REVIEW & INTERPRETATION

Bringing New Plant Varieties to Market: Plant Breeding and Selection Practices Advance Beneficial Characteristics while Minimizing Unintended Changes

Kevin C. Glenn,* Ben Alsop, Erin Bell, Mike Goley, Jonathan Jenkinson, Bing Liu, Cheryl Martin, Wayne Parrott, Chris Souder, Oscar Sparks, William Urquhart, Jason M. Ward, and John L. Vicini

ABSTRACT

Commercial-scale plant breeding is a complex process in which new crop varieties are continuously being developed to improve yield and agronomic performance over current varieties. A wide array of naturally occurring genetic changes are sources of new characteristics available to plant breeders. During conventional plant breeding, genetic material is exchanged that has the potential to beneficially or adversely affect plant characteristics. For this reason, commercial-scale breeders have implemented extensive plant selection practices to identify the top-performing candidates with the desired characteristics while minimizing the advancement of unintended changes. Selection practices in maize (*Zea mays* L.) breeding involve phenotypic assessments of thousands of candidate lines throughout hundreds of different environmental conditions over many years. Desirable characteristics can also be introduced through genetic modification. For genetically modified (GM) crops, molecular analysis is used to select transformed plants with a single copy of an intact DNA insert and without disruption of endogenous genes. All the while, GM crops go through the same extensive phenotypic characterization as conventionally bred crops. Data from both conventional and GM maize breeding programs are presented to show the similarities between these two processes.

K.C. Glenn, E. Bell, M. Goley, J. Jenkinson, B. Liu, C. Martin, C. Souder, O. Sparks, W. Urquhart, and J.L. Vicini, Monsanto Company, 800 North Lindbergh Blvd., St. Louis, MO 63167; B. Alsop, Benson Hill Biosystems, 1100 Corporate Square Dr., St., Louis, MO 63132; W. Parrott, Dep. of Crop and Soil Sciences, Univ. of Georgia, Athens, GA 30602; J.M. Ward, Royal Canin USA, 500 Fountain Lakes Blvd., Suite 100, St. Charles, MO 63301. Received 26 Mar. 2017. Accepted 26 June 2017. *Corresponding author (kcglen@monsanto.com). Assigned to Associate Editor Candice Hirsch.

Abbreviations: GM, genetically modified; PCR, polymerase chain reaction; QTL, quantitative trait locus; RM, relative maturity.

Farmers and consumers have an exceptional choice of crop varieties in the marketplace, with this abundance typically being unquestioned. The advent of genetically engineered crops brought the concept of genetic modification into public discourse and prompted a greater interest in how our food is produced. For some, there has arisen a misunderstanding on the origin of crop varieties and the different methods, including the original steps needed for domestication, by which humans have modified the genetics of our food sources (plants and animals). This misunderstanding has been reinforced through the media and, importantly, has affected government policies around the world.

Plant varieties can always be improved. They can yield more, better resist pests and diseases, survive shipping better, or simply taste better. Agricultural productivity in the United States increased 50% between 1982 and 2007 while using less land and labor (O'Donoghue et al., 2011). Productivity per unit input has increased 250% since 1948 (USDA-NASS, 2017). Plant breeding and technological advances in production practices contribute equally to these increases. Today's modern crop varieties were all derived from plant breeding. Plant breeding is an ongoing, cyclical process that involves identifying plants with desirable

Published in Crop Sci. 57:2906–2921 (2017).
doi: 10.2135/cropsci2017.03.0199

characteristics (yield, quality, resistance to abiotic and biotic stresses, etc.) and devising strategies to combine these characteristics to obtain superior varieties (Acquaah, 2012). In its simplest form, plant breeding results in improved crop varieties (the commercial product, also referred to as cultivars or hybrids, depending on the crop) through the mating of two or more parental lines that contain desirable characteristics. The target characteristics are measured over multiple generations throughout different environments and stress conditions. Offspring with desirable characteristics are selected, whereas offspring with undesired characteristics are eliminated from further breeding. The degree of improvement in the new variety depends on the level of genetic variation affecting the characteristics of interest and the ability to accurately measure the expression of these characteristics in many different environmental conditions (Fehr et al., 1987).

Genetic engineering, commonly referred to as genetic modification, is an additional tool that affords plant breeders new sources of characteristics, such as genes that confer abiotic or biotic stress tolerance, with many of these genes not available in the crop's genome (Prigge and Melchinger, 2012; Weber et al., 2012; Prado et al., 2014; Schnell et al., 2015). After a genetically modified (GM) line containing the desired DNA insert is chosen, the DNA insert is introduced (via backcrossing) into well-characterized, conventionally bred elite varieties. The selection process that follows is essentially the same as is used for conventionally bred crops.

Using hybrid maize (*Zea mays* L.) as the example, data are presented from Monsanto case studies to illustrate the commercial-scale breeding practices used to supply seed to farmers. The range of sources of genetic variation, extent of testing, and scope of plant selection processes used for conventional breeding are presented first, followed by a parallel overview for GM varieties.

CONVENTIONAL PLANT BREEDING

Sources of Genetic Variation Used for Conventional Breeding Programs

Plant breeders improve crops by identifying sources of genetic variation for the characteristics of interest. Plant genomes (the genetic material in each species) are highly variable, even within and among closely related species (Weber et al., 2012). Table 1 shows various natural biological processes that create genetic diversity. These include the movement of transposable elements, vertical gene flow via mating with wild relatives, horizontal gene flow (Bock, 2010) from unrelated plants, *Agrobacterium,* florendoviruses, pararetroviruses, and mutations such as single-nucleotide polymorphisms, chromosomal rearrangements, and the presence, absence, or copy number of germline genes.

Maize, in particular, has a high level of sequence and structural diversity (Buckler et al., 2006; Springer et al., 2009; Lai et al., 2010). A genomic comparison of two maize inbreds, B73 and Mo17, revealed an unprecedented level of genomic structural diversity compared with most higher eukaryotes studied thus far (Springer et al., 2009). For example, by a conservative estimate, at least 180 putative single-copy genes were present in one inbred but absent in the other, and >400 instances of putative sequence copy number variation between the two inbreds were observed (Springer et al., 2009). Likewise, a comparison (Hirsch et al., 2016) between B73 and PH207 found >2500 genes present only in one of those inbreds.

The traditional perspective has been that conventional breeding does not introduce new genes, only variations (alleles) of already existing genes. However, the emergence of the pangenome concept (Golicz et al., 2016) makes it clear that conventional breeding results in the introduction of additional genes and alleles, as well as novel combinations of genes. It is evident that the same mechanisms of genome instability found in nature that lead to genetic diversity (Table 1) are also active during conventional breeding (Weber et al., 2012; Schnell et al., 2015). One example is from a recent comparison of DNA structure (both large chromosomal changes and single-nucleotide polymorphisms) across a collection of soybean [*Glycine max* (L.) Merr.] cultivars, many derived by conventional breeding (Anderson et al., 2016). This study showed that genetic changes accumulated spontaneously across many conventional germplasm (i.e., standing variation). Another recent study of many maize varieties (both conventional and GM) showed that most of the observed compositional differences were associated with the backcrossing practices from conventional breeding (Venkatesh et al., 2015). Repetitive DNA sequences and structural variations in plants have the potential to contribute to genetic change. Similarly, transposable genetic elements in maize and many other plant species can mediate genetic changes (Hirsch and Springer, 2017). Transposable elements are DNA sequences that can change position within a genome, resulting in small insertions and deletions, as well as larger rearrangements such as inversions, deletions, and duplication of genes (Zhang and Peterson, 2004; Zhang et al., 2006; Weber et al., 2012).

Horizontal gene transfer across phylogenetic boundaries is another natural process that results in genetic variation in plants (Bock, 2010; Soucy et al., 2015), including transfer of DNA from bacteria, viruses, and unrelated plants (Bergthorsson et al., 2003; Staginnus et al., 2007; Liu et al., 2012; El Baidouri et al., 2014; Geering et al., 2014; Kyndt et al., 2015). One recent example was a study demonstrating that a wide selection of sweet potato [*Ipomoea batatas* (L.) Lam.] varieties contain *Agrobacterium* transfer DNA with expressed genes (Kyndt et

Table 1. Examples of naturally occurring genetic changes common in plants and the resulting characteristic.

Genetic change	Genotypic or phenotypic example	Reference
Transposable elements (transposons)	White grapes, blood oranges	Lisch (2013)
	>25,000 unique insertions detected across 31 varieties of soybean	Tian et al. (2012)
	Yellow maize	Palaisa et al. (2003)
	>50 new inserts of a transposon per rice plant per generation	Naito et al. (2006)
	Elongated tomato fruit	Xiao et al. (2008)
	Round or wrinkled peas (Mendel)	Ellis et al. (2011)
	2 million transposons exchanged between higher plants	El Baidouri et al. (2014)
Organellar DNA in nuclear DNA	Gain and loss of mitochondrial DNA common to maize inbred lines	Lough et al. (2008)
	Gain and loss of chloroplast DNA common to maize inbred lines	Roark et al. (2010)
Bacterial genes	Expression of several bacterial genes in sweet potatoes	Kyndt et al. (2015)
Crossing with wild relatives	>60 wild relatives have been used for >100 characteristics (80% involve pest or disease resistance) in 13 crops	Hajjar and Hodgkin (2007)
	Dozens of alien genes used in wheat breeding	Jones et al. (1995)
Pararetroviruses	Stable viral DNA in rice genome	Liu et al. (2012)
	Stable viral DNA in tomato (previously also seen in potato)	Staginnus et al. (2007)
Florendoviruses	Stable integrations in all plants	Geering et al. (2014)
Insertions and deletions	Submergence-tolerant rice	Xu et al. (2006)
	Dwarf sorghum	Multani et al. (2003)
	Yellow soybean seeds	Tuteja et al. (2004)
Single-nucleotide polymorphisms (SNPs)	Maize proteins (300–400 amino acids long) from 2 alleles differ by 3–4 amino acids	Tenaillon et al. (2001)
	Maize genome has 55 million SNPs	Gore et al. (2009)
	Green Revolution gene has 2 SNPs for dwarf wheat	Peng et al. (1999)
	One SNP caused loss of shattering in domestic rice	Konishi et al. (2006)
	Tall or short pea plants (Mendel)	Ellis et al. (2011)
	7 new SNPs created per meiosis per billion base pairs	Ossowski et al. (2010)
Presence, absence, or copy number of genes	856 wild-type soybean genes absent in cultivated varieties (and >186,000 DNA insertions or deletions)	Lam et al. (2010)
	>10^6 SNPs, 30,000 insertion or deletions, and a few large chromosomal deletions (>18 genes) in 6 elite maize varieties	Lai et al. (2010)
	Copy number variation relates to soybean cyst nematode resistance	Cook et al. (2012)
	Pinot Noir, Corvina, and Tannat wine grapes have 1873 genes not found in other wine grapes	Da Silva et al. (2013)
	Only 81% of *Brassica* genes are always present in the same number	Golicz et al. (2016)
	2500 genes found only in either B73 or PH207	Hirsch et al. (2016)
	G. soja genotypes can vary by 1000 to 3000 gene families from each other	Li et al. (2014)

al., 2015) as a consequence of natural transformation of an ancestral form of the plant.

Breeders commonly use locally adapted, domesticated germplasm (often called landraces) that exhibit exceptional performance in a specific group of geographic or management conditions, as well as international germplasm that are adapted to and have been selected for a wide range of environmental conditions (Hallauer et al., 1988; Lynch and Walsh, 1998; Doebley, 2004; Acquaah, 2012; Butruille et al., 2015). Locally adapted germplasm contain agronomic characteristics required for high yields in their environment and meet regional consumer preferences; therefore, they may be used more easily within a breeding program for their geographically distinct region.

Seed developers often have germplasm collections for crops that include most of the major cultivation regions of the world (Butruille et al., 2015; Smýkal et al., 2015). This germplasm is adapted for breeding in its region of origin and can also be used as a source of novel characteristics for breeding in other world regions. However, differences in agronomic characteristics, such as photoperiod sensitivity, temperature response, and disease susceptibility, must be recognized when using germplasm from other parts of the world.

To gain additional diversity not present in the existing germplasm, random genetic mutations can be induced using chemical or radiation mutagenesis (IOM/NRC, 2004; Curry, 2016). Breeders then select for agriculturally desirable genetic changes while selecting against the many unintended or unwanted changes that can occur with these methods (Bolon et al., 2011). Over 3000 plant varieties, mostly vegetables, fruits, grains, and ornamentals (IAEA-MVD, 2017), have been developed via mutagenesis. Well-known examples include the Star Ruby grapefruit (*Citrus × paradisi* Macfad) (Hensz, 1971) and high-oleic canola (*Brassica napus* L.) oil (Auld et al., 1992). Nevertheless, breeders are shifting from using these random methods of introducing genetic diversity over the past couple of decades to newer, more predictable methods like genetic engineering. Gene editing is emerging as a new tool for plant breeders to introduce alterations in genes to achieve desired plant characteristics; however, due to the scope of this review, gene editing is not discussed in detail here.

Interaction of Genotype and Environment on Plant Characteristics

A plant variety's appearance and performance (phenotype, *P*) is determined by an interaction between its genes (genotype, *G*) and the environment (*E*), commonly expressed as $P = G \times E$ (Lynch and Walsh, 1998). Traditionally, a major task of the plant breeder has been to differentiate between the effects of environment and genotype. The experimental design and selection strategy that breeders use to identify the most desirable genetic material is determined by the heritability, environment, and correlations between characteristics. Qualitative characteristics result in a limited number of possible phenotypes that can be placed into defined categories. For example, in simple Mendelian inheritance, only white or purple flower color, long or short stems, and yellow or green pod color are possible in pea (*Pisum sativum* L.) and are the result of a single gene controlling each characteristic (Ellis et al., 2011). Qualitative characteristics are more reliably expressed across different environments and are said to have high heritability. By comparison, quantitative characteristics are expressed as a continuous variation of phenotypes, in which few to multiple genes and the environment contribute to the expression of the characteristic. Examples of complex, quantitative characteristics in maize include yield and plant height, which are more strongly influenced by the environment than simpler qualitative characteristics (Hallauer et al., 1988).

Unintended Effects in Conventional Plant Breeding

Genomes are dynamic, and changes have occurred throughout evolution and during breeding that have the potential to affect plant characteristics. However, unforeseen hazardous effects (e.g., the production of an allergen, toxin, or other hazard that was not previously known to exist for the plant species) have not been documented (Parrott et al., 2012; Steiner et al., 2013; Ladics et al., 2015a, 2015b). Many crops have at least one characteristic known to be a hazard in some scenarios (IOM/NRC, 2004). For instance, potato (*Solanum tuberosum* L.) contains glycoalkaloids, such as solanine, and legumes contain lectins that are toxic when consumed in high doses or uncooked. Changes in the amounts of solanine in potato have occurred during breeding (Smith, 2013). Known anti-nutrients and toxins like these are typically monitored during conventional breeding to ensure that their levels are not increased (Zywicki et al., 2005; Shepherd et al., 2015). Increases in known toxicants, while unintended, are foreseeable and thus manageable in a breeding program.

Segregation of undesirable alleles already present in the breeding population also leads to unintended or undesirable effects. An example of a potentially undesirable agronomic characteristic is the degree to which the husk covers the ear in maize, which affects the protection against moisture and disease provided by the husk and in turn affects the geographic range of adaptation for maize hybrids. Maize grown in the southeastern United States require good husk coverage to prevent ear rot, whereas maize grown in northern latitudes benefit from reduced husk coverage to allow for faster ear dry down. Therefore, breeders developing short-season varieties (northern germplasm) from longer-season varieties (southern germplasm) must actively select against long, tight husks.

Other types of unintended effects from plant breeding may only be observed in certain environmental conditions. For example, a condition known as virescence (a temporary light-green color observed in maize seedlings) is triggered by cold and can result in delays in maturity (Fig. 1) (Hopkins and Walden, 1977). This phenotype is commonly observed in maize varieties that originated in warm climates, like those found in parts of South America. Plant breeding programs using South American maize varieties may contain this unintended virescence characteristic in their new varieties unless the allele responsible for this

Fig. 1. Unintended virescence phenotype observed during the development of new conventional maize varieties. Virescent plants have a chlorophyll deficiency and show a white to yellow phenotype. The virescent phenotype is induced by cool temperatures at germination and therefore was only observed in some field testing locations. The plants may appear normal as temperatures increase and the plants develop; however, the plants will likely have delayed maturity.

phenotype is selected against. Table 2 shows how adjusting the size of the breeding population and the intensity of selection can help to maintain genetic gain when a simply inherited characteristic, such as virescence, must be selected against in a breeding program. Typically, when an undesirable phenotype is present in the population at a high frequency, a plant breeding program will scale up the size of the beginning population to improve the chances of finding superior candidates with the desired characteristics that also lack the undesirable phenotype. It is important to note that unintended phenotypic changes, for the most part, do not equate with a food or feed safety risk.

BREEDING AND SELECTION OF INBRED LINES AND HYBRIDS

Breeding hybrid crops involves the creation of both the inbred lines, which may have improved characteristics important for hybrid seed production, and hybrids, which are the final commercial products that contain the combined characteristics from crossing two inbred lines. Hybrids usually show dramatic improvements over the inbred lines due to heterosis (Hallauer et al., 1988; Doebley, 2004; Schnable et al., 2009; Springer et al., 2009; Acquaah, 2012). Some crops, such as maize, have well-identified heterotic pools, making it easier to know which inbred line combinations tend to give superior hybrids. In other hybrid crops, such as tomato (*Solanum lycopersicum* L.) and cucumber (*Cucumis sativus* L.), there is less understanding of heterotic pools, and thus less predictability of which inbred line combinations will produce superior hybrids.

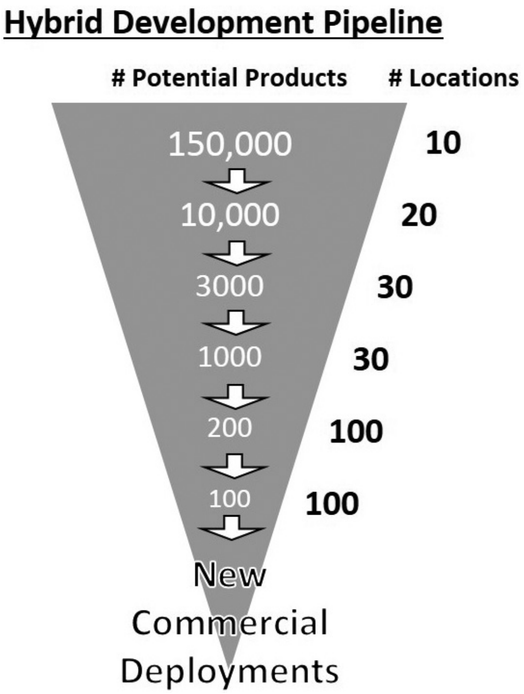


Fig. 2. Conventional maize breeding development process. The general process of developing new maize hybrids is shown and is the same process used in the selection of inbred lines. The number of individual plants tested is noted inside the triangle. Candidate lines are tested as both inbreds and hybrids in field trials. After each growing season, the best inbreds and hybrids are advanced into the next year's field trials. As plant breeding progresses, the size and number of the field plots increase. Throughout the selection process, individual inbred lines are screened using molecular markers for quality assurance and quality control to confirm the parent–offspring relationship and to test for homozygosity.

The selection process for inbreds and hybrids is very similar, with the hybrid development process illustrated in Fig. 2. In both cases, the process starts with large populations of candidate lines that are evaluated for numerous agronomic characteristics in a small number of locations. Each field season, the top-performing candidate lines within a breeding population are advanced into the next field season's testing. As materials advance, a reduced number of candidate populations are tested in an increasing number of locations and with larger plots at each location. This procedure often means that, when summed, hundreds of thousands of individual plants are grown in hundreds of different environments. This ensures that the potential commercial hybrids are tested in numerous environments

Table 2. Impact of inheritance of virescence on the size of the starting population and the selected number of individuals advanced from an F_2 to an F_3 generation. Plant breeders can adjust selection intensity and population size to account for undesirable characteristics when inheritance for the characteristic is predictable. The undesirable phenotype in this example is represented by the single-gene expression of virescence, and the desired phenotype is homozygous recessive.

Breeding strategy	Starting F_2 population size	Phenotypic selection intensity	Intensity of marker-assisted selection of nonvirescence	Individuals advanced to F_3 generation	Outcome
		%	%		
		Virescence locus homozygous (no segregation for virescence allele)			
Standard selection practice	2000	12	N/A	240	Adequate no. of individuals and genetic gain target
		Virescence locus heterozygous (segregation for virescence allele)			
No adjustment in selection intensity	2000	12	25	60	Inadequate no. of individuals to reach pipeline target
Relax selection intensity	2000	48	25	240	Reduced genetic gain
Increase population size	8000	12	25	240	Adequate no. of individuals and genetic gain target

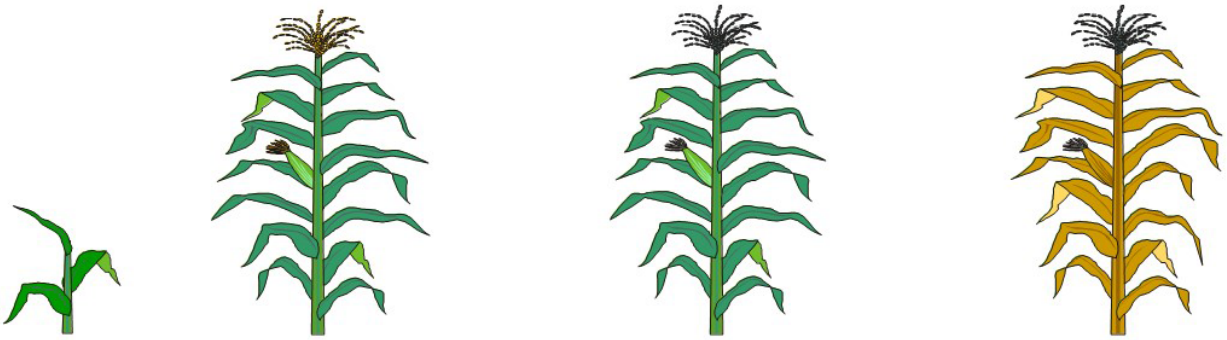
and stressors to understand the durability of the genetics and to monitor for unintended agronomic or performance effects that might only be observed in certain conditions.

In developing a new maize inbred, a breeder crosses two parental lines to produce a new segregating population that will be evaluated for numerous agronomic characteristics (Fig. 3). If one or both of these parental lines are elite (and possibly have been used in other commercial breeding programs), the offspring in this breeding population may not exhibit significant unintended or inferior characteristics but will simply need to be evaluated to identify the best performing amongst them. Sometimes a non-elite or nonadapted parent may need to be used to introduce the desired characteristics. In this case, a specific combination of parents can unexpectedly combine inferior genetics for a characteristic (e.g., delayed maturity, inadequate husk coverage, susceptibility to disease, virescence), and many of the individuals within the population will be discarded as unsuitable for further breeding due to the deficiency. As described below, marker-assisted selection can be applied to enable very specific selection of the genes of the elite germplasm with desirable characteristics while selecting against the undesired genes from the non-elite parent (Butruille et al., 2015, Chang and Coe, 2009, Eathington et al., 2007).

The top-performing candidates are repeatedly crossed back to the elite parent (backcrossing eliminates 50% of the non-elite genome with each cross), followed by ongoing evaluation and selection of the offspring with the most desired agronomic characteristics while removing any offspring with undesired characteristics. By the sixth backcross, the selected offspring will theoretically contain >99% of the DNA from the elite parent and <1% from the non-elite parent (Fig. 4). The offspring developed through this process should display the characteristics of interest from the non-elite parent, few or none of the undesirable characteristics from the non-elite parent, and all of the desired characteristics of the elite parent.

TECHNOLOGIES USED TO EXPEDITE THE BREEDING PROCESS

Creating inbred lines through self-pollination is a time-consuming process, as numerous generations of plants must be grown to maturity. Doubled-haploid technology is used to produce homozygous materials more quickly (Chang and Coe, 2009; Prigge and Melchinger, 2012). Doubled-haploid methods start from an initial population



Young Plant

- Percent emergence
- First leaf shape
- Plant vigor[1]
- Seedling height
- Seedling color

During Pollination

- Pre-anthesis brittle snap
- Pre-anthesis root lodging
- Density of spikelets
- Glume color
- Anther color
- Leaf color
- Silk color
- Days to pollen shed[2]
- Days to silking[2]

Post Pollination

- Brace root color
- Stalk anthocyanin
- Internode direction
- Internode length
- Leaf sheath pubescence
- Husk color fresh
- Plant height
- Ear height
- Stalk diameter
- Nodes with brace roots
- Leaf angle
- Upper leaf number
- Ear leaf length
- Ear leaf width
- Tassel length
- Tassel spike length
- Tassel peduncle length
- Tassel branch number
- Ear position
- Ears per stalk
- Days to husk opening[2]
- Husk color dry
- Husk cover
- Husk length

At Harvest

- Ear length
- Ear shape
- Ear weight
- # kernel rows
- # kernels per row longitudinally
- Ear diameter
- Cob diameter
- Cob color
- Kernel length
- Kernel width
- Kernel thickness
- Kernel row direction
- Kernel type
- Kernel cap color
- Kernel side color
- Endosperm color
- Endosperm type
- 100 kernel weight

Fig. 3. Agronomic characteristics observed during selection of inbreds and hybrids in conventional breeding. The agronomic characteristics listed are observed at very specific growth stages during the plant lifecycle. These agronomic and phenotypic measurements provide data for maize breeders to select varieties for advancement in breeding programs. [1] Visual appraisal of the vegetative plant on a 1-to-9 scale: 1 = excellent vigor to 9 = poor vigor; [2] Days from planting until ~50% of the plants are showing the characteristic.

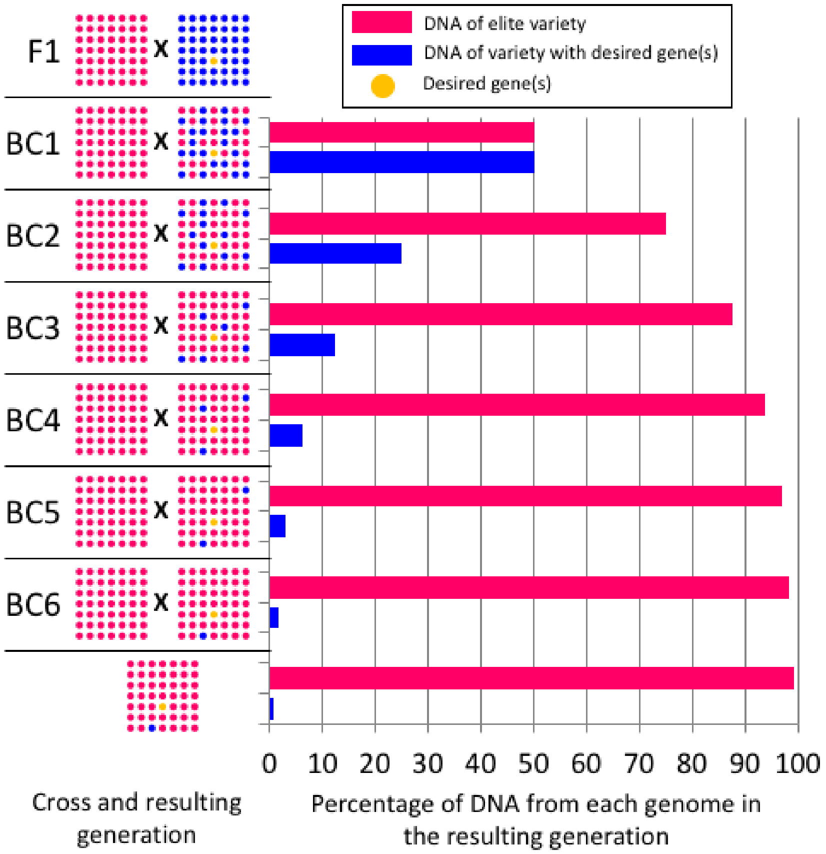


Fig. 4. Percentage of DNA from each genome in the resulting generation after backcrossing (BC) a variety (blue dots and blue bars) that has a beneficial characteristic (yellow dot) to elite germplasm (pink dots and pink bars). By the sixth backcross, the offspring will theoretically contain >99% of the DNA from the elite parent with <1% from the non-elite parent. The beneficial characteristic could be a native gene introduced during conventional breeding or a genetically modified trait incorporated during the trait introgression process.

of plants in which haploid variants (containing a single set of chromosomes) are developed and then treated to induce duplication of their chromosomes to produce diploid plants with identical chromosomal pairs. This new inbred is a genetically stable line that can undergo observation and testing to determine its value as a breeding parent.

During the generation of inbred lines, marker-assisted selection can be applied to enrich for characteristics of interest and the favorable quantitative trait loci (QTLs) that underlie the characteristics (Eathington et al., 2007). For instance, Monsanto's breeding program uses a subset of >50 different QTLs that exist for each of the characteristics of days to silking, days to anthesis, and grain oil content (Laurie et al., 2004; Buckler et al., 2009). Additionally, there are >100 QTLs associated with the development of root architecture in maize (Zurek et al., 2015), and breeding companies screen for a targeted subset of them within their breeding pipeline. Markers associated with virescence and other undesirable characteristics are used to identify, and subsequently discard, progeny that contain these undesired characteristics while advancing plants with the desirable QTLs. Seed chipping technology, combined with marker-assisted selection, allows plant breeders to select against undesired alleles with greater efficiency and accuracy, and at a reduced cost (Butruille et al., 2015). Seed chipping technology allows for identification of the genetics of the candidate lines from a small tissue sample taken from the seed. This negates the need to grow all the seeds in the field to identify the best lines.

SELECTION FOR ADAPTABILITY TO VARIABLE ENVIRONMENTAL CONDITIONS

Geographic Adaptation

The specific requirements for a given crop can differ greatly by the geographic location in which the crop is being grown. For example, some crop diseases can be found in many parts of the world, such as northern corn leaf blight [*Exserohilum turcicum* (Pass.) K.J. Leonard & E.G. Suggs] (Leonard et al., 1989), and resistance to these diseases is an important characteristic in most maize hybrids. Other diseases are specific to certain areas. Commercial-scale plant breeding programs must constantly evaluate the different geographic needs for each crop when determining which characteristics to select for in their breeding program (Butruille et al., 2015). For example, in the western US plains (e.g., Nebraska and Kansas), maize products with exceptional tolerance to Goss's wilt [*Clavibacter michiganensis* subsp. *nebraskensis* (Vidaver & Mandel) Davis et al.] are required by growers. The presence of the pathogen, environmental conditions (specifically the frequency of hail events), and management practices (such as continuous maize rotations) increase the importance of tolerance to this disease when developing maize hybrids for the western plains. Table 3 shows the shift in disease screening efforts to develop commercial products with improved Goss's wilt tolerance. A fivefold increase in early-generation screening for Goss's wilt resulted in approximately five times more commercial products with acceptable tolerance to the disease.

Environmental Effects and the Need for Testing across Multiple Seasons

Repeatability is an important factor when considering the commercial potential of new genetics. Temperature, water availability, solar radiation, and insect and disease pressures all vary by environment and by season, along with agronomic practices. Multienvironment and multiyear trials allow breeders to test their candidates for variations in performance among different geographies, year-to-year environmental variations, and response to different types of management

Table 3. Representation of a shift in screening efforts in the Monsanto North America corn breeding program with enhanced screening in 2008 to improve tolerance to Goss's wilt.†

First-year testing	Commercial deployment year	Early pipeline families screened	Available screening locations	Corn products with moderate resistance or greater
				%
2005	2011	9,800	3	13
2008	2014	49,000	6	61

† Based on Monsanto internal data and seed portfolio ratings.

(Butruille et al., 2015). For example, most farmers will plant products with a relative maturity (RM) that matches the historic RM average of their region, such as 105 RM. However, given weather fluctuations from year to year, the measure of growing degree units accumulation that represents a 105 RM zone actually occurs <40% of the time.

Testing and selection across multiple seasons is critical to ensure that products are well adapted to future weather conditions and fluctuations. Compared with US Midwest agricultural climatic conditions seen during 1971 to 2000, the projected climatic conditions in 2070 to 2099 for many important weather variables (e.g., the number of consecutive days with rain or the number of days with frost or high nighttime temperatures) will change by >30 d, requiring new varieties suited for those new weather conditions (Hatfield et al., 2014).

BREEDING OF GM CROPS

Sources of Candidate Genes for GM Crops and Early Testing

The use of transgenes has become a high-profile complement to conventional breeding. Conventional plant breeding has successfully improved crops through selection practices that capture genetic gains, even though information on the specific genes and genetic networks that contribute to the desired agronomic characteristics is usually limited. In contrast, genetic engineering requires prior knowledge of the desired gene(s) for introduction into the plant. Several recent articles provide an overview of the GM crop development process from discovery through commercialization (Privalle et al., 2012; Mumm, 2013; Prado et al., 2014; NAS, 2016). The first step in the development of a GM crop is to identify a gene that confers the desired characteristic (referred to as a "trait" in GM crops). For this publication, genetic modification and GM crops refer to plants with a particular gene (or multiple genes) isolated from its source and directly introduced into the plant's genome, and the end product is a new variety containing the desired trait (Ricroch and Hénard-Damave, 2016). Once inserted, the gene responsible for the desired trait is inherited by conventional Mendelian genetics during subsequent breeding (Weber et al., 2012). In most cases, the copied gene expresses a protein that confers the desired trait in the crop (Prigge and Melchinger, 2012; Prado et al., 2014). This review will focus on genes that encode proteins that confer the desired trait, using an herbicide resistance trait as an example.

The sources of genes range from the crop itself (and its close relatives) to more distantly related plant species or microbes, but not animal sources. Prior to making any GM crops, all potential proteins that would be produced by the inserted gene (called a "transgene") are screened using computer algorithms to ensure that they are not similar to known allergens, toxins, antinutritional proteins, disease agents, or pharmacologically active proteins (Delaney et al., 2008; Silvanovich et al., 2009; Hammond et al., 2013). For example, comparing candidate proteins to known allergens is typically done using sequence alignment programs, such as BLAST or FASTA, and by using a so-called sliding eight-amino-acid window search. Due to the precautionary nature of the sliding window search, ~15% of candidate sequences for transformation will yield a hit with proteins in an allergen database (Silvanovich et al., 2006). Any genes with hits with proteins in the allergen database are removed from further development.

During discovery, a large number of gene products are identified and tested by in vitro assays and/or in noncrop plants [such as *Arabidopsis thaliana* (L.) Heynh.], and ultimately within the intended crop (Mumm, 2013). The desired gene is cloned into a plasmid vector that can be transferred into a plant cell. A DNA construct (called an expression cassette) includes the gene(s) of interest and additional DNA sequences needed for gene expression in the plant. This expression cassette is part of the vector, as is a selectable marker that enables the identification of successfully transformed cells and progeny plants. In recent years, the trend is to include multiple genes of interest within a single DNA construct (Weber et al., 2012; NAS, 2016). For example, for effective insect control, different mechanisms of action can be identified by a combination of competitive binding experiments and testing new control agents against insects resistant to other insecticides (Pardo-López et al., 2013; Ladics et al., 2015a, 2015b; Chakroun et al., 2016; Jerga A et al., 2016; Jurat-Fuentes and Crickmore, 2017; Moar et al., 2017). Typically, dozens of versions of an expression cassette are tested in plants, both in greenhouse and confined field trials, to determine which best optimizes the gene expression needed to ensure commercial viability and reproducible efficacy of the trait. Furthermore, when the cassettes contain multiple genes (allowing more biotech traits to be expressed within a single inserted genetic locus), even more rigorous testing is conducted to ensure that all of the genes are expressed as expected. This process of gene discovery and experiments to confirm that the gene achieves the desired characteristic (often called "proof of concept" studies) can take 2 to 3 yr (Privalle et al., 2012; Mumm, 2013; Prado et al., 2014). For single-gene qualitative traits (e.g., herbicide tolerance with a well-understood plant biochemical pathway), a large percentage of candidate genes will advance out of this proof-of-concept stage.

BASICS OF *AGROBACTERIUM*-MEDIATED DNA INSERTION

Once the gene of interest in the expression cassette has passed the proof of concept studies, a transformation vector consisting of the transgene and a selectable marker is designed and used to transform cultured plant cells (Prado et al., 2014). *Agrobacterium tumefaciens* Smith & Townsend is a soil bacterium that naturally inserts pieces of DNA into relatively random spots within plant genomes (Mehrotra and Goyal, 2012; Bourras et al., 2015). *Agrobacterium* is unique among plant pathogens in that it causes disease (called "crown gall") by transferring genes into plant cells that produce compounds that aid further infection. These disease-causing *Agrobacterium* genes can be replaced with desirable genes (Mehrotra and Goyal, 2012; Bourras et al., 2015). This feature of *Agrobacterium* is exploited to insert the transgene and selectable marker into cultured plant cells. The plant cells are then screened on selection medium to identify transformed plant cells that can be regenerated into transgenic events (Schnell et al., 2015). The term "event" is commonly used to refer to each unique clone produced from DNA transformation. The random nature of the insertion into the plant's genome results in events with different molecular characteristics and expression profiles; therefore, additional screening is needed to select the most desirable event, a process that is similar to how germplasm variation is used to select the desired phenotype in conventional breeding.

THE EVENT SELECTION PROCESS

The goal of event selection is to use multiyear, multigenerational testing to identify and remove events that lack the desired trait for the product concept, and those that demonstrate undesired characteristics (Privalle et al., 2012; Mumm, 2013; Prado et al., 2014). Data from a recently completed event selection process for the development of a new herbicide-tolerant GM maize hybrid is used as a specific example.

SCREENING THE INITIAL GENERATION OF TRANSGENIC PLANTS

In the example shown in Fig. 5, transformation resulted in the generation of 5236 individual, unique transgenic events (the initial generation [R0]) that each contained a two-transgene insert and a selectable marker gene. During the initial screening process, high-throughput polymerase chain reaction (PCR) assays were performed on leaf samples to identify and eliminate events that contained either more than one inserted copy of the desired DNA and/or contained vector backbone DNA (part of the DNA used by the bacterium for the DNA insertion process but not part of the expression cassette). In this example, 3936 events were eliminated and 1300 events were advanced (Fig. 5).

Next, the remaining 1300 events were analyzed for gene expression and effectiveness (trait efficacy). Cultured cells were treated with the herbicide, and events were selected that showed tolerance to the herbicide. In addition, events were further screened for commercial-level herbicide tolerance by applying the target herbicide and propagating the tolerant plants. After the R0 trait efficacy evaluation, 642 events were advanced (Fig. 5). If the transgene had been for above-ground insect control, leaf bioassays could be used for early screening to select for efficacy against the target insect. For other characteristics that are not as easily selectable, gene expression is measured at this stage, and events that do not express the target proteins and genes at a determined threshold would be discarded.

Depending on the product concept and configuration of the transformation vector, it may be desirable to remove the selectable marker from the commercial product, in which case a linkage assay is run for each event to determine whether the insertion of the selectable marker is independent of the transgene. Linkage means that the transgene and selectable marker are inserted close together on a chromosome and are unlikely to segregate away from each other in the offspring; therefore, they would almost

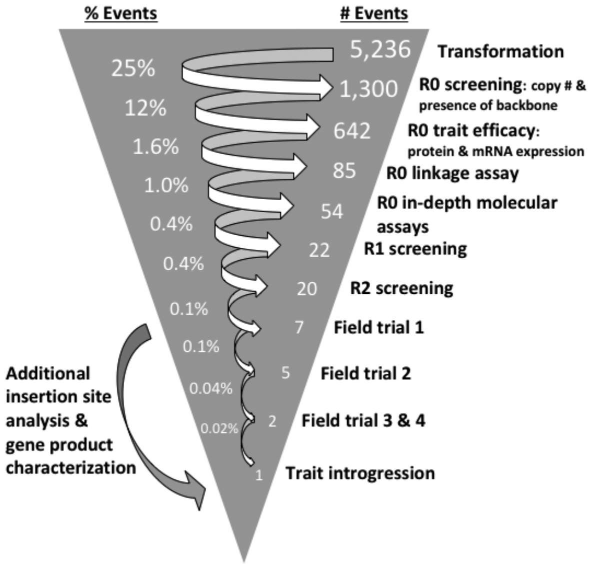


Fig. 5. The percentage and number of events advanced through each stage of a recently completed event selection process for the development of a new herbicide-tolerant genetically modified (GM) maize hybrid.

always be inherited together. In unlinked events, the marker is inserted far enough away from the transgene insertion that it will segregate during the next generation (R1) and can be eliminated (Matthews et al., 2001; Miller et al., 2002). In this example, removal of the selectable marker was desired, and events with linked selectable markers were discarded; 85 events had unlinked markers and were advanced (Fig. 5).

Additional molecular screening that is more comprehensive than the high-throughput PCR screening described above was applied in the R0 stage to ensure: (i) that only one copy of the DNA insert was in the selected events, (ii) the intactness and integrity of the DNA insert, (iii) the absence of any undesired DNA from the transformation process (e.g., the vector backbone), and (iv) that the DNA insertion had not interrupted endogenous genes (Kovalic et al., 2012). In this example, 54 events were advanced after in-depth R0 molecular analysis (Fig. 5). The discarded events contained either additional copies, fragments, or rearranged copies of the inserted DNA, or the inserted DNA interrupted endogenous genes. In addition to characterization of the DNA insertion, the RNA and/or proteins resulting from the inserted DNA were characterized to confirm that the intended gene products were being produced.

While R0 stage molecular screening and trait efficacy was progressing, general plant health was also monitored, and events with poor health and/or unintended phenotypes (“off-types”) were discarded. Examples of off-types include plants that show phenotypes such as poor germination, bleached tissues, discoloration, reduced plant height, or delays in silking or flowering. For this example, 5182 of the initial 5236 events (99%) were removed from the commercial product development pipeline prior to the completion of the initial generation (R0). The preliminary data for the remaining 54 R0 events showed that each had a single, intact, correct copy of the DNA insert, the insert had not disrupted any known endogenous genes, and the early field-testing results had shown acceptable levels of trait efficacy and no obvious “off-type” phenotypic characteristics.

FIELD SCREENING TO SELECT LEAD EVENT

The remaining 54 R0 events were self-pollinated (to generate R1 inbred seed) and outcrossed with elite lines (to generate F_1 hybrid seed). The resulting seeds were advanced to small field trial evaluations for continued measurement of trait efficacy and numerous agronomic characteristics (Fig. 5 and 6). Similar to the process illustrated in Fig. 2 for conventional breeding, as the number of GM events in the population pool decreased, the number of field trail locations and replications for the remaining events increased, with additional testing occurring at each stage (e.g., phenotypic, trait efficacy). In this example, 22 and then 20 events were advanced through small-scale R1 and R2 generation field screens, respectively (Fig. 5 and 6). This was followed by first seven, then five, and eventually only two of the initial 5236 events being extensively evaluated in four successive large-scale field trials (Fig. 5 and 6), with the outcome being selection of a single highly suitable event for advancement to commercial development. By the time the single event was selected for commercialization and entered both regulatory safety evaluations and trait introgression (the next stages of commercial development, discussed later, Fig. 6), >300,000 individual plants were grown and observed in >20 inbred or hybrid lines (Table 4) throughout many different environmental and stress conditions.

In addition to the event selection process, a rigorous seed quality process is in place to ensure that the seed used for regulatory safety studies and trait introgression is from the selected event and has not cross-pollinated with other events (Fig. 6). In maize, this process typically takes two generations and begins by tracing the selected event back to a single ear produced from a single homozygous plant. In the example detailed above, molecular assays were performed to confirm the identity of the plant producing this single ear and the resulting progeny, and plants whose identity could not be confirmed were discarded. In addition to the confirmation of identity, PCR analyses were performed for all other events grown within 200 m that might have shed pollen within a 3-wk time period of the pollination of the commercial event to ensure that no contamination had occurred. These analyses were performed several times throughout the seed quality process to ensure that the seed used for commercialization was from the selected event.

TRAIT INTROGRESSION

After event selection, the next phase of commercial development further reduces any risk of unintended effects in the final commercial variety. Trait introgression is the process of transferring the DNA insertion (with as little as possible of the event’s genomic sequence around this insertion) into numerous different lines that have agronomic characteristics optimized for their growing region or that meet specific agronomic demands.

Most maize varieties and inbred lines are not readily amenable to genetic engineering. Therefore, inbred lines that can readily incorporate the inserted DNA are used for the transformation. The inbred lines that can be genetically engineered are typically not elite lines; that is, they are older varieties that are not commercially competitive with current elite varieties or hybrids. Therefore, the DNA insertion is transferred from the selected event into elite inbred lines or varieties through a series of backcrosses. The GM trait in any chosen event will be marketed in at least 20 to 40 commercial varieties or hybrids.

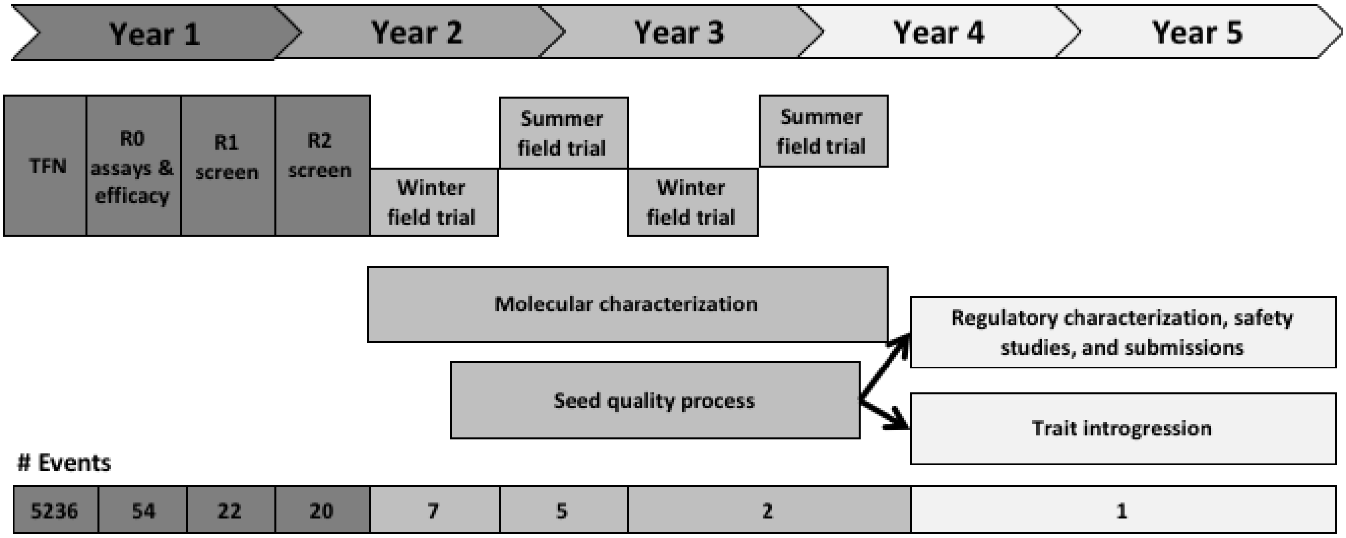


Fig. 6. Timeline of the event selection process. This figure illustrates the steps and approximate timing of the process of selecting an event for commercialization. The number of events from the example of the herbicide tolerant genetically modified maize variety is shown at the bottom. The process begins with the transformation (TFN) and progresses through an average of six field seasons (six generations) of evaluation and selection, which include two seasons of efficacy screening (R1 screen and R2 screen) and four seasons of field trials (two summer trials and two winter trials). All the while, molecular characterization of the events continues. The seed quality process typically begins after a full field season and is used to produce seed and confirm the identity and purity of the seed that will be provided for regulatory studies and trait introgression.

Table 4. Overview of greenhouse and field trials for the selection of a commercial genetically modified maize event.

Parameter	Greenhouse	Nursery	Field trial
Plants analyzed†	65	10,000–18,000	336,000–358,000
Generations analyzed	12	22	8
Agronomic characteristics observed	Plant vigor‡, days to silking§, days to pollen shed§	Germination, plant vigor‡, days to silking§, days to pollen shed§	Stand counts, trait efficacy, plant height, days to silking§, days to pollen shed§, shell weight, test weight, moisture, grain yield

† The number of plants in the nursery and field trials was estimated by multiplying the typical number of plants that are planted in each plot by the number of plots planted.

‡ Visual appraisal of the vegetative plant on a 1-to-9 scale: 1 = excellent vigor to 9 = poor vigor.

§ Days from planting until ~50% of the plants are showing the characteristic.

The trait introgression process can be expedited by incorporating the use of molecular markers. Molecular markers act as points of reference across the genomes of the event and the elite lines but are unique to each genome, allowing the rapid identification and selection of individual backcross progeny that have a higher proportion of elite line germplasm (Eathington et al., 2007). A typical trait introgression process includes approximately six backcross generations. The use of molecular markers can accelerate the process by facilitating the selection of progeny that contain the greatest proportion of elite line germplasm and reduce this to as few as three rounds of backcrossing.

As shown in Fig. 4, during successive backcrossing using current conventional plant breeding practices (including the use of molecular markers), theoretically >99% of the event's genome is removed from the offspring. DNA from the genetically engineered inbred that is chromosomally proximate to the DNA insertion may carry forward in the offspring ("linkage drag"), but the vast majority of the DNA in commercialized seed (and the harvested grain) is from the nontransformed, elite breeding lineage. In this regard, introgressing a transgene is analogous to backcrossing a trait from a landrace or nonadapted germplasm into an elite inbred. Therefore, these breeding processes segregate out any potential genomic mutations or epigenetic changes (variations in gene expression that are not from a genomic sequence change) that might have occurred during transformation or cell culture.

Any one cycle of trait introgression typically covers three to six generations; however, as is true for characteristics brought in through conventional breeding practices, the cycle is repeated yearly as new elite germplasm emerges from the breeding pipeline. Additionally, the DNA insertion is not integrated into the new suite of elite germplasm from the original event, but from the most recent integrated elite line with a similar genetic background. Thus, the genomic DNA around a given DNA insertion potentially becomes progressively more diluted with the DNA from the conventionally bred line at every generation of backcrossing. In the case of the DNA insertion in Table 5, the event's flanking DNA was no longer detectable even within 20 cM of the DNA insertion after 11 backcrosses. In this analysis, a value of 50 for the "average percentage of event DNA that is within 20 cM of the DNA insertion" would indicate that half of the DNA within 20 cM of the DNA insert is attributed to DNA from the initial transformation inbred (this analysis excludes the first centimorgan adjacent to the inserted DNA).

PLANT SELECTION PRACTICES MINIMIZE UNSAFE UNINTENDED CHANGES

This review of plant breeding and selection practices highlights how these methods are used to introduce beneficial characteristics into crops, whether from conventional breeding sources or from biotechnology. Importantly, these practices also minimize the advancement of unintended changes that might affect the safety of a new variety (European Commission, 2010; NAS, 2016).

When genetic engineering of plants was first being developed, it was hypothesized that this technology might induce potentially unintended changes that affect food or feed safety—for example, by activating previously dormant pathways in the plant (Kessler et al., 1992). As a result, extensive regulatory requirements for GM crops, which use a comparative safety assessment process, are now in place (König et al., 2004; Cellini et al., 2004; EFSA, 2006; Paoletti et al., 2008; CODEX, 2009; Privalle et al., 2012; Hoekenga et al., 2013; Prado et al., 2014). Since that time, numerous studies have found that GM varieties are compositionally equivalent to conventional crops (Parrott et al., 2012; Herman and Price, 2013; Hoekenga et al., 2013; Ricroch, 2013; Xu et al., 2014; Ladics et al., 2015a, 2015b; Curran et al., 2015; Venkatesh et al., 2015, 2016). The exceptions are a few cases where the desired trait confers an intended change in composition, such as improved nutrition (Chassy et al., 2008). Notably, >1300 independent global regulatory agency reviews of GM crops have concluded over the past 20+ years that, excluding GM crops with intentionally improved composition, all assessed characteristics of marketed varieties of GM crops (e.g., composition, agronomic and phenotypic) are equivalent to varieties with a history of safe use (European Commission, 2010; NAS, 2016; CLI, 2017). This is, in part, a result of the same plant selection practices being used by breeders to minimize undesirable unintended effects, whether derived from the spontaneous genetic changes that occur during conventional breeding (collectively also known as insertional effects; Schnell et al., 2015) or from the use of biotechnology to insert DNA into the plant genome.

Table 5. Dilution of DNA from the genome of a genetic modification event through successive backcrossing with elite varieties.

No. of backcrosses	No. of lines assessed	Average event DNA that is within 20 cM of the DNA insertion†
		%
0	1	100
6	13	29
7	6	26
8	18	22
9	32	9
10	25	8
11	7	0
12	10	0

† DNA within 1 cM of the DNA insertion is excluded from the analysis.

CONCLUDING REMARKS

Plant breeding is a process of crop improvement that continuously addresses changing needs by introducing new genetic diversity into product portfolios. The breeding and selection processes for both conventional and GM crops provide multiple opportunities to eliminate adverse unintended effects resulting from conventional breeding and/or the transformation process. The additional rigorous molecular and phenotypic characterization of GM crops further ensures that the inserted DNA performs as intended in the crop and does not confer adverse unintended effects. Combined with an inherently low ability of genomic changes to produce harmful effects, the creation of new plant varieties has one of the safest records of all human technologies.

Conflict of Interest

During the conduct of this study, many of the authors (K.C. Glenn, E. Bell, M. Goley, J. Jenkinson, B. Liu, C. Martin, C. Souder, O. Sparks, W. Urquhart, and J.L. Vicini) were employees of Monsanto Company and were provided financial support in the form of authors' salaries and research materials.

Acknowledgments

The authors thank Drs. Luis Burzio, Steven Halls, and Stefan Henkes for careful editorial review of this manuscript.

References

Acquaah, G. 2012. Principles of plant genetics and breeding. Wiley, Somerset, NJ. doi:10.1002/9781118313718

Anderson, J.E., J.M. Michno, T.J. Kono, A.O. Stec, B.W. Campbell, S.J. Curtin, and R.M. Stupar. 2016. Genomic variation and DNA repair associated with soybean transgenesis: A comparison to cultivars and mutagenized plants. BMC Biotechnol. 16:41. doi:10.1186/s12896-016-0271-z

Auld, D.L., M.K. Heikkinen, D.A. Erickson, J.L. Sernyk, and J.E. Romero. 1992. Rapeseed mutants with reduced levels of polyunsaturated fatty acids and increased levels of oleic acid. Crop Sci. 32:657–662. doi:10.2135/cropsci1992.0011183X003200030016x

Bergthorsson, U., K.L. Adams, B. Thomason, and J.D. Palmer. 2003. Widespread horizontal transfer of mitochondrial genes in flowering plants. Nature 424:197–201. doi:10.1038/nature01743

Bock, R. 2010. The give-and-take of DNA: Horizontal gene transfer in plants. Trends Plant Sci. 15:11–22. doi:10.1016/j.tplants.2009.10.001

Bolon, Y.T., W.J. Haun, W.W. Xu, D. Grant, M.G. Stacey, R.T. Nelson et al. 2011. Phenotypic and genomic analyses of a fast neutron mutant population resource in soybean. Plant Physiol. 156:240–253. doi:10.1104/pp.110.170811

Bourras, S., T. Rouxel, and M. Meyer. 2015. *Agrobacterium tumefaciens* gene transfer: How a plant pathogen hacks the nuclei of plant and nonplant organisms. Phytopathology 105:1288–1301. doi:10.1094/PHYTO-12-14-0380-RVW

Buckler, E.S., B.S. Gaut, and M.D. McMullen. 2006. Molecular and functional diversity of maize. Curr. Opin. Plant Biol. 9:172–176. doi:10.1016/j.pbi.2006.01.013

Buckler, E.S., J.B. Holland, P.J. Bradbury, C.B. Acharya, P.J. Brown, C. Browne et al. 2009. The genetic architecture of maize flowering time. Science 325:714–718. doi:10.1126/science.1174276

Butruille, D.V., F.H. Birru, M.L. Boerboom, E.J. Cargill, D.A. Davis, P. Dhungana et al. 2015. Maize breeding in the United States: Views from within Monsanto. In: J. Janick, editor, Plant breeding reviews. Vol. 39. John Wiley & Sons, New York p. 199–282.

Cellini, F., A. Chesson, I. Colquhoun, A. Constable, H.V. Davies, K.H. Engel et al. 2004. Unintended effects and their detection in genetically modified crops. Food Chem. Toxicol. 42:1089–1125. doi:10.1016/j.fct.2004.02.003

Chakroun, M., N. Banyuls, Y. Bel, B. Escriche, and J. Ferré. 2016. Bacterial vegetative insecticidal proteins (Vip) from entomopathogenic bacteria. Microbiol. Mol. Biol. Rev. 80:329–350. doi:10.1128/MMBR.00060-15 [erratum: 80:iii].

Chang, M.T., and E.H. Coe, Jr. 2009. Doubled haploids. In: A.L. Kriz and B.A. Larkins, editors, Molecular genetic approaches to maize improvement. Springer, Berlin. p. 127–142. doi:10.1007/978-3-540-68922-5_10

Chassy, B., M. Egnin, Y. Gao, K. Glenn, G.A. Kleter, P. Nestel et al. 2008. Nutritional and safety assessments of foods and feeds nutritionally improved through biotechnology: Case studies. Compr. Rev. Food Sci. Food Saf. 7:50–113.

CLI. 2017. BioTradeStatus. CropLife Int. www.biotradestatus.com (accessed 25 July 2017)

CODEX. 2009. Foods derived from modern biotechnology. FAO, WHO. http://www.fao.org/3/a-a1554e.pdf (accessed 8 Jan. 2017).

Cook, D.E., T.G. Lee, X. Guo, S. Melito, K. Wang, A.M. Bayless et al. 2012. Copy number variation of multiple genes at *Rhg1* mediates nematode resistance in soybean. Science 338:1206–1209. doi:10.1126/science.1228746

Curran, K.L., A.R. Festa, S.D. Goddard, G.G. Harrigan and M.L. Taylor. 2015. Kernel compositions of glyphosate-tolerant and corn rootworm-protected MON 88017 sweet corn and insect-protected MON 89034 sweet corn are equivalent to that of conventional sweet corn (*Zea mays*). J. Agric. Food Chem. 63:3046–3052. doi:10.1021/jf505687s

Curry, H.A. 2016. Evolution made to order: Plant breeding and technological innovation in twentieth-century America. Univ. of Chicago Press, Chicago. doi:10.7208/chicago/9780226390116.001.0001

Da Silva, C., G. Zamperin, A. Ferrarini, A. Minio, A. Dal Molin, L. Venturini et al. 2013. The high polyphenol content of grapevine cultivar tannat berries is conferred primarily by genes that are not shared with the reference genome. Plant Cell 25:4777–4788. doi:10.1105/tpc.113.118810 [erratum: 29:913].

Delaney, B., J.D. Astwood, H. Cunny, R. Eichen Conn, C. Herouet-Guicheney, S. MacIntosh et al. 2008. Evaluation of protein safety in the context of agricultural biotechnology. Food Chem. Toxicol. 46:S71–S97. doi:10.1016/j.fct.2008.01.045

Doebley, J. 2004. The genetics of maize evolution. Annu. Rev. Genet. 38:37–59. doi:10.1146/annurev.genet.38.072902.092425

Eathington, S.R., T.M. Crosbie, M.D. Edwards, R.S. Reiter, and J.K. Bull. 2007. Molecular markers in a commercial breeding program. Crop Sci. 47:S154–S163. doi:10.2135/cropsci2007.04.0015IPBS

EFSA. 2006. Guidance document for the risk assessment of genetically modified plants and derived food and feed by the Scientific Panel on Genetically Modified Organisms (GMO)—including draft document updated in 2008. EFSA J. 4:99. doi:10.2903/j.efsa.2006.99

El Baidouri, M., M.C. Carpentier, R. Cooke, D. Gao, E. Lasserre, C. Llauro et al. 2014. Widespread and frequent horizontal transfers of transposable elements in plants. Genome Res. 24:831–838. doi:10.1101/gr.164400.113

Ellis, T.H., J.M. Hofer, G.M. Timmerman-Vaughan, C.J. Coyne, and R.P. Hellens. 2011. Mendel, 150 years on. Trends Plant Sci. 16:590–596. doi:10.1016/j.tplants.2011.06.006

European Commission. 2010. A decade of EU-funded GMO research (2001–2010). Directorate-General for Research and Innovation, Biotechnologies, Agriculture, Food, European Union, Brussels. https://ec.europa.eu/research/biosociety/pdf/a_decade_of_eu-funded_gmo_research.pdf (accessed 24 Mar. 2017).

Fehr, W.R., E.L. Fehr, and H.J. Jessen. 1987. Principles of cultivar development. Macmillan, New York.

Geering, A.D., F. Maumus, D. Copetti, N. Choisne, D.J. Zwickl, M. Zytnicki et al. 2014. Endogenous florendoviruses are major components of plant genomes and hallmarks of virus evolution. Nat. Commun. 5:5269. doi:10.1038/ncomms6269

Golicz, A.A., P.E. Bayer, G.C. Barker, P.P. Edger, H. Kim, P.A. Martinez et al. 2016. The pangenome of an agronomically important crop plant *Brassica oleracea*. Nat. Commun. 7:13390. doi:10.1038/ncomms13390

Gore, M.A., J.-M. Chia, R.J. Elshire, Q. Sun, E.S. Ersoz, B.L. Hurwitz et al. 2009. A first-generation haplotype map of maize. Science 326:1115–1117. doi:10.1126/science.1177837

Hajjar, R., and T. Hodgkin. 2007. The use of wild relatives in crop improvement: A survey of developments over the last 20 years. Euphytica 156:1–13. doi:10.1007/s10681-007-9363-0

Hallauer, A.R., M.J. Carena, and J.B. Miranda-Filho. 1988. Quantitative genetics in maize breeding. Handbook of plant breeding. Springer, Washington, DC.

Hammond, B., J. Kough, C. Herouet-Guicheney, and J.M. Jez. 2013. Toxicological evaluation of proteins introduced into food crops. Crit. Rev. Toxicol. 43(Suppl. 2):25–42. doi:10.3109/10408444.2013.842956

Hatfield, J., G. Takle, R. Grotjahn, P. Holden, R.C. Izaurralde, T. Mader et al. 2014. Agriculture. In: J.M. Melillo, T.C. Richmond, and G.W. Yohe, editors, Climate change impacts in the United States: The third national climate assessment. U.S. Global Change Res. Prog., Washington, DC. p. 150–174.

Hensz, R.A. 1971. Mutation breeding and the development of the 'Star Ruby' grapefruit. Proc. Int. Soc. Citric. 2:582–585.

Herman, R.A. and W.D. Price. 2013. Unintended compositional changes in genetically modified (GM) crops: 20 years of research. J. Agric. Food Chem. 61:11695−11170. doi:10.1021/jf400135r

Hirsch, C.D., and N.M. Springer. 2017. Transposable element influences on gene expression in plants. Biochim. Biophys. Acta 1860:157–165. doi:10.1016/j.bbagrm.2016.05.010

Hirsch, C.N., C.D. Hirsch, A.B. Brohammer, M.J. Bowman, I. Soifer, O. Barad et al. 2016. Draft assembly of elite inbred line PH207 provides insights into genomic and transcriptome diversity in maize. Plant Cell 28:2700–2714. doi:10.1105/tpc.16.00353

Hoekenga, O.A., J. Srinivasan, G. Barry, and A. Bartholomaeus. 2013. Compositional analysis of genetically modified (GM) crops: Key issues and future needs. J. Agric. Food Chem. 61:8248–8253. doi:10.1021/jf401141r

Hopkins, W.G., and D.B. Walden. 1977. Temperature sensitivity of virescent mutants of maize. J. Hered. 68:283–286. doi:10.1093/oxfordjournals.jhered.a108837

IAEA-MVD. 2017. Joint FAO/IAEA mutant variety database. FAO, Int. Atomic Energy Agency. https://mvd.iaea.org/ (accessed 25 July 2017)

Institute of Medicine, National Research Council (IOM/NRC). 2004. Safety of genetically engineered foods: Approaches to assessing unintended health effects. The National Academies Press, Washington, DC.

Jerga, A., D. Chen, C. Zhang, J. Fu, J.L. Kouadio, Y. Wang et al. 2016. Mechanistic insights into the first Lygus-active β-pore forming protein. Arch. Biochem. Biophys. 600:1–11. doi:10.1016/j.abb.2016.03.016

Jones, S.S., T.D. Murray, and R.E. Allan. 1995. The development of disease resistance in wheat. Annu. Rev. Phytopathol. 33:429–443. doi:10.1146/annurev.py.33.090195.002241

Jurat-Fuentes, J.L. and N. Crickmore. 2017. Specificity determinants for Cry insecticidal proteins: Insights from their mode of action. J. Invertebr. Pathol. 142:5–10. doi:10.1016/j.jip.2016.07.018

Kessler, D.A., M.R. Taylor, J.H. Maryanski, E.L. Flamm and L.S. Kahl. 1992. The safety of foods developed by biotechnology. Science 256:1747–1749, 1832. doi:10.1126/science.1615315

König, A., A. Cockburn, R.W. Crevel, E. Debruyne, R. Grafstroem, U. Hammerling et al. 2004. Assessment of the safety of foods derived from genetically modified (GM) crops. Food Chem. Toxicol. 42:1047–1088. doi:10.1016/j.fct.2004.02.019

Konishi, S., T. Izawa, S.Y. Lin, K. Ebana, Y. Fukuta, T. Sasaki, and M. Yano. 2006. An SNP caused loss of seed shattering during rice domestication. Science 312:1392–1396. doi:10.1126/science.1126410

Kovalic, D., C. Garnaat, L. Guo, Y. Yan, J. Groat, A. Silvanovich et al. 2012. The use of next generation sequencing and junction sequence analysis bioinformatics to achieve molecular characterization of crops improved through modern biotechnology. Plant Genome 5:149–163. doi:10.3835/plantgenome2012.10.0026

Kyndt, T., D. Quispe, H. Zhai, R. Jarret, M. Ghislain, Q. Liu et al. 2015. The genome of cultivated sweet potato contains *Agrobacterium* T-DNAs with expressed genes: An example of a naturally transgenic food crop. Proc. Natl. Acad. Sci. USA 112:5844–5849. doi:10.1073/pnas.1419685112

Ladics, G.S., A. Bartholomaeus, P. Bregitzer, N.G. Doerrer, A. Gray, T. Holzhauser et al. 2015a. Genetic basis and detection of unintended effects in genetically modified crop plants. Transgenic Res. 24:587–603. doi:10.1007/s11248-015-9867-7

Ladics, G.S., A. Bartholomaeus, P. Bregitzer, N.G. Doerrer, A. Gray, T. Holzhauser et al. 2015b. Online supplementary material. Genetic basis and detection of unintended effects in genetically modified crop plants. Springer. https://static-content.springer.com/esm/art%3A10.1007%2Fs11248-015-9867-7/MediaObjects/11248_2015_9867_MOESM1_ESM.pdf (accessed 8 Jan. 2017).

Lai, J., R. Li, X. Xu, W. Jin, M. Xu, H. Zhao et al. 2010. Genome-wide patterns of genetic variation among elite maize inbred lines. Nat. Genet. 42:1027–1030. doi:10.1038/ng.684

Lam, H.-M., X. Xu, X. Liu, W. Chen, G. Yang, F.-L. Wong et al. 2010. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. Nat. Genet. 42:1053–1059. doi:10.1038/ng.715

Laurie, C.C., S.D. Chasalow, J.R. LeDeaux, R. McCarroll, D. Bush, B. Hauge et al. 2004. The genetic architecture of response to long-term artificial selection for oil concentration in the maize kernel. Genetics 168:2141–2155. doi:10.1534/genetics.104.029686

Leonard, K.J., Y. Levy, and D.R. Smith. 1989. Proposed nomenclature for pathogenic races of *Exserohilum turcicum* on corn. Plant Dis. 73:776–777.

Li, Y.H., G. Zhou, J. Ma, W. Jiang, L.G. Jin, Z. Zhang et al. 2014. De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. Nat. Biotechnol. 32:1045–1052. doi:10.1038/nbt.2979

Lisch, D. 2013. How important are transposons for plant evolution? Nat. Rev. Genet. 14:49–61. doi:10.1038/nrg3374

Liu, R., K.O. Koyanagi, S. Chen, and Y. Kishima. 2012. Evolutionary force of AT-rich repeats to trap genomic and episomal DNAs into the rice genome: Lessons from endogenous pararetrovirus. Plant J. 72:817–828. doi:10.1111/tpj.12002

Lough, A.N., L.M. Roark, A. Kato, T.S. Ream, J.C. Lamb, J.A. Birchler, and K.J. Newton. 2008. Mitochondrial DNA transfer to the nucleus generates extensive insertion site variation in maize. Genetics 178:47–55. doi:10.1534/genetics.107.079624

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, MA

Matthews, P.R., M.-B. Wang, P.M. Waterhouse, S. Thornton, S.J. Fieg, F. Gubler, and J.V. Jacobsen. 2001. Marker gene elimination from transgenic barley, using co-transformation with adjacent 'twin T-DNAs' on a standard *Agrobacterium* transformation vector. Mol. Breed. 7:195–202. doi:10.1023/A:1011333321893

Mehrotra, S., and V. Goyal. 2012. *Agrobacterium*-mediated gene transfer in plants and biosafety considerations. Appl. Biochem. Biotechnol. 168:1953–1975. doi:10.1007/s12010-012-9910-6

Miller, M., L. Tagliani, N. Wang, B. Berka, D. Bidney, and Z.Y. Zhao. 2002. High efficiency transgene segregation in co-transformed maize plants using an *Agrobacterium tumefaciens* 2 T-DNA binary system. Transgenic Res. 11:381–396. doi:10.1023/A:1016390621482

Moar, W., C. Khajuria, M. Pleau, O. Ilagan, M. Chen, C. Jiang et al. 2017. Cry3Bb1-resistant western corn rootworm, *Diabrotica virgifera virgifera* (LeConte) does not exhibit cross-resistance to DvSnf7 dsRNA. PLoS One 12:e0169175. doi:10.1371/journal.pone.0169175

Multani, D.S., S.P. Briggs, M.A. Chamberlin, J.J. Blakeslee, A.S. Murphy, and G.S. Johal. 2003. Loss of an MDR transporter in compact stalks of maize *br2* and sorghum *dw3* mutants. Science 302:81–84. doi:10.1126/science.1086072

Mumm, R.H. 2013. A look at product development with genetically modified crops: Examples from maize. J. Agric. Food Chem. 61:8254–8259. doi:10.1021/jf400685y

Naito, K., E. Cho, G. Yang, M.A. Campbell, K. Yano, Y. Okumoto et al. 2006. Dramatic amplification of a rice transposable element during recent domestication. Proc. Natl. Acad. Sci. USA 103:17620–17625. doi:10.1073/pnas.0605421103

National Academies of Science, Engineering, and Medicine (NAS), editor. 2016. Genetically engineered crops: Experiences and prospects. The National Academies Press, Washington, DC.

O'Donoghue, E., R. Hoppe, D. Banker, R. Ebel, K. Fuglie, P. Korb et al. 2011. The changing organization of U.S. farming. Econ. Res. Serv., Washington, DC.

Ossowski, S., K. Schneeberger, J.I. Lucas-Lledo, N. Warthmann, R.M. Clark, R.G. Shaw et al. 2010. The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. Science 327:92–94. doi:10.1126/science.1180677

Palaisa, K.A., M. Morgante, M. Williams, and A. Rafalski. 2003. Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. Plant Cell 15:1795–1806. doi:10.1105/tpc.012526

Paoletti, C., E. Flamm, W. Yan, S. Meek, S. Renckens, M. Fellous, and H. Kuiper. 2008. GMO risk assessment around the world: Some examples. Trends Food Sci. Technol. 19:S70–S78. doi:10.1016/j.tifs.2008.07.007

Pardo-López, L., M. Soberón, and A. Bravo. 2013. *Bacillus thuringiensis* insecticidal three-domain Cry toxins: Mode of action, insect resistance and consequences for crop protection. FEMS Microbiol. Rev. 37:3–22. doi:10.1111/j.1574-6976.2012.00341.x

Parrott, W.A., J.M. Jez, and L.C. Hannah. 2012. To be or not to be transgenic. Nat. Biotechnol. 30:825–826. doi:10.1038/nbt.2347

Peng, J., D.E. Richards, N.M. Hartley, G.P. Murphy, K.M. Devos, J.E. Flintham et al. 1999. 'Green revolution' genes encode mutant gibberellin response modulators. Nature 400:256–261. doi:10.1038/22307

Prado, J.R., G. Segers, T. Voelker, D. Carson, R. Dobert, J. Phillips et al. 2014. Genetically engineered crops: From idea to product. Annu. Rev. Plant Biol. 65:769–790. doi:10.1146/annurev-arplant-050213-040039

Prigge, V., and A.E. Melchinger. 2012. Production of haploids and doubled haploids in maize. In: V.M. Loyola-Vargas and N. Ochoa-Alejo, editors, Plant cell culture protocols. Humana Press, New York. p. 161–172. doi:10.1007/978-1-61779-818-4_13

Privalle, L.S., J. Chen, G. Clapper, P. Hunst, F. Spiegelhalter, and C.X. Zhong. 2012. Development of an agricultural biotechnology crop product: Testing from discovery to commercialization. J. Agric. Food Chem. 60:10179–10187. doi:10.1021/jf302706e

Ricroch, A.E. 2013. Assessment of GE food safety using '-omics' techniques and long-term animal feeding studies. New Biotechnol. 30:349–354. doi:10.1016/j.nbt.2012.12.001

Ricroch, A.E., and M.-C. Hénard-Damave. 2016. Next biotech plants: New traits, crops, developers and technologies for addressing global challenges. Crit. Rev. Biotechnol. 36:675–690. doi:10.3109/07388551.2015.1004521

Roark, L.M., A.Y. Hui, L. Donnelly, J.A. Birchler, and K.J. Newton. 2010. Recent and frequent insertions of chloroplast DNA into maize nuclear chromosomes. Cytogenet. Genome Res. 129:17–23. doi:10.1159/000312724

Schnable, P.S., D. Ware, R.S. Fulton, J.C. Stein, F. Wei, S. Pasternak et al. 2009. The B73 maize genome: Complexity, diversity, and dynamics. Science 326:1112–1115. doi:10.1126/science.1178534

Schnell, J., M. Steele, J. Bean, M. Neuspiel, C. Girard, N. Dormann et al. 2015. A comparative analysis of insertional effects in genetically engineered plants: Considerations for pre-market assessments. Transgenic Res. 24:1–17. doi:10.1007/s11248-014-9843-7

Shepherd, L.V.T., C.A. Hackett, C.J. Alexander, J.W. McNicol, J.A. Sungurtas, D. Stewart et al. 2015. Modifying glycoalkaloid content in transgenic potato: Metabolome impacts. Food Chem. 187:437–443. doi:10.1016/j.foodchem.2015.04.111

Silvanovich, A., G. Bannon, and S. McClain. 2009. The use of *E*-scores to determine the quality of protein alignments. Regul. Toxicol. Pharmacol. 54:S26–S31. doi:10.1016/j.yrtph.2009.02.004

Silvanovich, A., M.A. Nemeth, P. Song, R. Herman, L. Tagliani, and G.A. Bannon. 2006. The value of short amino acid sequence matches for prediction of protein allergenicity. Toxicol. Sci. 90:252–258. doi:10.1093/toxsci/kfj068

Smith, K.A. 2013. Horrific tales of potatoes that caused mass sickness and even death. Smithsonian. http://www.smithsonianmag.com/arts-culture/horrific-tales-of-potatoes-that-caused-mass-sickness-and-even-death-3162870/ (accessed 25 Jan. 2017).

Smýkal, P., C.J. Coyne, M.J. Ambrose, N. Maxted, H. Schaefer, M.W. Blair et al. 2015. Legume crops phylogeny and genetic diversity for science and breeding. Crit. Rev. Plant Sci. 34:43–104. doi:10.1080/07352689.2014.897904

Soucy, S.M., J. Huang, and J.P. Gogarten. 2015. Horizontal gene transfer: Building the web of life. Nat. Rev. Genet. 16:472–482. doi:10.1038/nrg3962

Springer, N.M., K. Ying, Y. Fu, T. Ji, C.T. Yeh, Y. Jia et al. 2009. Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. PLoS Genet. 5:e1000734. doi:10.1371/journal.pgen.1000734

Staginnus, C., W. Gregor, M.F. Mette, C.H. Teo, E.G. Borroto-Fernández, M.L.C. Machado et al. 2007. Endogenous pararetroviral sequences in tomato (*Solanum lycopersicum*) and related species. BMC Plant Biol. 7:24. doi:10.1186/1471-2229-7-24

Steiner, H.Y., C. Halpin, J.M. Jez, J. Kough, W. Parrott, L. Underhill et al. 2013. Editor's choice: Evaluating the potential for adverse interactions within genetically engineered breeding stacks. Plant Physiol. 161:1587–1594. doi:10.1104/pp.112.209817

Tenaillon, M.I., M.C. Sawkins, A.D. Long, R.L. Gaut, J.F. Doebley, and B.S. Gaut. 2001. Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). Proc. Natl. Acad. Sci. USA 98:9161–9166. doi:10.1073/pnas.151244298

Tian, Z., M. Zhao, M. She, J. Du, S.B. Cannon, X. Liu et al. 2012. Genome-wide characterization of nonreference transposons reveals evolutionary propensities of transposons in soybean. Plant Cell 24:4422–4436. doi:10.1105/tpc.112.103630

Tuteja, J.H., S.J. Clough, W.C. Chan, and L.O. Vodkin. 2004. Tissue-specific gene silencing mediated by a naturally occurring chalcone synthase gene cluster in *Glycine max*. Plant Cell 16:819–835. doi:10.1105/tpc.021352

USDA-NASS. 2017. Quick stats. USDA, Natl. Agric. Stat. Serv. https://quickstats.nass.usda.gov/results/90C69DEC-38D6-31B4-9953-4C6EB5E82D79?pivot=short_desc (accessed 22 May 2017).

Venkatesh, T.V., E. Bell, A. Bickel, K. Cook, B. Alsop, M. van de Mortel et al. 2016. Maize hybrids derived from GM positive and negative segregant inbreds are compositionally equivalent: any observed differences are associated with conventional backcrossing practices. Transgenic Res. 25:83–96. doi:10.1007/s11248-015-9910-8

Venkatesh, T.V., K. Cook, B. Liu, T. Perez, A. Willse, R. Tichich et al. 2015. Compositional differences between near-isogenic GM and conventional maize hybrids are associated with backcrossing practices in conventional breeding. Plant Biotechnol. J. 13:200–210. doi:10.1111/pbi.12248

Weber, N., C. Halpin, L.C. Hannah, J.M. Jez, J. Kough, and W. Parrott. 2012. Editor's choice: Crop genome plasticity and its relevance to food and feed safety of genetically engineered breeding stacks. Plant Physiol. 160:1842–1853. doi:10.1104/pp.112.204271

Xiao, H., N. Jiang, E. Schaffner, E.J. Stockinger, and E. van der Knaap. 2008. A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. Science 319:1527–1530. doi:10.1126/science.1153040

Xu, K., X. Xu, T. Fukao, P. Canlas, R. Maghirang-Rodriguez, S. Heuer, A.M. Ismail et al. 2006. *Sub1A* is an ethylene-response-factor-like gene that confers submergence tolerance to rice. Nature 442:705–708. doi:10.1038/nature04920

Xu, Y., R. Goodacre and G.G. Harrigan. 2014. Compositional equivalence of grain from multi-trait drought-tolerant maize hybrids to a conventional comparator: univariate and multivariate assessments. J. Agric. Food Chem. 62:9597–9608. doi:10.1021/jf5019609

Zhang, J., and T. Peterson. 2004. Transposition of reversed *Ac* element ends generates chromosome rearrangements in maize. Genetics 167:1929–1937. doi:10.1534/genetics.103.026229

Zhang, J., F. Zhang, and T. Peterson. 2006. Transposition of reversed *Ac* element ends generates novel chimeric genes in maize. PLoS Genet. 2:e164. doi:10.1371/journal.pgen.0020164

Zurek, P.R., C.N. Topp, and P.N. Benfey. 2015. Quantitative trait locus mapping reveals regions of the maize genome controlling root system architecture. Plant Physiol. 167:1487–1496. doi:10.1104/pp.114.251751

Zywicki, B., G. Catchpole, J. Draper, and O. Fiehn. 2005. Comparison of rapid liquid chromatography-electrospray ionization-tandem mass spectrometry methods for determination of glycoalkaloids in transgenic field-grown potatoes. Anal. Biochem. 336:178–186. doi:10.1016/j.ab.2004.10.013