

Complementing Plant Breeding Programs with Biotechnology

Dr Tom Elmo Clemente,
University of Nebraska-Lincoln



COMPLEMENTING PLANT BREEDING PROGRAMS WITH BIOTECHNOLOGY

By Dr Tom Elmo Clemente, University of Nebraska-Lincoln

There are three components that impact crop harvests: the genetics of the seed sown, the environment in which the crop is grown, and the inputs employed during the growing period. For most species, the genetic component is addressed through breeding programs. Breeding programs require genetic variation within the available crop population from which parental crosses are made. Breeding programs that develop elite cultivars for commodities such as maize, soybean, wheat and sorghum, have wonderful inherent genetic diversity to utilise. However, many desired traits that breeders seek are not present within existing crop populations. It is here where the tools of biotechnology can complement plant breeding programs, by introducing novel pieces of genetic variation that can impart these favourable traits.

Genes and Proteins

Our research group resides in the George W. Beadle Center on the campus of the University of Nebraska-Lincoln. George Beadle was a geneticist who in 1958 received the Nobel Prize in physiology and medicine, along with Edward Tatum, for the 'one gene, one protein' concept.

Genes, which can be simply defined as a piece of DNA with coding information to make a protein, are first transcribed to messenger RNA (mRNA) in the nucleus of a cell. The mRNA transcript is subsequently translated to a protein in the cytosol of the cell. It is the cellular protein make-up, or so-called proteome, that ultimately gives rise to traits.

The genome sequence of soybean, for example, predicts that the species has approximately 53,000 protein encoding genes, while that of maize approximately 35,000 protein encoding genes. One of the first novel pieces of genetic variation introduced into these two crops was a gene from a bacterium that imparted tolerance to

the herbicidal activity of glyphosate. Herbicide tolerance in agriculture is a desirable trait given that weed management protects crop yield. Given the source of the gene was outside of the plant species, the term 'transgene' is used. This transgene has been incorporated into soybean and maize breeding programs for over 20 years, and it remains a very powerful weed management tool for soybean and maize producers globally.

Introducing Genetic Variation

There are two requirements for successful introduction of novel genetic variation into a plant cell. The first is that a cell must be able to receive a piece of DNA, and the second is that same cell must have the capacity to become a plant that can pass the novel piece of DNA to the next generation. There are two general processes by which plant scientists deliver the novel piece of genetic variation: *Agrobacterium*-mediated delivery, and direct DNA introduction.

Agrobacterium-mediated delivery is carried out by exploiting the natural

transformation capacity of the bacterium *Agrobacterium tumefaciens*. This organism causes a plant disease called 'crown gall'. The elucidation of the genetic mechanism the microbe uses to trigger the disease began in the late 1950s and the primary mechanism was identified in the late 1970s. This herculean research task was carried out primarily in several public sector research groups led by Milt Gordon, Eugene Nester, Jeff Schell, Marc Van Montagu and Mary-Dell Chilton. Dr Chilton, in my opinion, was the one that pulled it all together; she initially worked with Nester and Gordon, then moved to the Washington University in St. Louis, and then to Syngenta



Shirley Jane Sato and Tom Elmo Clemente. The two have been coordinating research activities within the Plant Transformation Core Research Facility for over 25 years at the University of Nebraska-Lincoln.



Company. The outcome of these collective activities revealed that this bacterium evolved the capacity to transfer a piece of DNA (T-DNA) from its own cell into a plant cell during the host-microbe interaction.

The process starts when this soil microbe is in the vicinity of a wounded plant; the bacterium can sense wound metabolites and proceeds to migrate towards the wound site. In the initial steps of the interaction, a series of genes are expressed. These genes serve two roles: establishing a conduit to the plant cell, known as the *virB* pore, and processing the T-DNA. The T-DNA is quickly coated with binding proteins that protect it and transport it through the *virB* pore conduit into the plant cell, ultimately integrating it within one of the host's chromosomes.

The second process by which plant scientists introduce novel bits of genetic variation into plants is direct DNA delivery, which is primarily carried out via 'microprojectile bombardment'. Like the *Agrobacterium* research outcomes, this process was developed through public sector research activities led by a scientist from Cornell University, Dr John Sanford. This process involves precipitating pieces of DNA onto microprojectile spheres composed gold or tungsten, which range in size between 0.5 and 2.5 micrometres. These DNA-coated beads are then loaded onto a film that is situated above the plant cells within a biolistic device. A vacuum is generated and the film carrying the DNA-coated beads is accelerated towards a stopping screen. Upon hitting the screen, the DNA-coated beads are delivered into the plant cell, analogous to buckshot from a shot gun.

Once the plant cells have received the DNA, either through *Agrobacterium*-mediated transformation or the direct DNA method, the cells are placed in tissue culture to initiate the regeneration process. Depending on the media formulation

and culture conditions, the plant cells can follow one of two routes to transform into a whole plant: somatic embryogenesis or organogenesis. The former mirrors that of pollen fertilisation of an egg cell, while the latter will result in the differentiation of the cells into an organ, first a shoot, and the shoot subsequently induced to form roots.

Given the extremely low percentage of cells that have this dual ability to receive a piece of DNA and regenerate into a whole plant, the culturing environment must provide a selective advantage for these cells. For this reason, researchers will stack genes of interest with selectable marker genes. These genes impart the ability to grow on antibiotic or herbicide supplemented medium. Time frame from the delivery of DNA to the generation of a whole plant ranges from three to nine months, depending on the species.

Genome Editing Reagents

Plant scientists have the capacity to introduce genetic variation by the addition of a transgene, or edit a specific part of a chromosome (a locus), which can lead to either a reduction of expression of a gene or the generation of a 'null mutation', which will block the expression of a gene. These genome edits are another source of genetic variation a plant breeder can draw upon.

The ability to edit a specific genetic locus in a plant has significantly evolved over the past 15 years, the outcome being more precision. The first set of editing reagents relied upon specific proteins that can chaperone a fused molecular scissor to a genomic location. One such reagent is referred to as zinc finger nucleases. A zinc finger is a protein motif that can bind specific regions of DNA. The binding is influenced by amino acids present in the motif, and the stability of the binding



depends on a zinc ion. A motif is composed of an array of four amino acids. Each motif, depending on the array design, can recognise a specific set of three DNA base pairs. A zinc finger motif can then be assembled in a modular fashion, up to six, which will permit the recognition of a locus of 18 base pairs. The zinc finger module design is fused to a DNA scissor called *FokI*, which was isolated from a bacterium. This fused complex is referred to as a zinc finger nuclease (ZFN). Two such ZFN reagents are delivered that specifically bind the top and bottom of the target site, which then brings the *FokI* segments together and leads to a DNA cut (mutation).

A similar genome editing strategy was developed from knowledge about a mechanism that influences a host-microbe interaction. Many pathogens, both plant and animal, inject effector proteins into plant cells. These effectors interact with host factors, ultimately perturbing the plant's immune response in favour of pathogenicity. One class of effectors are called 'transcription activator-like' (TAL) due to their capacity to specifically bind to gene promoter regions, and activate gene transcription, which leads to mis-expression of these genes in such a way that facilitates pathogen ingress. TALs are protein modules that consist of 34 amino acid residues, varying only in the amino acids that reside at positions 12 and 13 of the module, which is called the repeat variable di-residue (RVD). The RVD dictates what part of the DNA the module will bind. Hence, one can assemble an array of modules, with altering RVDs, to bind to a specific stretch of DNA. For example, an 18-module array can be designed to preferentially bind a section of 18 base pairs.

Once the rules of TAL binding were determined, researchers used the *FokI* fusion strategy from the ZFN editing reagent approach to develop the next genome editing tool: transcription activator-like effector nucleases (TALENs). The advantage of this approach over ZFN is more flexibility in the genome targets that can be edited.

The third generation of genome editing tools is coined 'clustered regularly interspaced short palindromic repeats', or CRISPR. This tool was developed from what researchers learned from bacterial genome sequences. In the late 1980s when performing genome sequencing of bacterial species,

researchers observed an odd repetitive DNA sequence. By the early 2000s, it was realised that these repetitive DNA arrays contain segments of DNA from bacterial viruses called phages. Researchers also found repetitive clusters of DNA residing down-stream of a set of protein genes, which they termed CRISPR-associated genes (Cas).

In subsequent years, multiple research groups observed a correlation between the presence of phage DNA and the inability of that phage to subsequently infect the bacterium. This led the community to hypothesise that the Cas genes and down-stream CRISPR arrays are a type of bacterial immune system to protect against phage infections. The subsequent elucidation of this immune mechanism led to the development of CRISPR-Cas genome editing tools.

CRISPR-Cas editing reagents require a few key components: a Cas gene, and a guide RNA fused with a scaffold sequence. The scaffold sequence is what the Cas protein binds to, while the guide RNA, as the name implies, will bring the Cas protein to the genomic location targeted for editing. Regions of the genome that can be edited are dictated by a sequence referred to as the protospacer adjacent motif, which varies depending on Cas protein utilised. For example, one of the most widely used Cas proteins is Cas9, derived from the bacterial species *Streptococcus pyogenes*, which can edit a locus where the protospacer adjacent motif is NGG, where N is any base. The guide attaches to the DNA sequence adjacent to this motif, where Cas9 will then perform an edit. Compared to the ZFN and TALEN tools, CRISPR-Cas tool use RNA to guide the molecular scissor, which gives more precision, with significantly more editable loci.

Regardless of what editing tool a plant scientist decides to use to create a targeted mutation in a genome, the reagent design must be introduced into a host cell. For this, most approaches utilise *Agrobacterium*-mediated delivery or direct DNA introduction.

Bottlenecks in Getting Novel Genetic Variation on the Market

Most traits currently on the market that were developed through the tools of biotechnology are found in the major agriculture commodities: maize, soybean and cotton. There are a few exceptions to this, for example the transgenes present in papaya and yellow squash that impart virus resistance. It is noteworthy that papaya production on the Big Island of Hawaii would not be possible without the transgene that imparts resistance to papaya ringspot virus. This has been such a resounding success story, that when the state's legislators attempted to impose restrictions on transgenic crops, papaya was exempt.

So, how can it be that these tools of biotechnology, which have been continuously improved over the past 40 years, have led to



relatively few valuable traits on the marketplace? The answer lies in the global governmental regulatory agencies that govern the movement, release and market the availability of these traits. The global regulatory agencies are not harmonised, with each country having varying sets of regulatory hurdles that must be met to obtain de-regulated status. Secondly, and more importantly, the global regulatory agencies regulate the process by which genetic variation is introduced, not the trait derived from the novel genes.

This latter point is exemplified by an output trait – a trait that directly benefits the consumer, high oleic acid soybean oil. Oil biosynthesis during soybean seed development involves a fatty acid enzyme, Fad2-1, which converts the monounsaturated fatty acid, oleic acid, into the double saturated fatty acid, linoleic acid. If genetic variation introduced into soybean leads to the blocking of Fad2-1 activity, the outcome is a soybean oil high in oleic acid (>75%) and low in polyunsaturated fatty acids. Essentially, you get an oil with similar functionality to olive oil.

Genetic variation that leads to the perturbation of Fad2-1 enzymatic activity has been introduced into soybean via three processes. The first uses a chemical mutagen, ethyl methanesulfonate. When cells are exposed to this chemical, random mutations occur in the genome, and the scientist then needs to mine the plethora of mutations to find a genetic variant of interest. The second process involves the use of the genome editing reagent TALENs, while the third employs the addition of a transgene designed to down-regulate the expression of the Fad2-1 gene during seed development.

Regarding global regulation scrutiny, ethyl methanesulfonate mutagenesis is a process that has minimal regulatory hurdles across the globe. Genome editing reagents, such as TALENs, are ever changing, and regulatory hurdles are very different from country to country. Meanwhile, for the process by which a gene is added, major regulatory barriers remain.

The regulatory landscape that must be navigated to obtain a desirable trait in a crop is a major factor that is taken under consideration during the research and development stages of a product. The cost associated with securing de-regulation status of a novel gene introduced through the tools of biotechnology

can easily exceed \$20 million USD. In the high oleic acid soybean example, you have the same result, a soybean oil with greater than 75% oleic acid, but achieved through different processes.

Hence, the global regulatory agencies should follow the science, wherein a strong consensus of scientists agree that the process by which we add genetic variation to a plant is safe. Instead, what is needed is regulation of the trait, not the process approach to de-regulation of traits imparted by genes introduced into plants via the tools of biotechnology.

A second issue within most global regulatory agencies is that repeated introduction of a novel gene requires regulatory hurdles to be addressed, regardless of whether the gene has been de-regulated at an earlier date. For example, as mentioned above, the first herbicide tolerance trait imparted tolerance towards glyphosate. This trait was globally de-regulated in the late 1990s and remains a very powerful weed management tool in agriculture. Now if one were to take this same transgene and re-introduce it into soybean via biotechnology, regulatory hurdles would need to be overcome in many countries, while some countries are reassessing this approach. In the USA, the regulatory governance was recently updated, and such a re-introduction of the glyphosate tolerance gene would be expedited. However, we are in a global marketplace, and a harmonised regulatory process would greatly enhance our ability to get valuable traits on the market, and importantly, into other crops beyond the major commodities.

The challenge that lies ahead is addressing the caloric needs of an increasing population, with reduced arable land for crop production. The path forward will require scientists to continue to ask questions, in the constant pursuit of knowledge, addressing both management and genetic strategies that work together in an ever-changing environment, towards ensuring a plentiful, safe and sustainable food supply for future generations. To do this will require access to all available tools, including biotechnology, to add that novel bit of genetic variation that plant breeders require to do their job.

DR TOM ELMO CLEMENTE

Eugene W. Price Distinguished Professor of Biotechnology
Department of Agronomy & Horticulture
The Beadle Center for Genetics and Biomaterials Research
University of Nebraska-Lincoln
Lincoln, NE
USA

CONTACT

E: tclemente1@unl.edu

W: <https://www.unl.edu/clementelab/index.html>