (2019).
52. A. Kleinhofs, R. Warner, F. Murhlabauer, R. Nilan, Induction and selection of specific gene mutations in *Hordeum* and *Pisum*. *Mutat. Res. Fundam. Mol. Mech. Mutagen.* **51**, 29–35 (1978).
53. D. R. Page, U. Grossniklaus, The art and design of genetic screens: *Arabidopsis thaliana*. *Nat. Rev. Genet.* **3**, 124–136 (2002).
54. M. R. Green, J. Sambrook, Isolation of high-molecular-weight DNA using organic solvents. *Cold Spring Harb. Protoc.* **2017**, pdb.prot093450 (2017).
55. A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernysky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
56. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, H. Li, Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
57. J. T. Robinson, H. Thorvaldsdóttir, D. Turner, J. P. Mesirov, igv.js: An embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV). bioRxiv 2020.05.03.075499 [Preprint]. 5 May 2020. <https://doi.org/10.1101/2020.05.03.075499>.
58. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997 [q-bio.GN] (16 March 2013).
59. E. Garrison, G. Marth, Haplotype-based variant detection from short-read sequencing. arXiv:1207.3907 [q-bio.GN] (17 July 2012).
60. E. Kosman, K. J. Leonard, Similarity coefficients for molecular markers in studies of genetic relationships between individuals for haploid, diploid, and polyploid species. *Mol. Ecol.* **14**, 415–424 (2005).
61. J. T. Hill, B. L. Demarest, B. W. Bigrove, Y.-C. Su, M. Smith, H. J. Yost, Poly peak parser: Method and software for identification of unknown indels using sanger sequencing of polymerase chain reaction products. *Dev. Dyn.* **243**, 1632–1636 (2014).
62. R. C. Gentleman, V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A. J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J. Y. H. Yang, J. Zhang, Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biol.* **5**, 1–16 (2004).
63. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
64. D. Charif, J. R. Lobry, SeqinR 1.0-2: A contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis, in *Structural Approaches to Sequence Evolution* (Springer, 2007), pp. 207–232.
65. M. Krzywinski, J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones, M. A. Marra, Circo: An information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009).
66. M. Goel, H. Sun, W.-B. Jiao, K. Schneeberger, SyRI: Finding genomic rearrangements and local sequence differences from whole-genome assemblies. *Genome Biol.* **20**, 1–13 (2019).
67. S. M. Kale, A. W. Schulthess, S. Padmarasu, P. Boeven, J. Schacht, A. Himmelbach, B. Steuernagel, B. Wulff, J. Reif, N. Stein, A catalogue of resistance gene homologs and a chromosome-scale reference sequence support resistance gene mapping in winter wheat. *Plant Biotechnol. J.* pbi13843 (2022).
68. M. P. Van der Loo, The stringdist package for approximate string matching. *R J.* **6**, 111 (2014).
69. S. Knudsen, I. Due, S. Andersen, Components of response in barley anther culture. *Plant Breed.* **103**, 241–246 (1989).
70. M. Mascher, H. Gundlach, A. Himmelbach, S. Beier, S. O. Twardziok, T. Wicker, V. Radchuk, C. Dockter, P. E. Hedley, J. Russell, A chromosome conformation capture ordered sequence of the barley genome. *Nature* **544**, 427–433 (2017).
71. N. S. Betts, O. Berkowitz, R. Liu, H. M. Collins, B. Skadhauge, C. Dockter, R. A. Burton, J. Whelan, G. B. Fincher, Isolation of tissues and preservation of RNA from intact, germinated barley grain. *Plant J.* **91**, 754–765 (2017).

72. S. Astrup, G. C. Gibbons, L. Munck, A rapid method for estimating the degree of modification in barley malt by measurement of cell wall breakdown. *Carlsberg Res. Commun.* **46**, 77–86 (1981).
73. A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede, SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).
74. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
75. D. Burke, D. Dawson, T. Stearns, in *Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual* (Cold Spring Harbor Laboratory Press, 2000).
76. A. G. Lada, E. I. Stepchenkova, I. S. R. Waisretriger, V. N. Noskov, A. Dhar, J. D. Eudy, R. J. Boissy, M. Hirano, I. B. Rogozin, Y. I. Pavlov, Genome-wide mutation avalanches induced in diploid yeast cells by a base analog or an APOBEC deaminase. *PLOS Genet.* **9**, e1003736 (2013).
77. T. Den Abt, B. Souffriau, M. R. Foulquié-Moreno, J. Duitama, J. M. Thevelein, Genomic saturation mutagenesis and polygenic analysis identify novel yeast genes affecting ethyl acetate production, a non-selectable polygenic trait. *Microbial Cell* **3**, 159–175 (2016).
78. M. Lentz, The impact of simple phenolic compounds on beer aroma and flavor. *Fermentation* **4**, 20 (2018).
79. B. P. Grossird, E. J. Luesink, E. E. Vaughan, A. Arnaud, W. M. de Vos, Characterization, expression, and mutation of the *Lactococcus lactis* galPMKTE genes, involved in galactose utilization via the Leloir pathway. *J. Bacteriol. Res.* **185**, 870–878 (2003).
80. C. F. Iskandar, C. Cailliez-Grimal, F. Borges, A.-M. Revol-Junelles, Review of lactose and galactose metabolism in lactic acid bacteria dedicated to expert genomic annotation. *Trends Food Sci. Technol.* **88**, 121–132 (2019).
81. B. Keller, S. G. Krattinger, Plant science: Genomic compartments in barley. *Nature* **544**, 424–425 (2017).
82. J. Hernandez, B. Meints, P. Hayes, Introgression breeding in barley: Perspectives and case studies. *Front. Plant Sci.* **11**, 761 (2020).
83. A. Druka, J. Franckowiak, U. Lundqvist, N. Bonar, J. Alexander, K. Houston, S. Radovic, F. Shahinnia, V. Vendramin, M. Morgante, Genetic dissection of barley morphology and development. *Plant Physiol.* **155**, 617–627 (2011).
84. E. Tavakol, R. Okagaki, G. Verderio, V. S. J. A. Hussien, H. Bilgic, M. J. Scanlon, N. R. Todt, T. J. Close, A. Druka, R. Waugh, B. Steuernagel, R. Ariyadasa, A. Himmelbach, N. Stein, G. J. Muehlbauer, L. Rossini, The Barley *Uniculme4* gene encodes a BLADE-ON-PETIOLE-like protein that controls tillering and leaf patterning. *Plant Physiol.* **168**, 164–174 (2015).
85. C. Dockter, D. Gruszka, I. Braumann, A. Druka, J. Franckowiak, S. P. Gough, A. Janeczko, M. Kurowska, J. Lundqvist, U. Lundqvist, M. Marzec, I. Matyszczak, A. H. Müller, J. Oklestkova, B. Schulz, S. Zakhrebekova, M. Hansson, Induced variations in brassinosteroid genes define barley height and sturdiness, and expand the green revolution genetic toolkit. *Plant Physiol.* **166**, 1912–1927 (2014).
86. R. A. Burton, G. Ma, U. Baumann, A. J. Harvey, N. J. Shirley, J. Taylor, F. Pettolino, A. Bacic, M. Beatty, C. R. Simmons, K. S. Dhugga, J. A. Rafalski, S. V. Tingey, G. B. Fincher, A customized gene expression microarray reveals that the brittle stem phenotype fs2 of barley is attributable to a retroelement in the HvCesA4 cellulose synthase gene. *Plant Physiol.* **153**, 1716–1728 (2010).
87. M. Jost, S. Taketa, M. Mascher, A. Himmelbach, T. Yuo, F. Shahinnia, T. Rutten, A. Druka, T. Schmutzer, B. Steuernagel, S. Beier, S. Taudien, U. Scholz, M. Morgante, R. Waugh, N. Stein, A homolog of blade-on-petiole 1 and 2 (BOP1/2) controls internode length and homeotic changes of the barley inflorescence. *Plant Physiol.* **171**, 1113–1127 (2016).
88. S. Zakhrebekova, C. Dockter, K. Ahmann, I. Braumann, S. P. Gough, T. Wendt, U. Lundqvist, M. Mascher, N. Stein, M. Hansson, Genetic linkage facilitates cloning of Ert-m regulating plant architecture in barley and identified a strong candidate of Ant1 involved in anthocyanin biosynthesis. *Plant Mol. Biol.* **88**, 609–626 (2015).
89. K. Houston, S. M. McKim, J. Comadran, N. Bonar, I. Druka, N. Uzrek, E. Cirillo, J. Guzy-Wrobelska, N. C. Collins, C. Halpin, M. Hansson, C. Dockter, A. Druka, R. Waugh, Variation in the interaction between alleles of HvAPETALA2 and microRNA172 determines the density of grains on the barley inflorescence. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 16675–16680 (2013).
90. I. Matyszczak, M. Tominska, S. Zakhrebekova, C. Dockter, M. Hansson, Analysis of early-flowering genes at barley chromosome 2H expands the repertoire of mutant alleles at the Mat-c locus. *Plant Cell Rep.* **39**, 47–61 (2020).
91. T. Wendt, I. Holme, C. Dockter, A. Preuß, W. Thomas, A. Druka, R. Waugh, M. Hansson, I. Braumann, HvDep1 is a positive regulator of culm elongation and grain size in barley and impacts yield in an environment-dependent manner. *PLOS ONE* **11**, e0168924 (2016).
92. S. Ghosh, A. Watson, O. E. Gonzalez-Navarro, R. H. Ramirez-Gonzalez, L. Yanes, M. Mendoza-Suárez, J. Simmonds, R. Wells, T. Rayner, P. Green, A. Hafeez, S. Hayta, R. E. Melton, A. Steed, A. Sarkar, J. Carter, L. Perkins, J. Lord, M. Tester, A. Osbourn, M. J. Moscou, P. Nicholson, W. Harwood, C. Martin, C. Domoney, C. Uauy, B. Hazard, B. B. H. Wulff, L. T. Hickey, Speed breeding in growth chambers and glasshouses for crop breeding and model plant research. *Nat. Protoc.* **13**, 2944–2963 (2018).
93. T. Lawrenson, W. A. Harwood, Creating targeted gene knockouts in barley using CRISPR/Cas9. *Methods Mol. Biol.* **1900**, 217–232 (2019).
94. B. Orman-Ligeza, W. Harwood, P. E. Hedley, A. Hinchcliffe, M. Macaulay, C. Uauy, K. Trafford, TRA1: A locus responsible for controlling agrobacterium-mediated transformability in barley. *Front. Plant Sci.* **11**, 355 (2020).
95. X. Tang, S. Sretenovic, Q. Ren, X. Jia, M. Li, T. Fan, D. Yin, S. Xiang, Y. Guo, L. Liu, X. Zheng, Y. Qi, Y. Zhang, Plant prime editors enable precise gene editing in rice cells. *Mol. Plant* **13**, 667–670 (2020).
96. F. Sparla, G. Falini, E. Botticella, C. Pirone, V. Talame, R. Bovina, S. Salvi, R. Tuberosa, F. Sestili, P. Trost, New starch phenotypes produced by TILLING in Barley. *PLOS ONE* **9**, e107779 (2014).
97. Q. Li, Z. Pan, G. Deng, H. Long, Z. Li, X. Deng, J. Liang, Y. Tang, X. Zeng, N. Tashi, M. Yu, Effect of wide variation of the waxy gene on starch properties in hull-less barley from Qinghai-tibet plateau in China. *J. Agric. Food Chem.* **62**, 11369–11385 (2014).
98. S. Yu, H. Long, G. Deng, Z. Pan, J. Liang, X. Zeng, Y. Tang, N. Tashi, M. Yu, A single nucleotide polymorphism of nud converts the caryopsis type of barley (*Hordeum vulgare* L.). *Plant Mol. Biol. Rep.* **34**, 242–248 (2016).
99. J. Russell, M. Mascher, I. K. Dawson, S. Kyriakidis, C. Calixto, F. Freund, M. Bayer, I. Milne, T. Marshall-Griffiths, S. Heinen, A. Hofstad, R. Sharma, A. Himmelbach, M. Knauff, M. van Zonneveld, J. W. S. Brown, K. Schmid, B. Kilian, G. J. Muehlbauer, N. Stein, R. Waugh, Exome sequencing of geographically diverse barley landraces and wild relatives gives insights into environmental adaptation. *Nat. Genet.* **48**, 1024–1030 (2016).
100. W. Wang, R. Mauleon, Z. Hu, D. Chebotarov, S. Tai, Z. Wu, M. Li, T. Zheng, R. R. Fuentes, F. Zhang, L. Mansueti, D. Copetti, M. Sanciangco, K. C. Palis, J. Xu, C. Sun, B. Fu, H. Zhang, Y. Gao, Y. Zhao, F. Shen, X. Cui, H. Yu, Z. Li, M. Chen, J. Detras, Y. Zhou, X. Zhang, Y. Zhao, D. Kudrna, C. Wang, R. Li, B. Jia, J. Lu, X. He, Z. Dong, J. Xu, Y. Li, M. Wang, J. Shi, J. Li, D. Zhang, S. Lee, W. Hu, A. Poliakov, I. Dubchak, V. J. Ulat, F. N. Borja, J. R. Mendoza, J. Ali, Q. Gao, Y. Niu, Z. Yue, M. E. B. Naredo, J. Talag, X. Wang, J. Li, X. Fang, Y. Yin, J. C. Glaszmann, J. Zhang, J. Li, R. S. Hamilton, R. A. Wing, J. Ruan, G. Zhang, C. Wei, N. Alexandrov, K. L. McNally, Z. Li, H. Leung, Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* **557**, 43–49 (2018).
101. J. Simmonds, P. Scott, J. Brinton, T. C. Mestre, M. Bush, A. Del Blanco, J. Dubcovsky, C. Uauy, A splice acceptor site mutation in TaGW2-A1 increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. *Theor. Appl. Genet.* **129**, 1099–1112 (2016).
102. P. Hedden, The genes of the green revolution. *Trends Genet.* **19**, 5–9 (2003).
103. P. M. Chandler, C. A. Harding, 'Overgrowth' mutants in barley and wheat: New alleles and phenotypes of the 'Green Revolution' Della gene. *J. Exp. Bot.* **64**, 1603–1613 (2013).
104. Y. Xu, Q. Jia, G. Zhou, X. Q. Zhang, T. Angessa, S. Broughton, G. Yan, W. Zhang, C. Li, Characterization of the sdw1 semi-dwarf gene in barley. *BMC Plant Biol.* **17**, 11 (2017).
105. P. M. Chandler, C. A. Harding, R. A. Ashton, M. D. Mulcair, N. E. Dixon, L. N. Mander, Characterization of gibberellin receptor mutants of barley (*Hordeum vulgare* L.). *Mol. Plant* **1**, 285–294 (2008).
106. C. Marè, E. Mazzucotelli, C. Crosatti, E. Francia, A. m. Stanca, L. Cattivelli, Hv-WRKY38: A new transcription factor involved in cold- and drought-response in barley. *Plant Mol. Biol.* **55**, 399–416 (2004).
107. N. S. Betts, C. Dockter, O. Berkowitz, H. M. Collins, M. Hooij, Q. Lu, R. A. Burton, V. Bulone, B. Skadhauge, J. Whelan, G. B. Fincher, Transcriptional and biochemical analyses of gibberellin expression and content in germinated barley grain. *J. Exp. Bot.* **71**, 1870–1884 (2020).
108. Y. Wang, X. Cheng, Q. Shan, Y. Zhang, J. Liu, C. Gao, J.-L. Qiu, Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol.* **32**, 947–951 (2014).
109. C. Monat, S. Padmarasu, T. Lux, T. Wicker, H. Gundlach, A. Himmelbach, J. Ens, C. Li, G. J. Muehlbauer, A. H. Schulman, TRITEX: Chromosome-scale sequence assembly of Triticeae genomes with open-source tools. *Genome Biol.* **20**, 1–18 (2019).

#### Acknowledgments

**Funding:** We thank the Carlsberg Foundation for funding support: Carlsberg Foundation grant CF14-0461 (to B.L.M.) and Carlsberg Foundation grant CF15-0236 (to B.S.). **Author contributions:** S.K., T.W., and B.S. conceived the project. S.K., H.C.T., C.V., T.W., M.R., C.D., and P.R.P. were responsible for the mutagenesis and variant propagation. L.T.P., C.V., S.K., H.C.T., T.W., M.C., M.R., A.S., L.A.M., M.K.J., and J.A.C.-S. generated plant libraries. Q.L., C.D., M.E.J., J.T.Ø., K.N., J.O.H., and S.M.K. validated library and variant mutation load. H.C.T., L.T.P., C.V., M.C., M.R., and T.W. performed the ddPCR experiments. C.D., J.A.C.-S., I.B., S.B., and S.K. designed barley variant targets. C.D., S.B., J.A.C.-S., P.R.P., E.M., J.H., H.C.T., M.T.S.N., and S.K. analyzed barley variants. K.L., M.K., R.G.S., and N.S. prepared and screened the yeast library. R.T.F. and T.W. prepared and screened the bacterial library. Z.G., J.F., O.O., and B.L.M. assisted with experimental design, and C.D., M.E.J., H.C.T., M.R., B.L.M., G.B.F., and B.S. were responsible for

the preparation of the manuscript, with assistance from the other authors. **Competing interests:** S.K., T.W., H.C.T., M.R., M.C., A.S., O.O., and B.S. are inventors on a patent application related to this work filed by Carlsberg A/S (no. WO2018/001884, filed on 23 June 2017, published on 4 January 2018). S.K., T.W., M.R., J.T.Ø., and B.S. are inventors on a patent application related to this work filed by Carlsberg A/S (no. WO2021/069614, filed on 8 October 2020, published on 15 April 2021). S.K., S.B., O.O., H.C.T., T.W., and J.H. are inventors on a patent application related to this work filed by Carlsberg A/S (no. WO2019/129736, filed on 21 December 2018, published on 4 July 2019). C.D., P.R.P., S.K., O.O., T.W., M.C., H.C.T., and M.R. are inventors on a patent application related to this work filed by Carlsberg A/S (no. WO2019/129739, filed on 21 December 2018, published on 4 July 2019). The authors declare no other competing interests. All authors are previous or current Carlsberg A/S employees. B.L.M. is a distinguished professor at the Carlsberg Research Laboratory and serves as a consultant for the laboratory. G.B.F. is a senior scientific consultant and distinguished professor of raw materials at the Carlsberg Research Laboratory. The methodology presented in the manuscript is applied

commercially by Traitomic ([www.traitomic.com](http://www.traitomic.com)), a business unit at Carlsberg A/S. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Raw reads of whole-genome sequencing collected in this study have been deposited at the European Nucleotide Archive (ENA) under PRJEB51731. Isolated barley variant seeds obtained using the FIND-IT approach can be provided by the Carlsberg Research Laboratory pending scientific review and a completed material transfer agreement. Requests for FIND-IT variants should be submitted to Christoph.Docker@carlsberg.com.

Submitted 25 March 2022

Accepted 13 July 2022

Published 24 August 2022

10.1126/sciadv.abq2266