Precise, flexible and affordable gene stacking for crop improvement

Weiqiang Chen and David W. Ow

ARTICLE HISTORY
Received 14 December 2016
Accepted 20 December 2016

ABSTRACT
The genetic engineering of plants offers a revolutionary advance for crop improvement, and the incorporation of transgenes into crop species can impart new traits that would otherwise be difficult to obtain through conventional breeding. Transgenes introduced into plants, however, can only be useful when bred out to field cultivars. As new traits are continually added to further improve transgenic cultivars, clustering new DNA near previously introduced transgenes keeps from inflating the number of segregating units that breeders must assemble back into a breeding line. Here we discuss various options to introduce DNA site-specifically into an existing transgenic locus.

Introduction
The world population tripled from 2.5 billion in 1950 to 7.4 billion in 2015 and much of that is attributed to the less developed world that increased 3.6-fold from 1.7 billion in 1950 to 6.1 billion in 2015. With this trend, the projected population in less developed regions will reach up to 8.4 billion by 2050. Yet the world per capita arable land has been decreasing from 0.45 hectares in 1960 to a projected 0.18 hectares by 2050, and even lower for the less developed regions at 0.14 hectares. Not only will our world need to grow more food, but we must grow them with greater efficiency than in the past.

Although there are various solutions to increase food production, one solution that cannot be ignored is the bioengineering of crops through the tinkering of DNA. In 2015, global bioengineered (genetically modified) crops covered ~180 million hectares, of which a third of them contain more than a single trait, or stacked traits. More than half (54%) of global biotech crops were grown by developing countries from Latin America, Asia and Africa. Despite the fact that to commercialize a biotech crop is not trivial, requiring a decade and a hundred million US dollars, the investment in plant molecular genetics research in public and private sectors should produce a steady stream of new trait genes for crop improvement. This will inevitably offer the opportunity to bioengineer better crops through transgenesis or gene editing.

In transgenesis, a transgenic trait not only needs to be incorporated into a crop genome, but also must be bred out to elite lines that grow in the field. Due to the technical difficulty in transforming elite lines (field cultivars), transformation is typically conducted with a laboratory line (or a field line that can be transformed). The transgenic trait is then introgressed into a variety of different elite lines through repeated back-crossing to a recipient genotype. This practice may not necessarily be disadvantageous, as once a trait is deregulated, i.e., approved by the regulatory authorities, that one deregulated event can then be bred out to the many commercial cultivars that grow in...
different parts of the world. The conversion from a laboratory line to an elite line requires that non-elite traits are crossed out of the final field cultivar. For crops that behave genetically as diploids, if the relevant genetic difference between a laboratory line and an elite line (or between one elite line of one location and another elite line of another location) resides in 8 different unlinked loci, then the probability of assembling 8 segregating elite loci into a homozygous individual from their cross hybridization would be \( \left(\frac{1}{4}\right)^8 \) or 1 individual plant out of 65,536. Adding a first transgene, representing another independently segregating locus, would increase the probability to \( \left(\frac{1}{4}\right)^9 \) or 1 individual plant out of 262,144. As more new traits are discovered through the years, and more transgenic loci are added to the genome at new locations, this would make the line conversion process more and more difficult. For example, adding another transgenic locus to the above example would yield the probability of assembling 10 independently segregating loci as \( \left(\frac{1}{4}\right)^{10} \) or 1 homozygous individual out of 1,048,576, and this illustration does not even factor in the possibility that some traits exhibit linkage drag, i.e., they do not segregate independently.

In practice, a breeder cannot possibly obtain such large pool of progenies from a single cross, but instead could only assemble several traits from each round of backcross. Nonetheless, the statistical probabilities illustrate the rapidly increasing difficulty with introgressing more and more transgenic traits that segregate independently, resulting in more and more backcrosses needed. To make matters worse, line conversion is not conducted just with a single elite line, but with numerous elite lines grown commercially in many parts of the world. This makes the breeding process a bottleneck step in the bioengineering of crop plants.

There are several solutions around the problem of increasing transgenic loci. The first is the “do-over” strategy. That is, multiple transgenes can be packaged into one single transformation event. This can be done through stacking multiple transgenes in vitro into a single DNA cassette, and the cassette can be introduced into the plant genome as a large contiguous DNA fragment. This means that previously introduced traits will be combined with new traits to integrate the entire package into a new genomic location. Developers will have to screen the hundreds to thousands of random integration events to find an insertion with appropriate expression of all transgenes introduced, with an integration structure that is relatively precise, and without causing unwanted changes in the plant due to its integration event. Moreover, although old traits had received prior deregulation in previously introduced cultivars, they will nonetheless have to go through the deregulation process again since the new integration event is no longer the same as in previously deregulated cultivars. This approach would have developers expending greater effort in selecting for a line with field efficacy of the numerous transgenic traits, as well as in deregulating and rederegulating the numerous transgenes for commercial release.

Another solution is the “direct transformation into elite line” strategy. By advancing transformation efficiency of elite cultivars, or cultivars genetically closer to field cultivars, developers can exercise the option of introducing new transgenes directly into lines that will require fewer backcrosses to field cultivars. Success in transforming elite lines has been progressing steadily. However, with most crops, there are numerous different elite lines adapted to a wide range of habitats, and hence this approach would require numerous independent transformation events. The issue of regulatory approval would also magnify, as there would be a large number of independent integration events that must be deregulated, even though they all contain the same transgene.

A third solution would be the in vivo site specific integration of additional transgenes to the previously engineered transgenic locus, whether into an elite genome or into a laboratory variety that is easier to transform. As before, the appended transgenic locus from “site specific gene stacking” must be introgressed to the numerous locally adapted cultivars. However, this method maintains the original number of segregating loci to ease the downstream breeding effort. This approach could also ease regulatory approval of the new DNA as it resides in a previously deregulated genome location. Moreover, the previously placed transgenes would not be considered a new transgenic event that needs a fresh round of deregulation. In principle, site specific integration of additional DNA to a previously engineered transgenic locus could be accomplished using either plant endogenous homologous recombination or site specific recombination.

As the breakage of double stranded DNA enhances the normally low homologous recombination
efficiency, sequence specific nucleases (SSN) have been developed to break double stranded DNA at specified sequences. These SSN include zinc-finger nuclease (ZFN), transcription activator-like effector nuclease (TALEN), homing endonuclease/meganuclease, clustered regulatory interspaced short palindromic repeats (CRISPR)/CRISPR associated protein (Cas) and the recently described argonaute protein from Natronobacterium gregoryi (NgAgo). These SSN appear highly efficient for generating mutations through a non-homologous end-joining reaction, with the exception of the NgAgo system, in which there are currently some issues concerning its reliability. In all likelihood, the next generation of bioengineered crops would comprise not only of those obtained by transgenesis, but also by SSN-directed gene editing.

For inserting a transgene, the use of SSN requires that a double-stranded break be repaired from a homologous template of the new DNA. This repair pathway is not as efficient as the non-homologous end-joining pathway, but success has been reported for ZFN and meganuclease-mediated integration of a new DNA fragment into a preexisting transgenic locus in major crop plants. This demonstrates their potential use for repeated gene integration at the same genomic location.

Whether SSN is efficient for transgene insertion is not the most critical concern with commercial developers, as extra time and labor can compensate for operational inefficiency. The major concern is the issue of whether there is freedom to operate (FTO). Each of the SSN is patented and obtaining a license may be prohibitively expensive such that only major crop biotech companies can afford their use. For the commercial development of major crops, ZFN has been exclusively licensed since 2008 to Dow Agrosciences through Sangamo Biosciences, and recently Monsanto and Dow Agrosciences announced a non-exclusive global option and licensing agreement on using ZFN for crops. Through Vilnius University since June, 2015, Dupont has obtained exclusive licensing for all commercial use of CRISPR/Cas9, as well as announced a strategic alliance with Caribou Biosciences on CRISPR/Cas9. Likewise, Monsanto has received a license for use of CRISPR/Cas9 in agricultural applications from the MIT/Harvard Broad Institute. However, there is ongoing litigation on the CRISPR/Cas9 technology that could take years to sort out. The licensing of TALEN is less clear, although Calyxt/Cellectis Plant Sciences (Minnesota) has obtained rights to practice the invention. Whether NgAgo would work in crop plants is currently not known given its reproducibility problem in mammalian cells, as well as whether it will be deemed sufficiently novel and unobvious to be issued a patent. Thus, until patents expire, the accessibility of developing countries to use these SSN for the commercial development of bioengineered crop does not look promising. Yet it is in these developing countries where food security is most at risk.

In contrast to the use of SSN, the use of certain site-specific recombinases has FTO. Hou et al. have described a recombinase-mediated gene stacking (RMGS) system through demonstration of 2 rounds of gene stacking in tobacco, as well as an important step of removing unneeded DNA from the gene stack (Fig. 1A). The Mycobacteriophage Bxb1 integrase was used to integrate new DNA into the plant genome at specific attachment sites previously placed in the genome, either attB or attP. In the absence of an excisionase, the integrase-catalyzed recombination between attB and attP is not reversible. The Coliphage P1 Cre recombinase that performs reversible site-specific recombination on lox recombination sites was used subsequently to remove unneeded site-specific recombination of donor molecules, additional DNA can be precisely appended to the locus. Therefore, the Bxb1/Cre mediated gene stacking system provides the solution of not increasing the number of segregating loci when future crop plants are improved through the addition of more transgenes. As importantly, Cre mediated excision of unneeded DNA addresses biosafety issues, whether real or perceived, especially regarding selectable marker genes that encode antibiotic resistance.

In the article by Hou et al., an additional set of RS2 recombination sites was included to flank the transgenes, which could permit future deletion of the RS2-flanked transgenes by the Acinetobacter-derived CinH recombinase. For example, this might be useful for the root-specific deletion of transgenes in certain crops, such as for creating transgene-free carrots derived from transgenic plants.
Recently, we also developed a recombinase-mediated in vitro gene stacking system where the Streptomyces phage phiC31 (ΦC31) integrase was used to co-integrate separate circular DNA molecules, followed by using Cre to delete unnecessary plasmid backbone (Fig. 1B). The in vitro derived gene stack can be incorporated into an Agrobacterium vector by Bxb1 recombination in vitro, permitting the transfer of the transgene package to a single chromosome location, and in the process creates a new target site for the subsequent in vivo stacking of new DNA. Alternatively, the in vitro-derived gene stack can be integrated directly into the plant genome in vivo at a pre-existing chromosomal target. Being able to stack DNA in vitro as well as in vivo, and with compatibility between the 2 systems, brings new flexibility for using the RMGS system.

Most important of all, the RMGS system was designed with FTO. The Dupont Cre-Lox patent had expired and the patent applications on the Bxb1-att, and the CinH-RS2 systems, previously filed by the USDA had been abandoned. Although the particular method of DNA delivery described by Hou et al. was protoplast transformation, which has low applicability for major crop plants, recent work by Li et al. have
described a detailed protocol using biolistics to deliver the DNA into rice embryogenetic callus. Regenerated plants with precise site-specific integration events have been obtained and this biolistic-mediated site-specific integration method should be applicable to a wide variety of crop plants. As biolistics is a technology described since 1987 patent claims have long expired.27 Like the in vivo gene stacking system, the in vitro gene stacking system also has FTO. Although patents have been granted for the in vivo use of the phiC31 recombination system, the in vitro use of this system is not patented.

Recombinases, though proficient in site-specific recombination, cannot edit genes the same way as SSN. However, if obtaining a commercial license for SSN is not possible or affordable, there are other options. Though much less efficient than SSN, forward mutagenesis can generate most of the knockout or knock-down mutations, and these mutations can be screened by a gel electrophoresis method known as TILLING (targeting induced local lesions in genomes).30 Transgenes can also be engineered to suppress or enhance native gene expression. Thus, despite recent company mergers that further limit the control of food production to even fewer multinational agbio-tech corporations, developing countries and small biotech developers can still practice crop improvement through transgenesis and forward genetic approaches, and this open source RMGS system provides a tool for developing countries to solve their food security challenge.

Disclosure of potential conflicts of interest
No potential conflicts of interest were disclosed.

Funding
Supported by Chinese Ministry of Agriculture Grant 2010ZX08010-001 and Guangdong Province, China Talent Funds 2010.

References
[1] United Nations, Department of Economic and Social Affairs, Population Division. World Population Prospects: The 2015 Revision, DVD Edition. 2015.
[2] Alexandratos N, Bruinsma J. World agriculture towards 2030/2050: the 2012 revision. ESA Working paper No. 12-03. Rome, FAO. 2012.
[3] James C. 20th Anniversary of the Global Commercialization of Biotech Crops (1996 to 2015) and Biotech Crop Highlights in 2015. ISAAA Brief No. 51. ISAAA: Ithaca, NY. 2015.
[4] McDougall P. The cost and time involved in the discovery, development and authorisation of a new plant biotechnology derived trait. https://croplife.org/plant-biotechnology/regulatory-2/cost-of-bringing-a-biotech-crop-to-market/
[5] Voytas DF. Plant genome engineering with sequence-specific nucleases. Annu Rev Plant Biol, 2013; 64:327-50; PMID: 23451779; https://doi.org/10.1146/annurev-arplant-042811-105552
[6] Weeks DP, Spalding MH, Yang B. Use of designer nucleases for targeted gene and genome editing in plants. Plant Biotechnol J, 2016; 14(2):483-95; PMID: 26261084; https://doi.org/10.1111/pbi.12448
[7] Daboussi F, Stoddard TJ, Zhang F. Engineering meganuclease for precise plant genome modification. In: Zhang F, Puchta H, Thomson JG, editors. Advances in New Technology for Targeted Modification of Plant Genomes. 2015. p. 21-38; https://doi.org/10.1007/978-1-4939-2556-8_2
[8] Puchta H. Applying CRISPR/Cas for genome engineering in plants: the best is yet to come. Curr Opin Plant Biol, 2017; 36:1-8; https://doi.org/10.1016/j.copb.2016.11.011
[9] Gao F, Shen XZ, Jiang F, Wu Y, Han C. DNA-guided genome editing using the Natronobacterium gregoryi Argonaute. Nat Biotechnol, 2016; 34(7):678-73; PMID: 27136078; https://doi.org/10.1038/nbt.3547
[10] Cyranoski, D. Gene-editing row escalates. Nature, 2016; 540:20-1; (Updated: NgAgo gene-editing controversy escalates in peer-reviewed papers) http://go.nature.com/2fuJd8m; PMID: 27905463; https://doi.org/10.1038/nature.2016.21023
[11] Ainley WM, Sastry-Dent L, Welter ME, Murray MG, Zeiter B, Amora R, Corbin DR, Miles RR, Arnold NL, Strange TL, et al. Trait stacking via targeted genome editing. Plant Biotechnol J, 2013; 11(9):1126-34; PMID: 23953646; https://doi.org/10.1111/pbi.12107
[12] D’Halluin K, Vanderstraeten C, Van Hulle J, Rosolowska J, Van Den Brande I, Pennewaert A, D’Hnot K, Bossut M, Jantz D, Ruiter R, et al. Targeted molecular trait stacking in cotton through targeted double-strand break induction. Plant Biotechnol J 2013; 11(8):933-41; PMID: 23777410; https://doi.org/10.1111/pbi.12085
[13] Sangamo BioScience. Springfield, Illinois, United States. Dow agroSciences announces early exercise of option for commercial license with Sangamo BioScience for plants. 2008 Jun 18. http://investor.sangamo.com/releasedetail.cfm?releaseid=317375
[14] Monsanto. St. Louis, Missouri, United States. Monsanto and Dow AgroSciences announce global licensing agreement on EXZACT precision technology genome-editing platform. 2016 Oct 3. http://news.monsanto.com/press-release/corporate/monsanto-and-dow-agrosciences-announce-global-licensing-agreement-exzact-pre
[15] DuPont Pioneer. Wilmington, Delaware, United states. DuPonp pioneer gains exclusive license for genome-
[16] Caribou BioScience. Berkeley, California, United States. Caribou BioScience and Dupont announce strategic alliance. 2015 Oct 8. http://cariboubio.com/in-the-news/press-releases/caribou-biosciences-and-dupont-announce-strategic-alliance

[17] Monsanto. St. Louis, Missouri, United States. Monsanto announces global licensing agreement on Broad Institute on key genome-editing application. 2016 Sep 22. http://news.monsanto.com/press-release/corporate/monsanto-announces-global-licensing-agreement-broad-institute-key-genome-edi

[18] Ledford H. Bitter fight over CRISPR patent heats up. Nature 2016; 529:265; PMID: 26791698; https://doi.org/10.1038/nature.2015.17961

[19] Calyxt/Cellectis. Minnesota, United States. Cellectis Plant Science and Two Blades Foundation announce the execution of a cross-license agreement on TAL effector nuclease technologies. 2014 Dec 18. http://www.cellectis.com/en/content/cellectis-plant-sciences-and-two-blades-foundation-announce-cross-license-agreement-tal-effector-nuclease

[20] Hou L, Yau YY, Wei J, Han Z, Dong Z, Ow DW. An open-source system for in planta gene stacking by Bxb1 and Cre recombinases. Mol Plant, 2014; 7(12):1756-65; PMID: 25281665; https://doi.org/10.1093/mp/ssu107

[21] Dale EC, Ow DW. Gene-transfer with subsequent removal of the selection gene from the host genome. Proc Natl Acad Sci USA, 1991; 88(23):10558-62; PMID: 1660141; https://doi.org/10.1073/pnas.88.23.10558

[22] Chen W, Ow DW. Protocol for In Vitro Stacked Molecules Compatible with In Vivo Recombinase-Mediated Gene Stacking. In: Murata M, editor. Chromosome and Genomic Engineering in Plants: Methods and Protocols, Methods in Molecular Biology. 2016; p.31-47. https://doi.org/10.1007/978-1-4939-4931-1_3

[23] Ow DW. The long road to recombinase-mediated plant transformation. Biotechnol J 2016, 14:441-7. https://doi.org/10.1111/bti.12472

[24] Odell JT, Russell SH, Sauer BL, Hsu FC, Shen JB-J. Site-specific recombination of DNA in plant cells. U.S. patent 5,658,772. Issued Aug 19. 1997.

[25] Ow DW, Thomson JG. Site-specific recombination systems for use in eukaryotic cells. U.S. patent publication number US 2006/0046294 A1.

[26] Li R, Han Z, Hou L, Kaur G, Yin Q, Ow DW. Method for Biolistic Site-Specific Integration in Plants Catalyzed by Bxb1 Integrase. In: Murata M, editor. Chromosome and Genomic Engineering in Plants: Methods and Protocols, Methods in Molecular Biology. 2016; p.15-30. https://doi.org/10.1007/978-1-4939-4931-1_2

[27] Klein TM, Wolf ED, Wu R, Sanford JC. High-velocity microprojectiles for delivering nucleic-acids into living cells. Nature 1987; 327:70-3; https://doi.org/10.1038/327070a0

[28] Calos MP. Methods and composition for genomic modification. U. S. Patent 6,632,672, issued October 14, 2003. 2003.

[29] Ow DW, Calendar R, Thomason L. DNA recombination in eucaryotic cells by the bacteriophage phiC31 recombination system. U.S. Patent 6,746,870, issued June 8, 2004. 2004.

[30] Colbert T, Till BJ, Tompa R, Reynolds S, Steine MN, Yeung AT, McCallum CM, Comai L, Henikoff S. High-throughput screening for induced point mutations. Plant Physiol 2001; 126(2):480-4; PMID: 11402178; https://doi.org/10.1104/pp.126.2.480