Trends in Plant Science

CellPress REVIEWS

Review

Translational Research: Exploring and Creating Genetic Diversity

Pierre Jacob,[1] Adi Avni,[2] and Abdelhafid Bendahmane[1,*]

The crop selection process has created a genetic bottleneck ultimately restricting breeding output. Wild relatives of major crops as well as the so-called 'neglected plant' species represent a reservoir of genetic diversity that remains underutilized. These species could be used as a tool to discover new alleles of agronomic interest or could be the target of breeding programs. Targeted induced local lesions in the genome (TILLING) can be used to translate in neglected crops what has been discovered in major crops and reciprocally. However, random mutagenesis, used in TILLING approaches, provides only a limited density of mutational events at a defined target locus. Alternatively, clustered regularly interspaced short palindromic repeats (CRISPR) associated 9 (Cas9) fused to a cytidine deaminase could serve as a localized mutagenic agent to produce high-density mutant populations. Artificial evolution is at hand.

Genetic Improvement in the 21st Century

Yields have been greatly enhanced in the past 50 years, but extensive breeding has caused a reduction in the genetic variability of major crops. Consequently, plant breeding has narrowed the possibilities for further genetic improvement [1].

Biodiversity has proved to be a valuable reservoir of new alleles to counteract the erosion of this genetic variability (Figure 1) [2–4]. The poorly selected '**neglected crops**' (see Glossary), together with wild relatives of major crops, possess many traits that were lost during the domestication and breeding process. Those traits cannot always be found by conventional screenings in major crops. Neglected crops and wild relatives of major crops could therefore serve as a source of new traits to be transferred in major crops. Alternatively, neglected crops themselves could benefit from those results. The breeding programs developed for major crops have already discovered many valuable alleles and established strategies to modify important traits. Applying these to neglected crops, with present-day know-how, could lead to accelerated genetic improvement with limited investment.

Genetic diversity can also be created through targeted mutagenesis. Thanks to past decades' efforts in plant sciences, many of the genetic regulators controlling key agronomic traits [stress resistance (e.g., DREBs, HSFs), photosynthesis (cytochrome b/f complex), defense against pathogens (R genes)] are known [5–7]. Structure–function analyses have highlighted that many partial phenotypes, like a small modulation of an enzyme's activity, can be obtained from precise, punctual variations [8]. Although we may not always understand the complex underlying mechanisms, producing targeted, empiric genome modifications should allow the precise manipulation of agronomic traits in crops.

Trends

Modern plant breeding suffers from reduced genetic diversity. Investigation of neglected crops and gene editing could contribute to enhancing the genetic variability.

Molecular markers associated with important traits in major crops could be translated in neglected crops, resulting in large yield increases with limited investment.

Wild relatives of crop species could be used to identify new genetic markers associated with traits of interest.

Clustered regularly interspaced short palindromic repeats (CRISPR) associated 9 (Cas9) fused to a cytidine deaminase can perform C-to-T transition without double-strand breaks. This complex can induce multiple mutations in close proximity, raising the possibility of using CRISPR–Cas9 as a mutagenic agent for targeted induced local lesions in the genome (TILLING).

A combination of CRISPR–Cas9 and TILLING could be used to engineer alleles in major and neglected crops, translating known phenotypes or creating new ones.

[1]Institute of Plant Science – Paris-Saclay, INRA, 91190 Gif-sur-Yvette, France
[2]School of Plant Sciences and Food Security, Tel Aviv University, Tel Aviv, Israel

*Correspondence:
abdelhafid.bendahmane@inra.fr
(A. Bendahmane).

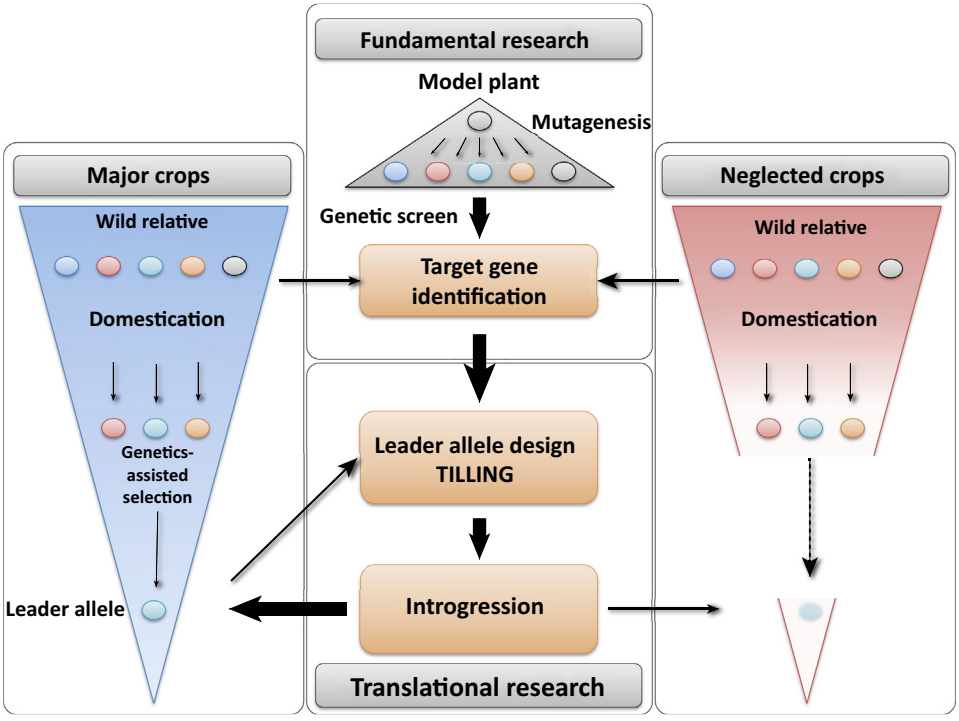

Figure 1. Translational Research: Application to Major and Neglected Crops. Breeding of major crops suffers from low genetic diversity. By contrast, neglected crops exhibit rich genetic biodiversity but suffer from low investment in research and breeding. Translational research can counteract the erosion of genetic biodiversity in major crops and accelerate the breeding of neglected crops. Translational research starts with the discovery of key genes of agronomic importance. Targeted mutagenesis and analysis of natural variations will lead to the identification of allelic series from which leader alleles can be selected. These leader alleles can then be combined to produce plant prototypes. Valuable allele combinations found in major crops can be readily translated to neglected crops using targeted mutagenesis approaches. Wild relatives and neglected plants can also serve as a platform for the discovery of new targets for crop improvement, benefiting both neglected and major crops.

Neglected Crops: A Source of Genetic Diversity and a Target for Genetic Improvement

Neglected crop species are a group of edible plants that used to be broadly cultivated or could be in the future. Approximately 30 plant species are widely cultivated, from the 30 000 edible plant species known [9]. Neglected crop landraces present highly variable genetic material and in many cases contain interesting traits that were lost during the selection process in major crops. They could be used in multiple ways, as illustrated in Figure 1. On the one hand, these neglected species could undergo crop improvement programs to increase yield while conserving a broad stress-resistance potential, among other interesting traits. Neglected crop breeding could benefit from the resources developed in other breeding programs. Genetic markers already identified in major crops could be translated and used for neglected crop breeding [10]. The increasing number of sequenced genomes underlines the high degree of synteny between members of the same family, as was shown for legumes [11]. In cereals, for instance, a cytokinin oxidase was shown to regulate the number of grains per spike [12] and was found to be conserved in wheat (*Triticum aestivum*) and barley (*Hordeum vulgare*) [13,14]. On the other hand, neglected crops could be used in genetic screens to identify new components of stress-resistance pathways. Many loci controlling important phenotypes cannot be identified in major crops because the associated alleles responsible for these traits were lost during domestication [15]. Using wild relatives of major crops with functional alleles should

Glossary

Clustered regularly interspaced short palindromic repeats (CRISPR)–CRISPR associated 9 (Cas9): ribonucleoprotein complex performing targeted double-strand DNA cleavage.
Neglected crops: domesticated plant species underutilized and poorly selected. Neglected crops represent the vast majority of edible plants but their use has been restricted due to particular growing, supply, or use constraints.
Plant base editor (PBE): ribonucleoprotein complex comprising a Cas9 nickase (D10A) fused to APOBEC1 (a cytidine deaminase) and UGI. PBE induces C-to-T transition without DSBs, allowing multiple genomic mutations in close proximity.
Targeted induced local lesions in the genome (TILLING): reverse genetics tool allowing the selection of non-GMO individuals randomly mutated at a locus of interest.

allow better exploitation of the considered trait and contribute to the genetic improvement of both neglected and major crops [16–18].

Bottle gourd (*Lagenaria siceraria* L.) is one such neglected crop. It is a close relative of many cultivated crops, such as melon (*Cucumis melo*), squash (*Cucurbita pepo*), cucumber (*Cucumis sativus*), and watermelon (*Citrullus lanatus*), that are widely spread in tropical, semiarid, and arid regions. The bottle gourd genome has already been partially sequenced [19]. Due to the extent of synteny, the genetic markers developed in the closely related cucurbit species can be easily applied for bottle gourd selection [20]. As a proof of concept, the genetic origin of the 'round fruit' trait was investigated in an F2 population. Synteny with the tomato (*Solanum lycopersicum*) genome and sequence comparison allowed the identification of the locus responsible for the round fruit shape in bottle gourd with minimal sequencing efforts [21]. In addition, many interesting traits had already been identified in the available landraces, such as disease resistance [22] and salt-stress resistance [23]. Thus, neglected crops may represent both a source of valuable traits that could be transferred to related crops and a target for genetic improvement.

Translational Research and Allele engineering

The principle of plant translational research is to apply insights from fundamental science to crop genetic improvement. To achieve this aim, the leader allele, whether engineered in the laboratory or identified in natural variation, must be transferred into the crop of interest. Direct t-DNA-mediated gene transfer would be convenient but faces strong opposition from the public [24,25]. Consequently, alternatives to transgenesis must be explored. Desired alleles can be created through direct crop genome modification (i.e., mutagenesis).

The **TILLING** technique can be used to isolate, from a mutant population, individuals harboring a mutation in the locus of interest [8,26–28]. This reverse genetics technique allows the production of allelic series comprising mutations of differing strength and penetrance, from silent to complete knockout (KO). In particular, intermediate phenotypes can be obtained from partial loss/gain-of-function mutations. These can be especially important when considering proteins with multiple functions. For instance, *Mildew resistance locus o* (*mlo*) mutants were recently selected through TILLING in bread wheat (*Triticum aestivum*) [29]. MLO is a protein with multiple functions beyond conferring sensitivity to powdery mildew (mainly in root development, pollen tube growth, and early leaf senescence [30]). Thus, *mlo*-null mutants are expected to have not only beneficial but also detrimental effects on the performance of corresponding crops. However, no negative effects were observed in wheat mutants engineered with TILLING [29]. Mutant populations are now available for most common crops, making TILLING a powerful tool (Table 1). Once the screening pipeline is set up, finding mutants is fast (a few weeks), and the same population can be used for any gene. Despite the potential of this technique, its use may be limited by the complex ploidy and the high degree of duplicated or functionally redundant genes in crops [8,31]. Another limitation of TILLING comes from the use of a random mutagenic agent. The most commonly used is ethyl methane sulfonate (EMS), which can principally produce G > A and C > T transitions (99%). The rate of EMS-induced mutation is typically around 1/100 kb (Table 1) [8]. On the one hand, a high number of mutations is required for population saturation [32]. On the other hand, a high mutation rate is detrimental to plant health and reproduction, which interferes with the population production process. Moreover, the genetic pool of elite varieties has been carefully optimized and previously selected traits must not be modified by random contaminant mutations. A number of backcrosses are therefore required to clear the genetic pool of unwanted mutations. Finally, mutant populations are produced from cultivars that often differ from the 'elite' cultivars used by breeders. The impact of the mutation on the phenotype may vary from one genetic background

Table 1. List of Every Mutant Population Developed for TILLING Published in a Peer-Reviewed Journal

Organism	Population size	Mutation density (mutation/kb)	Refs
Arabidopsis thaliana	3072	1/300	[56]
	6912	1/170	[57]
Arachis hypogaea	3420	1/967	[58]
Avena sativa	2600	1/40 to 1/20	[59]
Brachypodium distachyon	5731	1/396	[60]
Brassica napus	2604	1/130	[61]
	7110	1/41	
	3158	1/109	[62]
Brassica oleracea	8750	1/447	[63]
Brassica rapa	9216	1/60	[64]
Cucumis melo	4023	1/573	[65]
	2368	1/1500	[66]
Cucumis pepo	1464	1/133	[67]
Cucumis sativus	3331	1/1147	[68]
Eragrostis tef	21 210	1/115 to 1/370	[69]
Glycine max	768	1/140	[70]
	768	1/550	
	529	1/140	
Helianthus annuus	3651	1/475	[71]
	5000	1/480	[72]
Hordeum vulgare	9216	1/1000	[73]
	3148	1/374	[74]
	10 279	1/500	[75]
Linum usitatissimum	4894	1/41	[76]
Lotus japonicus	3697	NA	[77]
Oryza sativa	NA	1/1000	[78]
	2130	1/6190	[79]
	768	1/294	[80]
	768	1/265	
	767	1/135	[81]
Pisum sativum	8000	1/669	[82]
Solanum lycopersicum	8225	1/737	[83]
	13 000	NA	[84]
	1926	1/574	[85]
	4741	1/322	
	4759	1/574	[86]
	3052	1/1237	[87]
Solanum tuberosum	2748	1/91	[88]
Sorghum bicolor	1600	1/526	[89]
Triticum aestivum	10 000	1/24	[90]
	2348	1/37 to 1/23	[91]
	4244	NA	[92]

Table 1. (continued)

Organism	Population size	Mutation density (mutation/kb)	Refs
	630	1/13	[93]
	518	1/22	(N. Rothe, PhD thesis, Kansas State University, 2010)
	4500	1/84	[94]
	2610	1/34 to 1/47	[95]
Triticum monococcum	716	1/1000	(N. Rothe, PhD thesis, Kansas State University, 2010)
	1532	1/92	[96]
Triticum turgidum	8000	1/40	[90]
	1368	1/51	[97]
Zea mays	750	1/485	[98]

to another. Overall, the density of mutations in random mutagenesis is not high enough to conduct a saturated screen for a gene of interest and too high to be innocuous for the plant.

CRISPR–Cas9 as a Mutagenic Agent for Allele Engineering

The **CRISPR–Cas9** system is a two-component targeted-nuclease system. Multiple recent reviews explain the functioning of the CRISPR–Cas9 system in plants in detail [33–35]. The action of CRISPR–Cas9 depends on the recognition of a target sequence. The target genomic DNA sequence comprises 20 nucleotides followed by the protospacer adjacent motif (PAM) NGG. The system is directed to the genomic site using a programmed single-guide RNA (sgRNA) that base pairs with the DNA target, subsequently leading to a site-specific double-strand break (DSB) three or four nucleotides upstream of the PAM sequence. The Cas9 activity is dependent on two nuclease domains, each responsible for cutting one DNA strand. The C-terminal HNH domain cleaves the DNA strand complementary to the sgRNA while the RuvC domain cleaves the other DNA strand. Mutagenesis results when the broken chromosomes are repaired imprecisely through nonhomologous end joining (NHEJ) and small insertions/deletions (indels) are created at the break site. As many indels results in frameshifts, CRISPR–Cas9 is very efficient at inducing complete loss-of-function mutations. Alternatively, introducing two sgRNAs bordering the same locus will lead to complete deletion of the target DNA, ensuring that the loss of function is effective. The DNA editing is performed very quickly after the introduction of the CRISPR machinery, which gives a high probability that the editing mutation will be transmitted to the germline and to the next generation. NHEJ-mediated DNA repair does not allow genome editing as the end result is random. Homologous recombination repair (HRR) of DSB events is rarer in plants but allows precise, knowledge-based modifications. HRR was successfully used to generate targeted homologous recombination between chromosomes [36] and to replace a specific gene with a modified sequence [37,38].

In general, a major hurdle in the use of CRISPR–Cas9 technology in plants is the efficient co-delivery of nuclease, sgRNAs, and/or donor DNA. Protoplasts or callus can be transformed with plasmids containing donor DNA and/or Cas9/sgRNA [39]. Introduction of Cas9 can be performed by transient or stable transformation using *Agrobacterium*. Other delivery methods, such as bombardment or protoplast transformation, can be applied. In addition, an elegant method taking advantage of geminivirus-based replicons was used to provide plant cells with large amounts of donor DNA, allowing high-frequency knock in [40]. Viral vectors were also

used to deliver sgRNAs into a transgenic Cas9-containing tomato with high efficiency and reproducibility [41,42]. Recently, two groups reported the induction of heritable genome modification using *in vitro* preassembled sgRNA–Cas9 ribonucleoprotein complexes [43,44]. In plants many traits are redundantly controlled by gene families. Creating complex traits may need the editing of several genes simultaneously, but equal expression of multiple sgRNAs can be difficult to obtain using one promoter. Taking advantage of the cellular tRNA processing machinery, the introduction of a polycistronic tRNA–gRNA led to the editing of five target loci [45]. More recently, using multiple promoters to express sgRNA allowed the isolation of a sextuple *Arabidopsis* mutant from only 15 T1 plants [46]. If more sgRNAs are used, some 'hot-spot' effects may affect the random distribution of the mutations, as seen in t-DNA insertion mutant populations. Nevertheless, these effects could be mitigated by increasing the size of the population or by controlling the stoichiometry of the multiple sgRNAs introduced.

CRISPR–Cas9 is limited only by our knowledge of how mutations would impact phenotype. Thanks to advances in bioinformatics and genomics techniques, we can now predict and position vital gene features. Disrupting splicing sites, mutating highly conserved domains, and inducing a premature stop codon are all viable strategies to negatively affect the function of a given gene product with high reliability. For these applications CRISPR–Cas9 may be the most convenient strategy. However, mutations enhancing enzymatic activity or modifying a receptor's specificity or affinity cannot be predicted easily and require empirical studies. CRISPR–Cas9 induces mutations through DSB repair mechanisms and thus targeting multiple sites in a single locus would produce deletions. Consequently, classical CRISPR–Cas9 approaches are not well suited to the production of high-density mutant populations, and it cannot be achieved easily through random mutagenesis as has been done in microorganisms for decades. We propose here a novel method that would allow such genetic engineering: CRISPR–apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1 (APOBEC1) extensive allele mutagenesis and TILLING-mediated variability enrichment (CREATE).

The CREATE approach was made possible by a newly designed chimeric nuclease called the base editor (BE) complex. This chimeric protein comprises a deactivated Cas9 fused to a cytidine deaminase (APOBEC1) and can induce mutations without DSBs, allowing multiple simultaneous mutations in close proximity [47,48]. Cytidine deamination leads to the replacement of a cytidine by a uracil, ultimately replacing the G:C pairs with A:T. The dCas9–APOBEC1 complex [base editing 1 (BE1)] can perform sgRNA-guided DNA modifications of every cytidine in a 5-bp window. Deamination is offset by DNA repair mechanisms. DNA repair through the base excision repair (BER) mechanism was decreased by the addition of a uracil glycosylase inhibitor (UGI) to the BE complex. The restoration of the nuclease activity of the Cas9 HNH domain further avoided DNA repair through mismatch repair (MMR) mechanisms (Figure 2). The resulting BE3 complex (Cas9 RuvC D10A, APOBEC1, UGI) was shown to perform base editing with high efficiency (up to 37%) and specificity of action (around 5% indels). In plants this technique was recently successfully applied to tomato, wheat, rice, and maize, with efficiencies ranging from 43.48% in rice to 1.25% in wheat [49,50]. Interestingly, in plants the **plant base editor (PBE)** complex induced mutations from position 3 to 9 bp away from the sgRNA PAM site. Multiple sgRNAs can be used to span the entire locus encoding a domain of interest and produce a targeted mutant population. The stable transformation of one BE3 construct along with multiple sgRNAs should allow the production of a screening population representing immense variability in a very narrow region of the genome (Figure 2). A direct forward screen would allow the identification of new phenotypes or TILLING could be performed to define a subpopulation of interest for screening for new phenotypes.

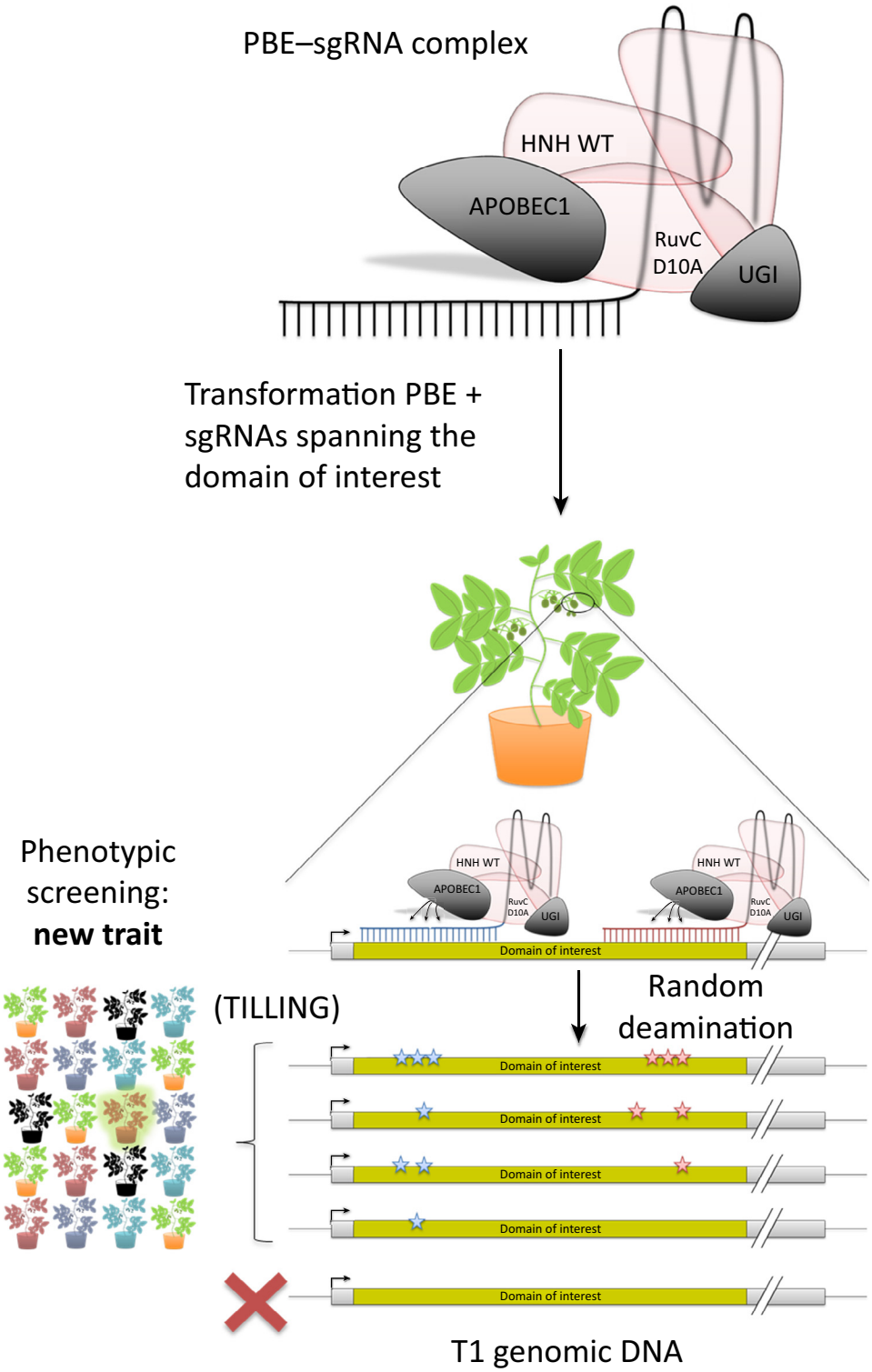

Figure 2. Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)–Apolipoprotein B mRNA Editing Enzyme, Catalytic Polypeptide 1 (APOBEC1) Extensive Allele Mutagenesis and Targeted Induced Local Lesions in the Genome (TILLING)-Mediated Variability Enrichment (CREATE). The plant base editor (PBE) comprises a CRISPR associated 9 (Cas9) nickase (RuvC D10A), a cytidine deaminase (APOBEC1), and a uracyl

(See figure legend on the bottom of the next page.)

Artificial Evolution with CRISPR–Cas9: Application to Disease-Resistance Genes

Using the CREATE method, nucleotide binding site–leucine-rich repeat (NBS-LRR) (also called NLR)-type immune receptors could be engineered to recognize novel pathogen effectors or transduce the signal more efficiently, as has been done with the NBS-LRR resistance to potato virus X (Rx), which confers specific resistance to potato virus X [51]. Using error-prone PCR and transient expression in tobacco, Farnham and Baulcombe were able to engineer a novel resistance specificity conferring on Rx the ability to recognize poplar mosaic virus [52]. NLR-mediated resistance is usually associated with the hypersensitive response (i.e., cell death), a very strong phenotype perfectly adequate for screening. The NBS-LRR-type receptors contain a C-terminal LRR domain and a central domain involved in nucleotide binding called nucleotide binding–apoptotic protease-activating factor 1 (Apaf1), R protein, and *Caenorhabditis elegans* death 4 (CED4) (NB-ARC). The variable N-terminal effector domain most often contains a coiled-coil (CC) or Toll interleukin receptor (TIR) domain in plants. [7]. The LRR domain is commonly thought to be responsible for direct or indirect pathogen recognition, although an unusual immune receptor comprising only a TIR domain was discovered [53]. On ligand binding, the latest models predict that the NLR oligomerizes through the N-terminal domain and this event drives the recruitment of downstream signaling components. In this regard both the N-terminal domain and the LRR domain appear as valuable targets. Resistance to *Pseudomonas syringae* pv maculicola 1 (RPM1) is one of the best studied NLRs in plants. An extensive structure–function study allowed the identification of 95 RPM1 loss-of-function alleles from a mutant population of 119 000 *Arabidopsis thaliana* (476 000 haplotypes) [54]. No missense alleles inducing loss of RPM1 function were found in residues 50–100 (in the CC domain), which suggests that this domain could be used for engineering a novel pathogen response without disrupting existing functions. Could those regions be efficiently targeted with CRISPR–Cas9? There are 13 PAMs in the DNA region encoding residues 50–100. As sgRNAs can be both in forward and reverse orientations, 26 sgRNAs could be used. As the BE3 complex allows deamination of cytidine ranging from 3 bp to 9 bp relative to the PAM sequence, the potential cytidine deamination window would reach a maximum of 182 bp without considering overlaps. Among the targetable nucleotides, 30 are cytidines, which brings the maximum number of mutant alleles to approximately 1 billion, just for the region corresponding to residues 50–100. In the study of Zong *et al*., 40 PBE-mutated rice plants were regenerated and exhibited seven different genotypes. Only four cytidines were present in the deamination window. The majority of those plants were single mutants (24/40; 60%) but the numbers of double and triple mutants were significant with, respectively, 11 (27.5%) and five (12.5%) plants. No quadruple mutant was found, suggesting that the total number of simultaneous mutations produced by PBE is limited. Considering an arbitrary limit of five co-occurring mutations at best would bring the number of alleles that could be generated through this method to approximately 175 000 for the 150-bp region encoding the RPM1 CC domain, using 26 sgRNAs.

TILLING populations over 10 000 individuals are frequently manipulated (Table 1). Producing a transgenic rice plant containing inducible PBE/sgRNA expression would enable the establishment of large mutant populations. If mutation rates as high as those obtained in the study of Zong *et al*. could be achieved, approximately 4000 mutants could be selected from 10 000 offspring of the PBE-induced plants. By comparison, an EMS population of 10 000 individuals typically contains a dozen individuals with mutational events at a single locus. In the study of

glycosylase inhibitor (UGI). Guided by multiple single-guide RNAs (sgRNAs) spanning the domain of interest, the PBE complex induces random mutations in the target sequence. Thanks to the cytidine deaminase activity of APOBEC1, G > A and C > T transitions are induced without double-strand breaks, allowing multiple simultaneous mutations. In the next generation, plant individuals exhibiting novel phenotypes could be identified by forward screens. Alternatively, plant individuals harboring new alleles could be identified by TILLING then phenotyped to identify leader alleles.

Farnham and Baulcombe, where they generated 1920 *Rx* mutants, 20 were gain of function and 13 of these truly acquired new pathogen specificity [52]. However, in this study a limited number of pathogens were tested so the gain-of-function mutation rate could be higher than reported. These numbers demonstrate that recovery of gain-of-function mutants is not a realistic task in EMS-based TILLING approaches. By contrast, the proposed CREATE approach should allow the selection of dozens of gain-of-function mutants. Once a *RPM1* variant population is set up, it could be challenged, in a forward screen, with any *RPM1*-resistant breaking isolate or any virulent pathogen. Alternatively, a subset of PBE-induced *RPM1* variants, or any *R* gene variants, could be first identified by sequencing and then challenged with the pathogen.

In this context DNA libraries from individual plants obtained from PBE-induced variants in *R* genes could be produced and pooled in a 3D pooling strategy. Mutant lines carrying key induced alleles could be identified by the NGS-based approach [55]. Consequently, the identification of individuals with enhanced RPM1 function could take only a few weeks and lead to multiple allele selection with a single deep-sequencing reaction. The introgression of multiple engineered loci could bring durable resistance to pathogens. Using this combination of PBE and TILLING should make genetic engineering a tangible reality.

Concluding Remarks and Future Perspectives

Since the 'green revolution' in the 1960s, classical breeding has been the main cause of yield increase in major crops. However, the greatest potential for improvement resides in neglected crops, which represent approximately 99.9% of edible plants [9]. Future translational research should establish work flows that focus on transfer of know-how not only from model plants to major crops but also from model plants and major crops to neglected crops. Progress in NGS that has dramatically reduced the cost of sequencing should greatly accelerate this process. Many alleles controlling key agronomic traits have already been identified and in many cases introgressed in elite cultivars in major crops, such as the shelf-life trait, one of the most important agronomic traits controlling fruit softening, fruit over-ripening, and susceptibility to opportunistic pathogens. The corresponding alleles should be readily identified and introgressed in neglected plants. Alleles controlling shelf-life in tomato could, for instance, be translated in most fleshy cucurbit fruits.

One could also expect that undiscovered genes/alleles controlling the same traits may be much more suitable for breeding. To identify these genes/alleles, one could take advantage of the biodiversity of neglected crops or wild relatives of major crops. To improve genetic biodiversity, TILLING has proved to be a valuable tool for the production of allelic series associated with a broad spectrum of phenotypes, directly in crops. Nevertheless, large TILLING populations are needed to reach saturation. The use of PBE should allow the production of high-density mutant populations with limited resources. CREATE could facilitate the creation of thousands of mutations at a locus of interest (see Outstanding Questions).

Outstanding Questions

What are the genes regulating key agronomic traits? Are these regulators conserved between model, major, and neglected crop species?

Can we find leader alleles in the existing genetic diversity? Are these the best possible alleles? If not, can they be created?

Can we modulate the activity of master regulators through epigenetic modifications? Can we transfer epigenetic regulation from one species to another?

Can CREATE, a combination of CRISPR–Cas9 and TILLING, be used to generate and identify leader alleles?

Can the deamination window of Cas9–APOBEC1 be increased? What are the alternatives to APOBEC1?

Acknowledgments

Financial support was provided by the European Research Council (ERC SEXYPARTH), the SPS labex (MULTICRISP), and the Plant Biology and Breeding Department of INRA. This work was also supported by Research Grant 4312-10 from BARD and Research Grant 2013227 from the US-Israel Binational Science Foundation.

References

1. Sikora, P. *et al.* (2011) Mutagenesis as a tool in plant genetics, functional genomics, and breeding. *Int. J. Plant Genomics* 2011, 314829
2. Nosenko, T. *et al.* (2016) Adaptation to low temperatures in the wild tomato species *Solanum chilense*. *Mol. Ecol.* 25, 2853–2869
3. Hutin, M. *et al.* (2015) A knowledge-based molecular screen uncovers a broad-spectrum OsSWEET14 resistance allele to bacterial blight from wild rice. *Plant J.* 84, 694–703
4. Yaniv, E. *et al.* (2015) Evaluation of marker-assisted selection for the stripe rust resistance gene Yr15, introgressed from wild emmer wheat. *Mol. Breed.* 35, 1–12

5. Evans, J.R. (2013) Improving photosynthesis. *Plant Physiol.* 162, 1780–1793
6. Jacob, P. *et al.* (2017) The heat-shock protein/chaperone network and multiple stress resistance. *Plant Biotechnol. J.* 15, 405–414
7. Jones, J.D.G. *et al.* (2016) Intracellular innate immune surveillance devices in plants and animals. *Science* 354, aaf6395
8. Chen, L. *et al.* (2014) Progress in TILLING as a tool for functional genomics and improvement of crops. *J. Integr. Plant Biol.* 56, 425–443
9. Longin, C.F.H. and Würschum, T. (2016) Back to the future – tapping into ancient grains for food diversity. *Trends Plant Sci.* 21, 731–737
10. Ellstrand, N.C. *et al.* (2013) Introgression of crop alleles into wild or weedy populations. *Annu. Rev. Ecol. Evol. Syst.* 44, 325–345
11. Cannon, S.B. *et al.* (2009) Three sequenced legume genomes and many crop species: rich opportunities for translational genomics. *Plant Physiol.* 151, 970–977
12. Ashikari, M. (2005) Cytokinin oxidase regulates rice grain production. *Science* 309, 741–745
13. Zhang, L. *et al.* (2012) TaCKX6-D1, the ortholog of rice OsCKX2, is associated with grain weight in hexaploid wheat. *New Phytol.* 195, 574–584
14. Zalewski, W. *et al.* (2010) Silencing of the HvCKX1 gene decreases the cytokinin oxidase/dehydrogenase level in barley and leads to higher plant productivity. *J. Exp. Bot.* 61, 1839–1851
15. Soyk, S. *et al.* (2017) Bypassing negative epistasis on yield in tomato imposed by a domestication gene. *Cell* 169, 1142–1149 e12
16. Yang, S. *et al.* (2016) Identification of QTLs for red fruit firmness using the wild tomato species *Solanum pennellii* LA716 introgression lines. *Plant Breed.* 135, 728–734
17. Millet, E. *et al.* (2014) Introgression of leaf rust and stripe rust resistance from Sharon goatgrass (*Aegilops sharonensis* Eig) into bread wheat (*Triticum aestivum* L.). *Genome* 57, 309–316
18. Soyk, S. *et al.* (2016) Variation in the flowering gene SELF PRUNING 5G promotes day-neutrality and early yield in tomato. *Nat. Genet.* 49, 162–168
19. Xu, P. *et al.* (2011) Partial sequencing of the bottle gourd genome reveals markers useful for phylogenetic analysis and breeding. *BMC Genomics* 12, 467
20. Mashilo, J. *et al.* (2017) Phenotypic and genotypic characterization of bottle gourd [*Lagenaria siceraria* (Molina) Standl.] and implications for breeding: a review. *Sci. Hortic. (Amsterdam)* 222, 136–144
21. Xu, P. *et al.* (2014) Population genomic analyses from low-coverage RAD-seq data: a case study on the non-model cucurbit bottle gourd. *Plant J.* 77, 430–442
22. Ling, K.-S. *et al.* (2013) Development and field evaluation of multiple virus-resistant bottle gourd (*Lagenaria siceraria*). *Plant Dis.* 97, 1057–1062
23. Yang, Y. *et al.* (2015) Bottle gourd rootstock-grafting promotes photosynthesis by regulating the stomata and non-stomata performances in leaves of watermelon seedlings under NaCl stress. *J. Plant Physiol.* 186–187, 50–58
24. Domingo, J.L. (2016) Safety assessment of GM plants: an updated review of the scientific literature. *Food Chem. Toxicol.* 95, 12–18
25. Blancke, S. *et al.* (2015) Fatal attraction: the intuitive appeal of GMO opposition. *Trends Plant Sci.* 20, 414–418
26. McCallum, C.M. *et al.* (2000) Targeting induced local lesions in genomes (TILLING) for plant functional genomics. *Plant Physiol.* 123, 439–442
27. Sharp, P. and Dong, C. (2014) TILLING for plant breeding. *Methods Mol. Biol.* 1145, 155–165
28. Wang, N. and Shi, L. (2015) Screening of mutations by TILLING in plants. *Methods Mol. Biol.* 1245, 193–203
29. Acevedo-Garcia, J. *et al.* (2017) *mlo*-based powdery mildew resistance in hexaploid bread wheat generated by a non-transgenic TILLING approach. *Plant Biotechnol. J.* 15, 367–378
30. Wang, Y. *et al.* (2014) Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol.* 32, 947–951
31. Gauffier, C. *et al.* (2016) A TILLING approach to generate broad-spectrum resistance to potyviruses in tomato is hampered by eIF4E gene redundancy. *Plant J.* 85, 717–729
32. Dalmais, M. *et al.* (2008) UTILLdb, a *Pisum sativum in silico* forward and reverse genetics tool. *Genome Biol.* 9, R43
33. Paquet, D. *et al.* (2016) Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* 533, 125–129
34. Puchta, H. (2016) Using CRISPR/Cas in three dimensions: towards synthetic plant genomes, transcriptomes and epigenomes. *Plant J.* 87, 5–15
35. Schiml, S. and Puchta, H. (2016) Revolutionizing plant biology: multiple ways of genome engineering by CRISPR/Cas. *Plant Methods* 12, 8
36. Filler Hayut, S. *et al.* (2017) Targeted recombination between homologous chromosomes for precise breeding in tomato. *Nat. Commun.* 8, 15605
37. Čermák, T. *et al.* (2015) High-frequency, precise modification of the tomato genome. *Genome Biol.* 16, 232
38. Butt, H. *et al.* (2017) Efficient CRISPR/Cas9-mediated genome editing using a chimeric single-guide RNA molecule. *Front. Plant Sci.* 8, 1441
39. Sauer, N.J. *et al.* (2016) Oligonucleotide-mediated genome editing provides precision and function to engineered nucleases and antibiotics in plants. *Plant Physiol.* 170, 1917–1928
40. Baltes, N.J. *et al.* (2014) DNA replicons for plant genome engineering. *Plant Cell* 26, 151–163
41. Mahfouz, M.M. *et al.* (2014) Genome engineering via TALENs and CRISPR/Cas9 systems: challenges and perspectives. *Plant Biotechnol. J.* 12, 1006–1014
42. Ali, Z. *et al.* (2016) CRISPR/Cas9-mediated immunity to geminiviruses: differential interference and evasion. *Sci. Rep.* 6, 26912
43. Svitashev, S. *et al.* (2016) Genome editing in maize directed by CRISPR–Cas9 ribonucleoprotein complexes. *Nat. Commun.* 7, 13274
44. Woo, J.W. *et al.* (2015) DNA-free genome editing in plants with preassembled CRISPR–CAS9 ribonucleoproteins. *Nat. Biotechnol.* 33, 1162–1164
45. Xie, K. *et al.* (2015) Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc. Natl. Acad. Sci. U. S. A.* 112, 3570–3575
46. Zhang, Z. *et al.* (2016) A multiplex CRISPR/Cas9 platform for fast and efficient editing of multiple genes in Arabidopsis. *Plant Cell Rep.* 35, 1519–1533
47. Komor, A.C. *et al.* (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424
48. Plosky, B.S. (2016) CRISPR-mediated base editing without DNA double-strand breaks. *Mol. Cell* 62, 477–478
49. Zong, Y. *et al.* (2017) Precise base editing in rice, wheat and maize with a Cas9–cytidine deaminase fusion. *Nat. Biotechnol.* 35, 438–440
50. Shimatani, Z. *et al.* (2017) Targeted base editing in rice and tomato using a CRISPR–Cas9 cytidine deaminase fusion. *Nat. Biotechnol.* 35, 441–443
51. Bendahmane, A. (1999) The *Rx* gene from potato controls separate virus resistance and cell death responses. *Plant Cell Online* 11, 781–792
52. Farnham, G. and Baulcombe, D.C. (2006) Artificial evolution extends the spectrum of viruses that are targeted by a disease-resistance gene from potato. *Proc. Natl. Acad. Sci. U. S. A.* 103, 18828–18833
53. Nishimura, M.T. *et al.* (2017) TIR-only protein RBA1 recognizes a pathogen effector to regulate cell death in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.* 114, E2053–E2062
54. Tornero, P. *et al.* (2002) Large-scale structure–function analysis of the *Arabidopsis* RPM1 disease resistance protein. *Plant Cell Online* 14, 435–450

55. Tsai, H. *et al.* (2011) Discovery of rare mutations in populations: TILLING by sequencing. *Plant Physiol.* 156, 1257–1268
56. Greene, E.A. *et al.* (2003) Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in *Arabidopsis*. *Genetics* 164, 731–740
57. Till, B.J. *et al.* (2003) Large-scale discovery of induced point mutations with high-throughput TILLING. *Genome Res.* 13, 524–530
58. Knoll, J.E. *et al.* (2011) TILLING for allergen reduction and improvement of quality traits in peanut (*Arachis hypogaea* L.). *BMC Plant Biol.* 11, 81
59. Chawade, A. *et al.* (2010) Development and characterization of an oat TILLING-population and identification of mutations in lignin and β-glucan biosynthesis genes. *BMC Plant Biol.* 10, 86
60. Dalmais, M. *et al.* (2013) A TILLING platform for functional genomics in *Brachypodium distachyon*. *PLoS One* 8, e65503
61. Wang, N. *et al.* (2008) A functional genomics resource for *Brassica napus*: development of an EMS mutagenized population and discovery of *FAE1* point mutations by TILLING. *New Phytol.* 180, 751–765
62. Gilchrist, E.J. *et al.* (2013) A mutant *Brassica napus* (canola) population for the identification of new genetic diversity via TILLING and next generation sequencing. *PLoS One* 8, e84303
63. Himelblau, E. *et al.* (2009) Forward and reverse genetics of rapid-cycling *Brassica oleracea*. *Theor. Appl. Genet.* 118, 953–961
64. Stephenson, P. *et al.* (2010) A rich TILLING resource for studying gene function in *Brassica rapa*. *BMC Plant Biol.* 10, 62
65. Dahmani-Mardas, F. *et al.* (2010) Engineering melon plants with improved fruit shelf life using the TILLING approach. *PLoS One* 5, e15776
66. González, M. *et al.* (2011) Towards a TILLING platform for functional genomics in Piel de Sapo melons. *BMC Res. Notes* 4, 289
67. Vicente-Dólera, N. *et al.* (2014) First TILLING platform in *Cucurbita pepo*: a new mutant resource for gene function and crop improvement. *PLoS One* 9, e112743
68. Boualem, A. *et al.* (2014) Development of a *Cucumis sativus* TILLinG platform for forward and reverse genetics. *PLoS One* 9, e97963
69. Zhu, Q. *et al.* (2012) High-throughput discovery of mutations in tef semi-dwarfing genes by next-generation sequencing analysis. *Genetics* 192, 819–829
70. Cooper, J.L. *et al.* (2008) TILLING to detect induced mutations in soybean. *BMC Plant Biol.* 8, 9
71. Sabetta, W. *et al.* (2011) sunTILL: a TILLING resource for gene function analysis in sunflower. *Plant Methods* 7, 20
72. Kumar, A.P. *et al.* (2013) SMART – sunflower mutant population and reverse genetic tool for crop improvement. *BMC Plant Biol.* 13, 38
73. Caldwell, D.G. *et al.* (2004) A structured mutant population for forward and reverse genetics in barley (*Hordeum vulgare* L.). *Plant J.* 40, 143–150
74. Talamè, V. *et al.* (2008) TILLMore, a resource for the discovery of chemically induced mutants in barley. *Plant Biotechnol. J.* 6, 477–485
75. Gottwald, S. *et al.* (2009) TILLING in the two-rowed barley cultivar 'Barke' reveals preferred sites of functional diversity in the gene *HvHox1*. *BMC Res. Notes* 2, 258
76. Chantreau, M. *et al.* (2013) PT-Flax (phenotyping and TILLinG of flax): development of a flax (*Linum usitatissimum* L.) mutant population and TILLinG platform for forward and reverse genetics. *BMC Plant Biol.* 13, 159
77. Perry, J.A. *et al.* (2003) A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus*. *Plant Physiol.* 131, 866–871
78. Wu, J.-L. *et al.* (2005) Chemical- and irradiation-induced mutants of Indica rice IR64 for forward and reverse genetics. *Plant Mol. Biol.* 59, 85–97
79. Sato, Y. *et al.* (2006) Mutant selection from progeny of gamma-ray-irradiated rice by DNA heteroduplex cleavage using *Brassica* petiole extract. *Breed. Sci.* 56, 179–183
80. Till, B.J. *et al.* (2007) Discovery of chemically induced mutations in rice by TILLING. *BMC Plant Biol.* 7, 19
81. Suzuki, T. *et al.* (2008) MNU-induced mutant pools and high performance TILLING enable finding of any gene mutation in rice. *Mol. Genet. Genomics* 279, 213–223
82. Triques, K. *et al.* (2007) Characterization of *Arabidopsis thaliana* mismatch specific endonucleases: application to mutation discovery by TILLING in pea. *Plant J.* 51, 1116–1125
83. Gady, A.L. *et al.* (2009) Implementation of two high through-put techniques in a novel application: detecting point mutations in large EMS mutated plant populations. *Plant Methods* 5, 13
84. Menda, N. *et al.* (2004) *In silico* screening of a saturated mutation library of tomato. *Plant J.* 38, 861–872
85. Minoia, S. *et al.* (2010) A new mutant genetic resource for tomato crop improvement by TILLING technology. *BMC Res. Notes* 3, 69
86. Piron, F. *et al.* (2010) An induced mutation in tomato eIF4E leads to immunity to two potyviruses. *PLoS One* 5, e11313
87. Okabe, Y. *et al.* (2012) Availability of Micro-Tom mutant library combined with TILLING in molecular breeding of tomato fruit shelf-life. *Breed. Sci.* 62, 202–208
88. Muth, J. *et al.* (2008) Precision breeding for novel starch variants in potato. *Plant Biotechnol. J.* 6, 576–584
89. Xin, Z. *et al.* (2008) Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population. *BMC Plant Biol.* 8, 103
90. Slade, A.J. *et al.* (2005) A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nat. Biotechnol.* 23, 75–81
91. Dong, C. *et al.* (2009) Simultaneous mutation detection of three homoeologous genes in wheat by high resolution melting analysis and Mutation Surveyor®. *BMC Plant Biol.* 9, 143
92. Sestili, F. *et al.* (2010) Production of novel allelic variation for genes involved in starch biosynthesis through mutagenesis. *Mol. Breed.* 25, 145–154
93. Feiz, L. *et al.* (2009) Creation and functional analysis of new *Puroindoline* alleles in *Triticum aestivum*. *Theor. Appl. Genet.* 118, 247–257
94. Fitzgerald, T.L. *et al.* (2010) A high-throughput method for the detection of homoeologous gene deletions in hexaploid wheat. *BMC Plant Biol.* 10, 264
95. Chen, L. *et al.* (2012) Development and characterization of a new TILLING population of common bread wheat (*Triticum aestivum* L.). *PLoS One* 7, e41570
96. Rawat, N. *et al.* (2012) A diploid wheat TILLING resource for wheat functional genomics. *BMC Plant Biol.* 12, 205
97. Uauy, C. (2013) An *in-silico* functional genomics resource: targeted re-sequencing of wheat TILLING mutant populations. In *Plant and Animal Genome Conference XXI 2013*. International Plant and Animal Genome Conference
98. Till, B.J. *et al.* (2004) Discovery of induced point mutations in maize genes by TILLING. *BMC Plant Biol.* 4, 12