GM Crops 1:3, 120-127; May/June 2010 © 2010 Landes Bioscience

Using genetic variability available in the breeder's pool to engineer fruit quality

Diego Orzaez, Antonio J. Monforte and Antonio Granell*

Fruit Genomics and Biotechnology and Molecular Markers Labs; Instituto de Biología Molecular y Celular de Plantas; Valencia, Spain

Key words: genomics, plant breeding, QTL, gene-of-interest, cisgenesis, intragenesis, seamless cloning, phytobrick

We substantiate here the opinion that experts in biotechnology and natural biodiversity can work together on the production of successive waves of next-generation GM fruit crops to improve organoleptic and nutritional quality and therefore generate wider public acceptance. In this scenario genetic engineering becomes a faster and more precise way of transferring genes of interest to fruit crop plants from the same or sexually compatible species (intra- or cisgenesis) than more traditional methods, such as MASPB. The availability of complete genome sequences for an increasing number of crop plants, as well as the results from genomics studies, can assist in the identification of gene-to-trait association. The next wave of GM crops will be able to take full advantage of a Synthetic Biology-based strategy in the development of new fruit varieties by using DNA not necessarily present in the breeder's pool for a wide range of applications.

There are still a number of challenges which require attention, such as identifying genes and allelic forms associated with traits of interest and improving the precision and stability of the transferred DNA.

Introduction

The availability of sufficient genetic variability is a prerequisite for success for any breeding program. The need to preserve and make that genetic variability accessible was already identified in the early 20th century. Nowadays, breeders have access to enormous amounts of genetic variability from cultivated and wild species stored in international germplasm banks.[1] New breeding methods using molecular markers and Quantitative Trait Locus (QTL) analysis, such as Advanced Backcross QTL analysis,[2] have been developed to efficiently incorporate new genetic variability into modern varieties. However, this potential for the improvement of varieties has only been used for a small number of traits, such as increased productivity, fruit size, durability, disease resistance and abiotic stress tolerance.[3] Quality has not only been neglected as an objective, but, in many crops, the increase in productivity has been accompanied by a decline in quality. This is the case, for instance, of the tomato, where the organoleptic quality is low in varieties that have been selected mainly for yield, resistance to pathogens and long shelf life. The reasons for this decline in fruit quality include: (1) traits other than fruit quality being used as the major selection force; (2) a failure to understand the nature of the genes responsible for complex traits such as quality traits; and (3) the limited precision of traditional breeding methods when introgressing one gene from a given source into an elite inbred line. The latter phenomenon is also called "linkage drag", and is a consequence of the linkage of the target gene with other genes with undesirable effects that are also dragged during the breeding process when the target gene is selected. Multi-resistant modern tomato cultivars are a clear example of the possible detrimental effects of linkage drag. Most resistance genes are introgressed from wild species, and for each resistance gene additional wild species genes are introgressed, so one may expect undesirable linkage drag effects in those varieties that have multiple resistance genes from wild species.

There have been efforts to move the emphasis in traditional breeding to quality, at least in several large projects pushed by public interest (in Solanaceae, for instance, see: SOL http://solgenomics.net/solanaceae-project/, EUSOL http://www.eu-sol.net/science, ESPSOL https://chirimoyo.ac.uma.es/espsol/, etc.). From these and other projects we have learned how to map QTLs involved in important quality traits, such as metabolite content, fruit size, etc., in order to find linked markers and discover new useful allelic variability from wild species.[4] Databases, such as those in **Table 1**, that gather results from multiple independent QTL experiments, have been created in order to facilitate breeders' access to the data. However, all this information has not yet been incorporated into applied breeding programs, except for very few examples such as in reference five. This is primarily due to the low resolution of the QTL positions which usually expands large chromosome regions that may contain several hundred genes, making their marker-assisted introgressions less efficient and involving many more genes apart from the one responsible for the trait.

In this opinion paper we propose that new tools for higher-precision breeding mediated by genetic engineering be created; we revisit where we stand with respect to the major obstacles, how we can overcome those obstacles, and what are still the major challenges for precise, quality genetically modified (GM) fruit crops.

*Correspondence to: Antonio Granell; Email: agranell@ibmcp.upv.es
Submitted: 01/03/10; Revised: 04/26/10; Accepted: 04/26/10
Previously published online:
www.landesbioscience.com/journals/gmcrops/article/12327

Table 1. Databases integrating genotypic and phenotypic data useful to finding candidate genomic regions involved in agronomic traits of interest.

Crop plant or family	World Wide Web address
Grasses	http://www.gramene.org/qtl/index.html
Grape	http://www.vitaceae.org
Tomato	http://164.107.85.47:8004/cgi-bin/qtl_information.pl http://zamir.sgn.cornell.edu/Qtl/Html/home.htm
Potato	http://www.scri.ac.uk/research/genetics/GeneticsAndBreeding/potatoes/mappingqtls
Cucurbitacea	www.icugi.org
Rosaceae	http://www.bioinfo.wsu.edu/gdr/
Various	http://www.phenome-networks.com/

Consumer-Oriented Fruit Traits and the Use of Intra- or Cisgenic Approaches as a Possible Means of Overcoming Public Concerns

Genetic engineering is now feasible for many crops, and the transfer of precise regions of the genome or precise genes or allelic forms from any organism into crop plants is a daily routine in many labs and companies. A number of examples supporting the successful use of transgenic approaches to increasing fruit quality have been described in the literature,[6-8] but have so far failed to make it to market.

This is due to many reasons, including the fact that GM crops lack general consumer acceptance.[9] GM crops still face fierce and increased opposition from consumers worldwide, and this is not expected to change in the next few years. The fact that no consumer-oriented traits were the basis for the first genetically engineered examples, that antibiotic resistance genes were introduced along with the traits and that genes from outside the traditional breeder's pool (i.e., from the same or related species) were used contribute to this lack of acceptance. A number of surveys support the view that naturalness is a key factor for the acceptance of GM crops, such that most people would eat food from genetically modified plants if the DNA introduced came from another variety of the same plant or a sexually compatible species and no foreign DNA remained after the biotechnological process[9] (see **Table 2**). The basis of the methodology for overcoming these concerns has been developed and is usually defined as cis- or intragenesis, and is designed to use genetic material from the same or related species, thus producing engineered plants that theoretically could have also been obtained by traditional breeding.[9,10]

As is the case for most traits of interest, the other main obstacle for GM crops with improved fruit quality is the difficulty in identifying the region and allelic form at the gene level in a species that is sexually compatible or related to that which we are interested in engineering.

The study of natural genetic variation will help overcome these difficulties and objections, as is reflected, for instance, in the nature of plant projects funded by the EU (http://ec.europa.eu/research/agriculture/index_en.html). The knowledge obtained from these studies will be very useful for the next generation of GM crops.

Understanding the Nature of the Gene Underlying a Trait of Interest

Knowing the nature of the gene underlying a trait of interest is not necessary for classical plant breeding but is crucial when a targeted, genetically engineered approach is proposed. The identification of such genes has been elusive until recently, but the integration of new genomics technologies with currently available biotechnological and genetic tools is helping to discover these genes. The identification of genes/gene variants underlying many of the traits of interest is still pending.

In a simple scenario, a gene with effects of interest may be identified in a model species, and transferred later to a cultivated species. This approach has had as yet limited success in the delivery of new varieties.[11] On the other hand, a new wave of engineered plants based on the cis- (www.cisgenesis.com/) or intragenic approaches[12] is emerging. The orthologue gene of that in the model species can be identified in the cultivated species and then engineered either directly from the same species or even ectopically or temporally expressed by an intragenic approach. As the resulting plant does not contain any DNA sequences from outside the traditional breeders' pools, this approach should receive more favorable public acceptance and hopefully less regulatory opposition.[11,13-15]

But often times, as is the case in fruit quality traits (aroma, flavor, nutrition content, fruit morphology, etc.), the study of model species cannot shed light on the genes involved in a given process because the quality trait is not present in the model species. The experiments must then be carried out in the target species, where the genomic tools and information about gene function are far fewer as compared with model species. When the precise nature of the gene is not known or when reliable candidate genes cannot be obtained, the first step usually involves the mapping of QTLs. The information revealed by QTL maps indicates that often several independent genome regions are responsible for the trait and contribute to the trait to varying degrees. Further, the precision of the QTL map is limited (depending on population size, heritability of the trait, number of QTLs, accuracy of the phenotyping, etc.). QTL intervals often span large genomic regions, many times covering 10–20 cM that may correspond to up to approximately 2,000 genes,[16] making it virtually impossible to identify the gene(s) in early experiments. Fine mapping can be used to narrow down a QTL more accurately, i.e., to a region of <1 cM containing a few tens of genes. Finally, the gene underlying the QTL may be cloned (i.e., *fw2.2*,[17] *brix9.2*;[18] reviewed in ref. 19). Currently, these strategies require the investment of large amounts of time and resources and are not always feasible. This means that fine mapping and elimination of linkage drag is still costly and sometimes impossible.

At this point the question is: what genes to engineer into a crop plant? We still do not know the sequences of most, but we do know that there is enough genetic variability to improve quality traits, currently defined just as QTLs. We know where

Table 2. Technologies with different levels of "engineering" or "foreigness"/"naturalness" and level of precision (linkage drag)

Technology	Gene or region to transfer	Introgressed region/linkage drag
Xenogenesis	Foreign including "artificial"	highly defined/no linkage
Transgenesis	Across species boundaries	highly defined/no linkage
Cisgenesis	Within species or related/breeder's pool (no further rearrangement)	highly defined/no linkage
Intragenesis	Different parts from the genome of same or s-related organism*	Highly defined/no linkage
Transplantomics	Xeno o transgenesis of the plastid genome	Highly defined/no linkage
Cisplantomics	Within species plastid genome	Highly defined/no linkage
MASPB	Region between markers from sexually compatible	Low to medium resolution/yes linkage
Classical PB	Region associated to trait from sexually compatible	Low resolution/yes linkage

*i.e., mutant allele transferred to another cultivar/breeder's pool.

to obtain some of these allelic differences (the genotypes from those QTLs where mapped), and so the next two questions are: how to identify the gene in question and how to engineer it into the plant?

Golden Era for Gene Discovery Associated with Quality

Although still in the early stages, we are currently experiencing a dramatic paradigm shift in how biological information is obtained due to the large amount of genomic information and tools available for a number of crop plants (for instance, Solanaceae, rice and Graminae, Rosaceae, etc.). The possibility of identifying genes associated with or responsible for a trait of interest is increasing dramatically with the increasing number of genomic tools that can simultaneously interrogate the participation of thousands of genes, markers or several hundred metabolites in a given phenotype or trait. This information is often centralized and integrated in one-stop shop repositories with tools that analyze and correlate large phenotypic and genotypic datasets (**Table 1**). At present, the main bottleneck is obtaining accurate phenotypes. A new field of Phenomics is even emerging as a powerful tool to understand the complex relationships between phenotype and genotype.

Further to this, the complete genomes for a number of important crops (rice, tomato, potato, grape, papaya, cucumber, maize) have been elucidated and many more are yet to come (many during 2010). The technological advances in next-generation sequencing and assembly protocols indicate that, in the case of Solanaceae, for instance, the sequencing of tomato and potato will be followed by the sequencing of several other Solanacea (including eggplant, pepper, etc.). In the SOL100 initiative alone, Dr. Sanwen in Beijing plans to sequence 100 genes related to tomato. The high levels of synteny, microsynteny and sequence homology that exist within Solanaceae crops indicate that there will be rapid progress in the assembly of those other species once the reference tomato and potato genomes are finished (first version was released in Dec. 2009 on the SGN website; http://potatogenomics.plantbiology.msu.edu/ftp://ftp.solgenomics.net/tomato_genome/wgs/assembly). Similar situations are envisioned for crops related to other plants currently being sequenced.

Most important for gene discovery directed towards fruit quality improvement is that there are already ongoing sequencing programs for accessions of *Solanum pennelli* (Langhorst K, personal communication), *S. pimpinellifolium* (Granell A, personal communication) and up to one hundred more (Sanwen, personal communication), with the peculiarity that some of them are parentals used in the construction of Recombinant Inbred Line or Introgression Line, populations used previously to identify a large number of QTLs involved in crop quality. Sequence information will allow the rapid development of new markers and map-based cloning without the need of tedious large genome library screenings and chromosome walking.

The number of cloned QTLs has been increasing over the last few years,[20] with eight cloned QTLs identified in crop species, and in just the past five years more than thirty QTLs have been elucidated. We expect the number of cloned QTLs to increase dramatically in the next few years. Furthermore, the availability of variability at the genome sequence level and the low cost of high throughput genotyping methods will also allow the study of the correlation between sequence polymorphism at specific genes and phenotypic variability by association analysis in natural populations or germplasm collections.[21] This will result in a dramatic increase in the pool of genes and alleles of interest for both traditional and biotechnological breeding in the next few years (**Fig. 1**).

A rapid assay will be necessary in order to verify the gene function of candidate genes. The technology is already available. For example, transient assays by Virus Induced Gene Silencing[22,23] enable the testing of a large number of candidate genes. An example of the success of the combination of natural variants and genomic tools to identify bona fide candidate genes was obtained with the transcriptional factor from tomato myb12 which confers the pink fruit phenotype.[24] Out of a huge number of Myb and bhlh transcriptional factors expressed in fruits, myb12 was the only one that showed a transcript profile similar to the profiles of accumulation flavonols and to the levels of expression of structural genes in the flavonol biosynthesis pathway in both wild type and the *y* mutant. A polymorphism detected in the myb12 gene in the y accession mapped to the same location, indicating that myb12 is likely the y mutation and thereby provides a breeding tool. Existing alleles of myb12 can now be searched for in the traditional breeder pool and transferred by crossing and

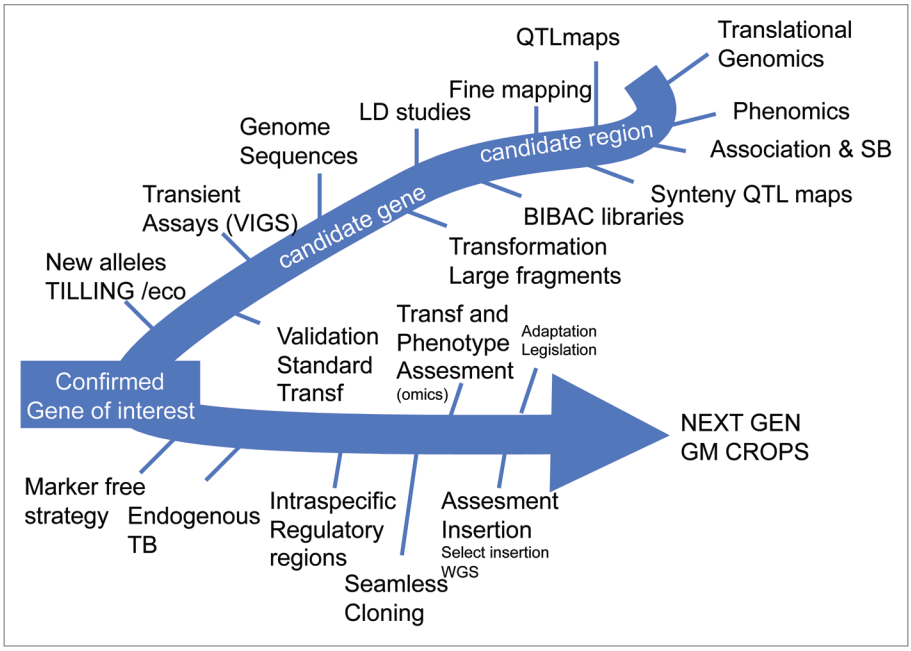

Figure 1. General scheme for the selection, confirmation and introduction of alleles from the breeder's gene pool into crop plants by seamless marker-free genetic engineering.

selection, gain-of-function mutations can be transferred by transformation to the elite background or an antisense construct may be made using all Solanum regulatory regions to provide an array of effects on flavonol levels.

These combined approaches are still emerging; they are expensive to implement and the expertise to use them is still dispersed among different labs. As the price of many of these technologies (next-gen sequencing, large-scale genotyping, etc.) keeps falling, the number of people using them will increase, as will the involvement of breeding companies in the large research consortiums where these technologies are integrated, and we will see more and more success cases.

Cisgenic and Intragenic Engineering by Design

In a more powerful approach, we propose that once a number trait to gene association has been untangled, natural variability be used as a source of alleles of interest for biotechnological breeding by target (design) introgression of new alleles using cisgenic/intragenic engineering.[13] In the simplest scenario, the gene allele conferring the best trait would come from the same species or a closely related sexually compatible species to be engineered into the elite background as is (cisgenic). A limitation is that this must be a dominant allele since, with genetic engineering, the allele of interest will be inserted at an unpredictable, though identifiable, location in the plant genome, while the endogenous allele still remains there (see future challenges section on how to improve this).

In a more elaborate process, ectopic expression of the gene for specific expression of a trait may be obtained by selecting the promoter sequence driving the desired pattern of expression and shuffled for the endogenous promoter or regulator (this has been dubbed intragenesis: selecting the pericarp promoter, anthocyanin regulator gene or having the expression of the gene of interest downregulated by antisense technology, for instance, still using a plant-only strategy).

A further elaboration of this approach can be proposed with targeted introgression by the genetic engineering of several genes (as achieved for resistance genes) from the breeder's pool and its pyramidization in the elite background without the undesired linkage drag effects associated with the breeding by design™ strategy[25] that was initially proposed[26] to transfer several regions of interest by MASPB.

One Step Further—GM Crops Become Synthetic Biology Oriented: Phytobricks, the Building Blocks for the Genetic Engineering of Fruit Crops

A way to envision the future of fruit crop design by means of genetic engineering is to follow the current trends of genetic engineering in organisms simpler than plants, and more precisely that of the emerging field of Synthetic Biology (SB).[27,28] Synthetic biologists construct interchangeable standardized components (building blocks or biobricks) which can be assembled together following certain assembly rules or "assembly standards."

Some laboratories are beginning to apply synthetic biology principles to plants. The concept of phytobricks has been proposed to refer to plant genetic building blocks (www.symbio.org.uk). These include gene coding sequences, such as enzymes and transcription factors mined from collections of newly sequenced plant genomes, gene expression databases and gene functional analysis, but also non-coding regulatory regions, including promoters,

matrix-attachment regions, enhancers, microRNAs, etc., as well as higher-order elements, such as regulatory networks, genetic modules or even genetic loci arising from systems biology and genomic projects.

In order to make synthetic crop design feasible, the increasing availability of phytobricks for fruit trait improvement needs to go hand in hand with the technical developments that facilitate the construction and assembly of interchangeable biological parts in plant binary vectors. The increasing affordability of gene synthesis, which has contributed to making even the synthesis of a complete genome possible,[29] as well as the development of new cloning methods (assembly standards, following the synthetic biology nomenclature) adapted to plant transformation vectors, are becoming widespread. In the most widely used assembly standard, the flanking regions are designed in such a way that the assembly reactions are idempotent, that is, the resulting assembled element contains identical flanking regions to the original components. This feature, which greatly facilitates the exchange of elements and the growth of increasingly complex structures, is achieved by the use of restriction enzymes with compatible overhangs.[30]

A common standard for the assembly of genetic elements in plant expression cassettes is yet to be defined, but the strategy is necessarily more complex than in microbial biotechnology. One of the reasons for this is that DNA constructs cannot be delivered as self-standing plasmids, as in bacteria. Instead, they need to be integrated within a T-DNA, flanked by T-DNA borders and located in a binary plasmid. In addition, phytobricks are on average bigger than bacterial components, making the discovery of unique restriction sites more difficult. The combination of all these factors makes the use of idempotent strategies very difficult in plants. Nevertheless, plant biotechnologists have almost unintentionally created a sort of rudimentary assembly standard for the cloning of genetic elements into plant vectors. Some examples are the use of rare-cutters for the exchange of expression cassettes between multicopy and binary vectors,[31-33] or the development of multipurpose vector collections for plant transformation.[34-36]

In recent years, the use of recombination-based cloning, and particularly the Gateway proprietary technology, has become widespread in plant biotechnology. This has boosted the generation of gateway-adapted multipurpose vector collections, made of destination binary vectors decorated with fixed phytobricks (usually a reporter gene or a promoter), which can easily incorporate a new phytobrick by homologous recombination (usually one's gene-of-interest).[37,38] A step forward towards the generation of truly exchangeable modular parts is provided by Multisite Gateway (MGW) technology. Based on homologous recombination, MGW allows tandem assembly of up to four basic "phytobricks" in a binary destination vector. The advantages of modular assembly provided by MGW technology have been recently demonstrated with the generation of fruit-dedicated collections of phytobricks, which include fruit promoters, reporters, tags and terminators.[39] This is probably the first example of a modular collection of interchangeable parts made in plants that resembles those efforts made in SB for prokaryotic organisms.

Also relevant to recombinant crop design is yet another recently developed cloning strategy based on Type IIS restriction enzymes, known as GoldenGate cloning.[40,41] In contrast to traditional Type II RE or Gateway, this strategy allows the assembly of several DNA bricks together without leaving residual foreign sequences in between. In traditional Type II cloning, the recognition sites for regular Type II REs cannot be removed and the assembly of DNA fragments leaves a "scar" or "seam" in the joining region between the original fragments. Gateway recombination is an even "dirtier" strategy in this sense, as it leaves long "scars" of 12 nucleotides between assembled bricks. This "foreign" DNA sequence may have several disadvantages, such as disturbed protein fusions, interference with transcriptional fusions or simply the introduction of exogenous DNA sequences that may ruin "clean" cloning strategies, such as those designed for recombinant drug production or cis/intragenic approaches. In contrast, Type IIS REs' main feature is that they cut a few nucleotides away from their recognition site, so that the digested stretch has no sequence requirements and can be custom-tailored. If the cut sequence (usually four nucleotides in length) in one brick is designed in such way that it matches the flanks of the next adjacent brick, this results in a seamless fusion. By carefully designing the flanking regions of each DNA brick, seamless multi-component constructs can be easily and efficiently built in binary vectors from interchangeable basic phytobricks.

Synthetic Biology principles might be applied well to fruit crops for multiple purposes. Perhaps the most obvious is to enhance metabolic engineering for biofortification or the bioproduction of industrial or pharmaceutical metabolites. In addition, in the field of fruit molecular farming, the production of multimeric proteins of increasing complexity may benefit from SB-like approaches. It is worth mentioning that the use of seamless assembly standards is particularly of interest in the field of recombinant plant breeding, since it may facilitate the design of intragenic constructs. The ability to fuse the intraspecific phytobricks without leaving exogenous scars may facilitate the development and acceptance of this technology. This SB approach is still emerging in fruit crop plants and although it is promising, especially for the purposes indicated above, it still faces many difficulties, such as how to coordinate expression of multiple genes as well as unpredictability regarding the stability of large insertions as well as the behavior of the inserted DNA in the context of the whole genome.

Present and Future Challenges

Even in the best scenario of an all-plant-DNA-engineered GM crop plant,[42,43] and even presupposing their acceptance by society and its legislators,[14] the generation of GM crops of interest for the different stakeholders still faces a number of technological challenges to its optimal implementation.

The first is identifying all genes associated with the traits of interest and finding the best alleles or patterns of expression that lead to a superior product. As indicated earlier, further progress in genomics and genome sequencing will assist in this area.

The second is overcoming the difficulties encountered with certain crops and especially with certain elite materials for

important crops that are still recalcitrant to Agrobacterium-mediated transformation. Here, using a cultivar amenable to transformation, and mobilizing the transformation event to the elite cultivar by backcrossing followed by marker-assisted selection using the same gene is one possibility, but understanding the mechanisms underlying the differences in susceptibility to transformation is an important problem to be solved in order to expand the technology to whatever crop or variety may be of interest.[44]

The third is to understand the molecular/genetic basis for the differences in stability, but also in the levels of expression achieved by different events of transformation (molecular/structural basis of position effects, chromatin structure, epigenetic effects, etc.). This could help overcome some technical and legislative difficulties associated with unstable traits (it is difficult to register something that changes with time,[45] and it would be beneficial for the optimization of trait expression).

The fourth, but likely most important challenge is to develop an efficient transformation system that would direct the insertion of the gene of interest to specific sites in the plant genome. Randomness and spurious integration of backbone sequences in the GM crop plants contribute to a negative view of transgenic plants. In one scenario, this could take advantage of homologous recombination as has been described with almost all other organisms from bacteria to animal cells. In all these other cases genomic engineering tools have been developed that permit directed and highly efficient modifications of a chosen genomic locus into virtually any desired mutant allele (reviewed in refs. 46–48). Currently, the transformation efficiency in some cases in plants is good, but the insertion site is unpredictable, and thus the frequency of homologous recombination is less than 1 in 100,000.[49-51] Further work in this direction is needed as homologous recombination in the case of cisgenesis would increase the precision of the transformation, but most importantly could overcome the presence of the endogenous allele that, in the case of homologous recombination, would be substituted by the engineered allele-introduced trait (this is even more important when the introduced allele is recessive). This is a badly needed technology for both the applied and basic understanding of genes in plants and where, despite our well-developed transformation technology, we are lagging behind animal systems. Limited success has been obtained more recently in that direction, such as recombinase-mediated cassette exchange (RMCE), which allows the integration of transgenes directionally into predefined plant genome sites and therefore gene replacement.[52,53] A fifth additional challenge has to do with the size of the DNA to be introduced in a GM crop. In the best scenario, this size would be compatible with transferring not just one but several genes, like in the case of multiprotein complexes for the modification of metabolic pathways, by selecting the best candidates, for instance, from different mutant backgrounds and collating them in a single stretch of DNA. The size of the engineered DNA should also be big enough as to be useful for testing several candidate genes in a QTL region in a time-efficient manner. This latter possibility is important as the resolution of most of the QTL maps spans over a hundred genes or a few hundred kbs.[16] Transfer of up to 150–200 kb DNA has been reported, although there are some questions about the stability of the transferred DNA.[54,55] Transferred DNA sizes of around 40 kb appear to be more stable.[56,57] Interestingly enough, this is approximately the size of the transferred DNA in wild Agrobacterium T-plasmid. In order to define and use the gene(s) underlying a QTL, the genome of the parental providing the allele of interest could be cloned in large fragments (higher than 40 kb) in a type of plasmid (BIBAC system) and then be used for plant transformation of the related sexually compatible crop cultivar. BIBACs in the QTL region can be identified by screening or by using HTP; the BIBAC library can be end-sequenced and a database produced to tag the BIBAC[58] or artificial chromosome[59,60] vector collection and identify those that map in the QTL region, provided a physical map is available or the genome of the plant is known. Once confirmed by transformation that the BIBAC contains the gene of interest, the region can be scanned at a higher resolution by targeting individual genes in independent transformations. The availability of complete genome sequences for the crop of interest would facilitate this approach and, together with genomic experiments, would additionally make the identification of operons of interest[61] easy.

Lastly, there is the challenge of satisfying the need to generate and release clean DNA-engineered events. Even though Agrobacterium-mediated transformation allows defining the precise piece of DNA to be engineered into the plant, thus avoiding the linkage drag normally associated with the introgression of DNA regions by MASPB, there are reports of unexpected and undesirable/spurious plasmid vector backbone introgression. Binary vector backbone sequences can be found in a high percentage of transgenic plants,[62] and this may increase with the size of the transferred DNA.[56] Although the transfer of vector backbone sequences to plant genome appears to be a natural consequence of the intrinsic mechanism of DNA transfer by Agrobacterium[62] and therefore unavoidable, a number of strategies have been developed to counter select the transfer of vector pieces by selection against markers incorporated in that region.[63] The availability of methods to rapidly screen different transgenic events and to select those with no vector introgression in combination with the use of positive selection strategies may help in the identification of clean events of transformation. The linkage drag in any case is much smaller than in the case of QTL introgression by MAS, although the non-plant nature of the former is to be considered a negative carryover (see **Table 2**). The good news is that they are easy to score and a strategy to select the modified plants with no vector sequences attached is feasible. The detection of viral remnants and even cryptic transfer DNA sequences in plant genomes also indicates that during evolution the transfer of DNA from other organisms has occurred.

Other challenges, partially satisfied with the current technologies but requiring additional developments include: the development of marker-free engineered plants,[11,64-67] to cope mainly with the concern of spreading antibiotic resistance genes used generally as selection markers, and the identification of plant sequences that can work as transfer T-DNA borders.[68]

References

1. Hammer K, Arrowsmith N, Gladis T. Agrobiodiversity with emphasis on plant genetic resources. Naturwissenschaften 2003; 90:241-50.
2. Tanksley SD, Nelson JC. Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor Appl Genet 1996; 92:191-203.
3. Hajjar R, Hodgkin T. The use of wild relatives in crop improvement: a survey of developments over the last 20 years. Euphytica 2007; 156:1-13.
4. McCouch S. Diversifying selection in plant breeding. PLoS Biol 2004; 2:347.
5. Gur A, Zamir D. Unused natural variation can lift yield barriers in plant breeding. PLoS Biol 2004; 2:245.
6. Qin Y, Teixeira da Silva JA, Zhang L, Zhang S. Transgenic strawberry: state of the art for improved traits. Biotechnol Adv 2008; 26:219-32.
7. Fraser PD, Enfissi EM, Bramley PM. Genetic engineering of carotenoid formation in tomato fruit and the potential application of systems and synthetic biology approaches. Arch Biochem Biophys 2009; 483:196-204.
8. Azari R, Tadmor Y, Meir A, Reuveni M, Evenor D, Nahon S, et al. Light signaling genes and their manipulation towards modulation of phytonutrient content in tomato fruits. Biotechnol Adv 2010; 28:108-18.
9. Rozin P. Naturalness judgments by lay Americans: Process dominates content in judgments of food or water acceptability and naturalness. Judgm Decis Mak 2006; 1:91-7.
10. Lusk JL, Rozan A. Consumer acceptance of ingenic foods. Biotech J 2006; 1:1433-4.
11. Rommens CM. Intragenic crop improvement: combining the benefits of traditional breeding and genetic engineering. J Agric Food Chem 2007; 55:4281-8.
12. Rommens CM, Haring MA, Swords K, Davies HV, Belknap WR. The intragenic approach as a new extension to traditional plant breeding. Trends Plant Sci 2007; 12:397-403.
13. Schouten HJ, Krens FA, Jacobsen E. Cisgenic plants are similar to traditionally bred plants: international regulations for genetically modified organisms should be altered to exempt cisgenesis. EMBO Rep 2006; 7:750-3.
14. Schouten HJ, Krens FA, Jacobsen E. Do cisgenic plants warrant less stringent oversight? Nat Biotechnol 2006; 24:753.
15. Jacobsen E, Schouten HJ. Cisgenesis strongly improves introgression breeding and induced translocation breeding of plants. Trends Biotechnol 2007; 25:219-23.
16. Price AH. Believe it or not, QTLs are accurate! Trends Plant Sci 2006; 11:213-6.
17. Frary A, Nesbitt TC, Grandillo S, Knaap E, Cong B, Liu J, et al. fw2.2: a quantitative trait locus key to the evolution of tomato fruit size. Science 2000; 289:85-8.
18. Fridman E, Pleban T, Zamir D. A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. Proc Natl Acad Sci USA 2000; 97:4718-23.
19. Doebley JF, Gaut BS, Smith BD. The molecular genetics of crop domestication. Cell 2006; 127:1309-21.
20. Salvi S, Tuberosa R. To clone or not to clone plant QTLs: present and future challenges. Trends Plant Sci 2005; 10:297-304.
21. Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, et al. Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell 2009; 21:2194-202.
22. Orzaez D, Medina A, Torre S, Fernandez-Moreno JP, Rambla JL, Fernandez-del-Carmen A, et al. A visual reporter system for virus-induced gene silencing in tomato fruit based on anthocyanin accumulation. Plant Physiol 2009; 150:1122-34.
23. Orzaez D, Granell A. Reverse genetics and transient gene expression in fleshy fruits: overcoming plant stable transformation. Plant Signal Behav 2009; 4:864-7.
24. Ballester AR, Molthoff J, de Vos R, Hekkert BL, Orzaez D, Fernandez-Moreno JP, et al. Biochemical and molecular analysis of pink tomatoes: deregulated expression of the gene encoding transcription factor SlMYB12 leads to pink tomato fruit color. Plant Physiol 2010; 152:71-84.
25. Collard BC, Mackill DJ. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philos Trans R Soc Lond B Biol Sci 2008; 363:557-72.
26. Peleman JD, van der Voort JR. Breeding by design. Trends Plant Sci 2003; 8:330-4.
27. Haseloff J, Ajioka J. Synthetic biology: history, challenges and prospects. J R Soc 2009; 6:389-91.
28. Benner SA, Sismour AM. Synthetic biology. Nat Rev Genet 2005; 6:533-43.
29. Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, Zaveri J, et al. Complete chemical synthesis, assembly and cloning of a Mycoplasma genitalium genome. Science 2008; 319:1215-20.
30. Knight TF. Idempotent vector design for standard assembly of biobricks. MIT Synthetic Biology Working Group Technical Reports 2003.
31. Chung SM, Frankman EL, Tzfira T. A versatile vector system for multiple gene expression in plants. Trends Plant Sci 2005; 10:357-61.
32. Dafny-Yelin M, Chung SM, Frankman EL, Tzfira T. pSAT RNA interference vectors: a modular series for multiple gene downregulation in plants. Plant Physiol 2007; 145:1272-81.
33. Dafny-Yelin M, Tzfira T. Delivery of multiple transgenes to plant cells. Plant Physiol 2007; 145:1118-28.
34. Hellens RP, Edwards EA, Leyland NR, Bean S, Mullineaux PM. pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. Plant Mol Biol 2000; 42:819-32.
35. Thole V, Worland B, Snape JW, Vain P. The pCLEAN dual binary vector system for Agrobacterium-mediated plant transformation. Plant Physiol 2007; 145:1211-9.
36. Tzfira T, Tian GW, Lacroix B, Vyas S, Li J, Leitner-Dagan Y, et al. pSAT vectors: a modular series of plasmids for autofluorescent protein tagging and expression of multiple genes in plants. Plant Mol Biol 2005; 57:503-16.
37. Curtis MD, Grossniklaus U. A gateway cloning vector set for high-throughput functional analysis of genes in planta. Plant Physiol 2003; 133:462-9.
38. Karimi M, Inze D, Depicker A. GATEWAY((TM)) vectors for Agrobacterium-mediated plant transformation. Trends Plant Sci 2002; 7:193-5.
39. Estornell LH, Orzaez D, Lopez-Pena L, Pineda B, Anton MT, Moreno V, et al. A multisite gateway-based toolkit for targeted gene expression and hairpin RNA silencing in tomato fruits. Plant Biotech J 2009; 7:298-309.
40. Engler C, Gruetzner R, Kandzia R, Marillonnet S. Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. PloS one 2009; 4:5553.
41. Engler C, Kandzia R, Marillonnet S. A one pot, one step, precision cloning method with high throughput capability. PloS One 2008; 3:3647.
42. Rommens CM, Humara JM, Ye J, Yan H, Richael C, Zhang L, et al. Crop improvement through modification of the plant's own genome. Plant Physiol 2004; 135:421-31.
43. Rommens CM. All-native DNA transformation: a new approach to plant genetic engineering. Trends Plant Sci 2004; 9:457-64.
44. Ke J, Khan R, Johnson T, Somers DA, Das A. High-efficiency gene transfer to recalcitrant plants by *Agrobacterium tumefaciens*. Plant Cell Rep 2001; 20:150-6.
45. Schmid A. Biotechnology, plant variety protection and changing property institutions in agriculture. N Cent J Agr Econ 1985; 7:9.
46. Huang J, Zhou W, Dong W, Watson AM, Hong Y. From the Cover: Directed, efficient and versatile modifications of the Drosophila genome by genomic engineering. Proc Natl Acad Sci USA 2009; 106:8284-9.
47. Sorrell DA, Kolb AF. Targeted modification of mammalian genomes. Biotechnol Adv 2005; 23:431-69.
48. Kolb AF, Coates CJ, Kaminski JM, Summers JB, Miller AD, Segal DJ. Site-directed genome modification: nucleic acid and protein modules for targeted integration and gene correction. Trends Biotechnol 2005; 23:399-406.
49. Kempin SA, Liljegren SJ, Block LM, Rounsley SD, Yanofsky MF, Lam E. Targeted disruption in Arabidopsis. Nature 1997; 389:802-3.
50. Risseeuw E, Franke-van Dijk ME, Hooykaas PJ. Gene targeting and instability of Agrobacterium T-DNA loci in the plant genome. Plant J 1997; 11:717-28.
51. Risseeuw E, Offringa R, Franke-van Dijk ME, Hooykaas PJ. Targeted recombination in plants using Agrobacterium coincides with additional rearrangements at the target locus. Plant J 1995; 7:109-19.
52. Nanto K, Yamada-Watanabe K, Ebinuma H. Agrobacterium-mediated RMCE approach for gene replacement. Plant Biotechnol J 2005; 3:203-14.
53. Li Z, Xing A, Moon BP, McCardell RP, Mills K, Falco SC. Site-specific integration of transgenes in soybean via recombinase-mediated DNA cassette exchange. Plant Physiol 2009; 151:1087-95.
54. Hamilton CM, Frary A, Xu Y, Tanksley SD, Zhang H-B. Construction of tomato genomic DNA libraries in a binary-BAC (BIBAC) vector. Plant J 1999; 18:223-9.
55. Miranda A, Janssen G, Hodges L, Peralta EG, Ream W. *Agrobacterium tumefaciens* transfers extremely long T-DNAs by a unidirectional mechanism. J Bacteriol 1992; 174:2288-97.
56. Frary A, Hamilton CM. Efficiency and stability of high molecular weight DNA transformation: an analysis in tomato. Transgenic Res 2001; 10:121-32.
57. Song J, Bradeen JM, Naess SK, Helgeson JP, Jiang J. BIBAC and TAC clones containing potato genomic DNA fragments larger than 100 kb are not stable in Agrobacterium. Theor Appl Genet 2003; 107:958-64.
58. Hamilton CM, Frary A, Lewis C, Tanksley SD. Stable transfer of intact high molecular weight DNA into plant chromosomes. Proc Natl Acad Sci USA 1996; 93:9975-9.
59. Liu YG, Shirano Y, Fukaki H, Yanai Y, Tasaka M, Tabata S, et al. Complementation of plant mutants with large genomic DNA fragments by a transformation-competent artificial chromosome vector accelerates positional cloning. Proc Natl Acad Sci USA 1999; 96:6535-40.
60. Liu YG, Nagaki K, Fujita M, Kawaura K, Uozumi M, Ogihara Y. Development of an efficient maintenance and screening system for large-insert genomic DNA libraries of hexaploid wheat in a transformation-competent artificial chromosome (TAC) vector. Plant J 2000; 23:687-95.
61. Barone A, Di Matteo A, Carputo D, Frusciante L. High-throughput genomics enhances tomato breeding efficiency. Curr Genomics 2009; 10:1-9.
62. Kononov ME, Bassuner B, Gelvin SB. Integration of T-DNA binary vector 'backbone' sequences into the tobacco genome: evidence for multiple complex patterns of integration. Plant J 1997; 11:945-57.
63. Hanson B, Engler D, Moy Y, Newman B, Ralston E, Gutterson N. A simple method to enrich an Agrobacterium-transformed population for plants containing only T-DNA sequences. Plant J 1999; 19:727-34.
64. Zuo J, Niu QW, Moller SG, Chua NH. Chemical-regulated, site-specific DNA excision in transgenic plants. Nat Biotechnol 2001; 19:157-61.

65. Miller M, Tagliani L, Wang N, Berka B, Bidney D, Zhao ZY. High efficiency transgene segregation in co-transformed maize plants using an *Agrobacterium tumefaciens* 2 T-DNA binary system. Transgenic Res 2002; 11:381-96.
66. Klaus SMJ, Huang FC, Golds TJ, Koop HU. Generation of marker-free plastid transformants using a transiently cointegrated selection gene. Nat Biotechnol 2004; 22:225-9.
67. Daniell H, Muthukumar B, Lee SB. Marker free transgenic plants: engineering the chloroplast genome without the use of antibiotic selection. Curr Genet 2001; 39:109-16.
68. Rommens CM, Bougri O, Yan H, Humara JM, Owen J, Swords K, et al. Plant-derived transfer DNAs. Plant Physiol 2005; 139:1338-49.