Introgression of Crop Alleles into Wild or Weedy Populations

Norman C. Ellstrand,[1] Patrick Meirmans,[2] Jun Rong,[3]
Detlef Bartsch,[4] Atiyo Ghosh,[5] Tom J. de Jong,[6]
Patsy Haccou,[7] Bao-Rong Lu,[8] Allison A. Snow,[9]
C. Neal Stewart Jr,[10] Jared L. Strasburg,[11]
Peter H. van Tienderen,[12] Klaas Vrieling,[13]
and Danny Hooftman[14]

[1] Department of Botany and Plant Sciences, University of California, Riverside,
California 92521; email: ellstrand@ucr.edu

[2] Instituut voor Biodiversiteit en Ecosysteem Dynamica, Universiteit van Amsterdam,
1098 XH Amsterdam, The Netherlands; email: p.g.meirmans@uva.nl

[3] Center for Watershed Ecology, Institute of Life Science and Key Laboratory of Poyang Lake
Environment and Resource Utilization, Ministry of Education, Nanchang University,
330031 Honggutan Nanchang, People's Republic of China; email: rong_jun@hotmail.com

[4] Federal Office of Consumer Protection and Food Safety, 10117 Berlin, Germany;
email: detlef.bartsch@bvl.bund.de

[5] Integrative Systems Biology, Okinawa Institute of Science and Technology,
Okinawa 904-0495, Japan; email: atiyo.ghosh@oist.jp

[6] Institute of Biology Leiden, Leiden University, 2333 BE Leiden, The Netherlands;
email: T.J.de.Jong@biology.leidenuniv.nl

[7] Leiden University College The Hague, Leiden University, 2514 EG The Hague,
The Netherlands; email: p.haccou@luc.leidenuniv.nl

[8] Ministry of Education Key Laboratory for Biodiversity and Ecological Engineering,
Department of Ecology and Evolutionary Biology, Fudan University, Shanghai 200433,
People's Republic of China; email: brlu@fudan.edu.cn, lbaorong25@hotmail.com

[9] Department of Evolution, Ecology, and Organismal Biology, The Ohio State University,
Columbus, Ohio 43210; email: snow.1@osu.edu

[10] Department of Plant Sciences, University of Tennessee at Knoxville, Knoxville,
Tennessee 37996; email: nealstewart@utk.edu

[11] Department of Biology, University of Minnesota-Duluth, Minnesota 55812;
email: jstrasbu@d.umn.edu

[12] Instituut voor Biodiversiteit en Ecosysteem Dynamica, Universiteit van Amsterdam,
1090 GE Amsterdam, The Netherlands; email: P.H.vanTienderen@uva.nl

[13] Institute of Biology, Leiden University 2333 BE Leiden, The Netherlands;
email: k.vrieling@biology.leidenuniv.nl

[14] Center for Ecology and Hydrology, National Environmental Research Council, Wallingford,
Oxfordshire OX10 8BB, United Kingdom; email: dann1@ceh.ac.uk

Keywords

gene flow, hybridization, plant conservation genetics, transgene flow, evolution, invasiveness

Abstract

The evolutionary significance of introgression has been discussed for decades. Questions about potential impacts of transgene flow into wild and weedy populations brought renewed attention to the introgression of crop alleles into those populations. In the past two decades, the field has advanced with considerable descriptive, experimental, and theoretical activity on the dynamics of crop gene introgression and its consequences. As illustrated by five case studies employing an array of different approaches, introgression of crop alleles has occurred for a wide array of species, sometimes without significant consequence, but on occasion leading to the evolution of increased weediness. A new theoretical context has emerged for analyzing empirical data, identifying factors that influence introgression, and predicting introgression's progress. With emerging molecular techniques and analyses, research on crop allele introgression into wild and weedy populations is positioned to make contributions to both transgene risk assessment and reticulate evolution.

HISTORICAL PERSPECTIVE

Most crops were domesticated from wild plants centuries or even millennia ago (Warwick & Stewart 2005). Early evolution under anthropogenic selection produced domesticated plants that were locally adapted and more productive. With the advent of formal plant breeding, progress in plant improvement depended on selecting from (*a*) pre-existing variation in the evolving crop and (*b*) alleles obtained by intentional hybridization with the crop progenitor and other wild or weedy (henceforth, WW) relatives. Recently, techniques available to breeders have expanded to include methods such as human-mediated intertaxon crosses, hybrid embryo rescue, protoplast fusion, induced mutations, and transgenesis. Still, spontaneous and intentional hybridization between WW populations and locally adapted landraces continues to be a source of variation for crop improvement both in formal breeding and in traditionally managed agroecosystems (Hajjar & Hodgkin 2007, Jarvis & Hodgkin 1999).

Conversely, spontaneous hybridization can be the first step for the flow of novel crop alleles in the other direction, i.e., into WW relatives (Ellstrand 2003). Subsequent establishment of those alleles is known as introgression, "the permanent incorporation of genes from one set of differentiated populations (species, subspecies, race, and so on) into another" (Stewart et al. 2003, p. 806). The process of introgression begins when a fertile or semifertile hybrid (or even an F_2 or later segregant) successfully backcrosses with one of the parental species. Unless opposed by selection or drift, further introgression proceeds under repeated backcrossing or selfing.

The importance of crop allele introgression in the evolution of WW populations has been controversial. In the late twentieth century, deWet & Harlan (1975) recognized peripatric and sympatric wild-weed-crop complexes as zones of significant gene exchange across reproductive isolating barriers of varying permeability. They argued that two of the three avenues of exchange would be significant: (*a*) Such evolutionary hotbeds would favor the flow of beneficial alleles from the wild to the crop, providing farmers with a constant source of genetic variation useful for plant improvement and (*b*) gene flow from crops into weedy populations could stimulate the evolution of crop mimics. But their view was that "selection almost completely prevents gene flow in the

direction of the wild race" (deWet & Harlan 1975, p. 106). Others had a contrasting view, e.g., "...gene flow, if it exists, is apparently more effective in the direction from the cultivated to the wild populations" (Ladizinsky 1985, p. 191). At a time when gene flow and hybridization were often inferred from morphology alone, rather than genetically based markers, conclusive evidence for intertaxon gene flow was often lacking.

The flow of crop alleles into WW populations might have remained solely a topic for academic discussion; but applied evolutionists subsequently recognize that intertaxon gene flow can lead to evolutionary changes that impact human affairs. Hybridization can sometimes stimulate the evolution of new weeds or invasives (Schierenbeck & Ellstrand 2009) or contribute to the risk of extinction (Ellstrand & Elam 1993, Levin et al. 1996). Given these potential problems, the advent of genetically engineered crops and subsequent questions about possible consequences of transgene flow brought renewed attention to the flow of crop alleles into WW populations.

Stimulated by questions regarding transgene dispersal, dozens of evolutionary geneticists and ecologists took to the field in the 1990s to conduct experimental or descriptive studies addressing the likelihood of such gene flow, usually using nontransgenic plants as model systems. Their initial focus was to determine whether spontaneous hybridization between crops and WW relatives occurred under field conditions. Secondary questions included whether gene flow occurred at distances and rates large enough to permit crop genes (and by extension, transgenes) to enter WW populations and, if so, what consequences were expected to ensue. Subsequent introgression was often assumed, yet introgression-related data, such as the fitness of advanced hybrid generations and the effects of specific crop traits, were largely neglected. The most frequent relevant data from that early research were measurements of the relative fitness of the F_1s versus that of the WW parents (Ellstrand 2003).

Ellstrand (2003) reviewed these and earlier relevant hybridization studies. He examined data for the world's 25 most important domesticated plants in terms of area planted; for 22 of these plants, he found substantial empirical evidence for some spontaneous hybridization with WW relatives somewhere in the world. Hybridization patterns were idiosyncratic for the crops. For some, such as coffee, hybridization apparently occurs rarely and in a few locations. For others, low levels of hybridization typically occur over much of the globe; in the case of wheat and its weedy *Aegilops* relatives, hybridization occurs whenever the species co-occur in temperate regions. For a few, such as cultivated sunflower, hybridization can occur at relatively high rates. Several years later, a book by Andersson & de Vicente (2010) re-evaluated and updated what is known regarding opportunities for crop gene flow to WW relatives for a similar list of 20 important food crops. With rich detail describing the crops' reproductive and dispersal biology, compatible relatives, and hybrid fitness, each case study is presented as a crop-specific guide for transgene flow assessment. Both books (Andersson & de Vicente 2010, Ellstrand 2003) mention introgression when they found supporting data.

Both books focus strongly on hybridization, and hybridization is not introgression. If significant evolutionary or ecological impacts occur, they typically occur through introgression (Arnold 2006). As the significance of introgression became clear, the research focus of the twenty-first century changed considerably, and the research emphasis shifted from "Does hybridization occur?" to the next step of "How and when does introgression occur and at what levels?" The context broadened from focusing on the transgene escape to the introgression of any domesticated alleles (den Nijs et al. 2004). The question of "When and how will introgression have any significant impact?" has also become paramount. Although the introgression of domesticated plant genes has received attention from publications in the context of transgene flow (den Nijs et al. 2004, Kwit et al. 2011, Stewart et al. 2003), the wider burst of research activity that started this century awaits a thorough review.

Thus, we address introgression from domesticated plants to their WW relatives. As mentioned above, such gene flow provides examples of contemporary microevolution that may have important implications. Below we ask, "What do we know about crop gene introgression into WW populations?" including when evolution by introgression creates a negative impact (Natl. Res. Counc. 2002, Organ. Econ. Co-op. Dev. 1993).

Our specific focus is the establishment of alleles from domesticated plants in WW populations. How alleles are naturally introgressed may follow a variety of pathways. The pollen parent in the initial hybridization event (or events) could involve a plant in cultivation, a volunteer left from a previous planting or from seed spillage into a natural population, or a recent escape from cultivation. An alternate evolutionary pathway could start with an uncultivated plant as pollen parent hybridizing with a plant in cultivation. If the hybrid seed from the maternal plant naturally disperses or is harvested and sown in the same location, it can variously self-pollinate, cross with other hybrids, or spontaneously backcross with its parental parent (depending, in part, on the breeding system of each of the parental taxa).

We define wild plants as those capable of growing and reproducing in ecosystems that are largely undisturbed by humans. Weeds are those whose populations persist only under anthropogenic disturbance. Of course, intermediate cases exist; also, some taxa may include both wild and weedy populations.

First, we present a brief overview of hybridization and introgression in plants. Next, we examine various molecular approaches for identifying and studying crop allele introgression. Some case studies of introgression of domesticated alleles into WW populations follow, including the few described examples of spontaneous transgene introgression. We conclude with a look to the future.

HYBRIDIZATION AND INTROGRESSION IN PLANTS

Hybridization is an important component of plant evolution, occurring in roughly 25% of plant species (Baack & Rieseberg 2007). Nonetheless, hybridization occurs unevenly among plant taxa; a higher propensity occurs in certain genera, families, and orders (Whitney et al. 2010). In a few families, natural intergeneric hybridization occasionally occurs (Stace 2010). Even for readily hybridizing species, hybridization and introgression rarely occur at high enough rates to jeopardize the integrity of a species (Levin et al. 1996).

Thus, although hybrids are produced, most of them have no further evolutionary impact. However, sometimes a little hybridization has considerable evolutionarily significance (Arnold 2006). Mechanisms such as alloploidy or apomixis can fix hybridity, leading to the evolution of new hybrid-derived species from one or just a few founders (Arnold 2006). Likewise, homoploid hybrid speciation can yield new species without a change in ploidy level or mating system (Buerkle et al. 2000). More frequently, the evolutionary impact of hybridization is mediated through introgression.

Edgar Anderson's 1949 book *Introgressive Hybridization* championed introgression as a potentially important process for introducing adaptive variation into a population. At that early date, hybrids and hybrid derivatives were largely assigned by morphology. Nowadays, multilocus molecular markers have greatly enhanced the ability to detect introgression (Rieseberg & Wendel 1993). The last major review, twenty years ago, examined hundreds of putative cases of introgression, identifying about three dozen cases (almost all bolstered by molecular data) of probable natural introgression in plants (Rieseberg & Wendel 1993). Today, that number might be larger by an order of magnitude (examples below and in **Table 1**). By definition, introgression cannot be as common as hybridization. Determining how common introgression in plants is and its real evolutionary significance remains challenging and generally requires molecular marker-based methods.

17.4   Ellstrand et al.

Table 1   Examples of introgressed crop alleles in wild or weedy populations with strong support

Cultivated Ancestor	Example of wild or weedy (WW) taxon with one or more populations with introgressed crop alleles	Reference
Beta vulgaris vulgaris, beet	Europe's weed beet (descendant of beet × *B. v. maritima*)	Case study in this review
Brassica napus, oilseed rape	*B. rapa*	Andersson & de Vicente 2010
Cichorium intybus, chicory	WW *C. intybus*	Kiaer et al. 2007
Cynara cardunculus var. *scolymus*, artichoke	Some populations of California's artichoke thistle (descendant of artichoke × *C. c. cardunculus*)	Leak-Garcia et al. 2013
Glycine max, soybean	*G. soja*	Andersson & de Vicente 2010
Gossypium hirsutum, cotton	*G. barbadense*	Andersson & de Vicente 2010
Helianthus annuus, sunflower	*H. petiolaris*	Case study in this review
Lactuca sativa, lettuce	*L. serriola*	Case study in this review
Oryza glaberrima, African domesticated rice	*O. barthii*	Andersson & de Vicente 2010
Oryza sativa, Asian domesticated rice	*O. rufipogon*	Case study in this review
Phaseolus vulgaris, common bean	*P. vulgaris*	Andersson & de Vicente 2010
Raphanus sativus, radish	*R. raphanistrum*	Case study in this review
Solanum tuberosum, potato	*S. edinense* (a stabilized clonal hybrid of potato × *S. demissum*)	Ellstrand 2003
Sorghum bicolor, sorghum	*S. halepense*	Andersson & de Vicente 2010
Triticum turgidum wheat	*Aegilops peregrina*	Kwit et al. 2011
Ulmus pumila, Siberian elm	*U. minor*	Ellstrand 2003
Zea mays mays, maize	*Z. m. mexicana*	Andersson & de Vicente 2010

IDENTIFYING HYBRIDS AND INTROGRESSANTS

Determining whether an individual is a hybrid or has hybrid ancestry is not straightforward in the absence of genetic markers. Phenotypic intermediacy might indicate an early-generation hybrid, a segregant, or a later generation backcross retaining some characters from both parental species. Also, phenotypic response to environmental variation or developmental instability may create unexpected morphologies that mimic hybrid ancestry. Furthermore, experimental research has revealed that, more often than not, for any given morphological or otherwise quantitative trait, F_1s are not necessarily intermediate compared to their parents (Rieseberg & Ellstrand 1993).

For detecting introgression, the problem becomes worse. In his seminal book, Edgar Anderson, believing introgression became increasingly important as the number of introgressed alleles decreased, bemoaned:

> How important is introgressive hybridization? I do not know. One point seems fairly certain; its importance is paradoxical. The more imperceptible introgression becomes, the greater is its biological significance. It may be of the greatest fundamental importance when by our crude methods we can do no more than demonstrate its existence (Anderson 1949).

Co-dominant genetic-based markers can provide more certainty than phenotypic measurements. For hybridizing taxa fixed for the alternate alleles, *a1* and *a2*, if an individual is heterozygous for both alleles, the individual can be assigned hybrid ancestry. But information from a single locus is insufficient to distinguish an introgressant from a first generation hybrid.

When few individuals are surveyed, it is often uncertain whether both parental taxa are actually fully fixed for alternate alleles. If not, an allele might be wrongly assigned as introgressed. To illustrate, imagine both putative parents are not fixed for alternative alleles but have retained alleles from a common ancestor (symplesiomorphy or incomplete lineage sorting). One species has the *a1* allele at 97% frequency and the *a2* allele at 3%; the other has *a1* at 1% and *a2* at 99%. If an individual is found that is heterozygous at that locus, does it have hybrid ancestry or is it just one of the rare heterozygotes within either species? Thus, thorough sampling of putative parental taxa is crucial, especially when few loci are assayed. Populations that have never been in reproductive contact with crops are particularly useful but are difficult to identify for widespread ancient crops whose distributions have changed. The accuracy of assigning hybrid ancestry and introgression increases for each additional independent marker assayed and as more data are available from other sources of information.

Alternatively, transgenes, the result of genetic engineering, are evolutionarily unique to crops. If we find an engineered gene in a WW plant, we can be certain that it has hybrid ancestry. At the moment, evidence for transgene introgression is known from only a few cases (Warwick et al. 2008, Wegier et al. 2011).

In most cases, the ideal tool kit is a set of genetically based markers for large numbers of loci dispersed throughout the whole genome that can be easily and unambiguously scored. The most useful markers are codominant, allowing for homozygous loci to be distinguished from heterozygous loci. Furthermore, access to a set of anonymous loci with no a priori expectations of selection history can provide a baseline expectation of patterns of introgression to compare with a locus of interest. However, fully dominant molecular markers and/or markers whose genetic basis is unclear have limited value in elucidating ancestry (Whitkus et al. 1994).

Allozymes were the first widely used biochemical marker system, an improvement over morphological markers, but they were constrained by the low number of loci sampled. For a given taxon, polymorphic allozyme loci rarely number more than a dozen, each with two or a few more alleles. DNA-based markers are much more powerful. For example, polymorphic microsatellite loci may have several to dozens of alleles per locus. Their primary limitation is their relatively high back-mutation rate, resulting in alleles with different histories appearing to be identical (homoplasy). Microsatellite mutation rate should be less problematic for crop allele introgression research when such introgression has occurred recently in evolutionary time, in many cases, less than a century (Ellstrand et al. 2010).

Single nucleotide polymorphisms (SNPs) are emerging as highly informative markers for characterizing introgression. SNPs can be confidently scored and provide high throughput analysis; SNP microarrays can survey polymorphism at up to tens of thousands of sites across a plant's genome and are now available at relatively affordable prices. Alternative methods using reduced genomic libraries include RADtag sequencing, which might also be useful for larger genomes (Baird et al. 2008). Also, the genotyping-by-sequencing approach allows direct evaluation of large, complex (e.g., polyploid) genomes without prior development of other molecular tools (Elshire et al. 2011). Although such an approach is feasible for any species, currently genome assembly remains labor-intensive and costly for many nonmodel species, especially those with large genomes.

Direct SNP genotyping as well as single molecule sequencing are advancing rapidly, and cheaper flexible genotyping methods are becoming available (Maughan et al. 2011). Even sequencing of polyploid plants is within reach because of single molecule sequencing (Kovalic et al. 2012), yielding allele dosage and haplotype information.

With access to a large number of loci of known genome location, variation in introgression rates throughout the genome can be estimated. Genome-level approaches have already revealed that introgression occurs unevenly across the genome of the species involved

17.6   Ellstrand et al.

(Baack & Rieseberg 2007, Hooftman et al. 2011, Twyford & Ennos 2012). Information on linkage disequilibrium (LD) between markers can aid in making inferences about introgression because, in cases of recent introgression, tightly-linked taxon-specific markers would still be in LD. LD should decrease in later generation hybrids, and is expected to be very low if similarity between taxa resulted from a common but ancient ancestry—unless selection acting on allele combinations is extremely strong (e.g., Linder et al. 1998). Linkage information can be obtained by crossing the taxa under study and analyzing segregation in the F_2 or BC_1 generation. Because most important crops have been already extensively genetically mapped, new markers can easily be placed on that map and tested for LD. In particular, haplotype data are ideal for reconstruction of introgression as they provide information about linkage relationships, that is, whether genes are in coupling or repulsion phase, which in turn gives information on the ancestral genotypes.

The most detailed genetic information regarding loci and linkage is obtained via genome sequencing. One example of detection of introgression through sequencing is the discovery of Neanderthal DNA in modern humans (Green et al. 2010) revealed by next-generation sequencing techniques combined with what is known regarding the migration patterns of modern humans.

MEASURING INTROGRESSION

Once introgression has been identified, more questions follow: What fraction of the recipient population is introgressed? Are introgression rates roughly equivalent across the genome or are some loci or genomic regions significantly more or less introgressed relative to neutral expectations; that is, does selection appear to be involved in which loci have passed from one taxon to the next? As the number of marker loci increases, so too will the ability to answer these questions, especially if the genomic location of those loci is known.

Another important factor is the timescale over which introgression has occurred. How many years or generations of introgression have occurred? For example, for transgenes the timescale of introgression into unmanaged populations is necessarily recent, no longer than since the first field release of the transgenic crop, obtainable from public regulatory records. However, the detection methodology depends on publically available transgenic sequences. Other kinds of historic information may be useful for establishing time since contact between potentially hybridizing data—ranging from newspaper accounts to archeological artifacts (e.g., Leak-Garcia et al. 2013).

The combination of historic information and genetic data can be used to infer the amount of gene exchange between populations. Traditional population genetic methods for estimating genetic differentiation, such as F statistics (Wright 1931) and their analogs, have often been used to estimate levels of gene flow. These methods measure genetic differentiation, which can be used to estimate historic gene flow (average number of successful immigrants per generation) but make some potentially biologically unrealistic assumptions (Whitlock & McCauley 1999).

Recently, methods have been developed for assigning individuals into groups based on patterns of linkage equilibrium or genetic differentiation (e.g., Corander et al. 2004, Pritchard et al. 2000). Such methods can identify admixed individuals (those whose genetic content is derived from more than one of the inferred groups). Such admixture implies recent introgression. Another method, BAYESASS (Wilson & Rannala 2003), explicitly estimates population-level introgression rates rather than individual migrant ancestries and does not assume Hardy-Weinberg equilibrium (often the case with other methods). BAYESASS and related approaches are appropriate for detecting introgression that has occurred relatively recently (<20 generations ago). The relatively new (at least to population genetics) method of Approximate Bayesian Computation (Beaumont 2010) has been successfully applied in crop-wild introgression during domestication (e.g., Ross-Ibarra et al. 2009) and has potential regarding more recent crop-WW questions. This method is versatile, given

the range of evolutionary simulation programs now available (partially reviewed by Hoban et al. 2012), with relatively few limitations for demographic scenarios that can be accommodated (e.g., some cases of asymmetrical gene flow). However, it requires large amounts of data, especially for complex scenarios. Just how far into the past these methods are informative needs to be addressed with additional simulation or controlled experimental work.

The foregoing methods are often applied to microsatellite data. As discussed above, the high mutability of microsatellites can result in homoplasy, a problem for estimating long-term introgression. Questions of long-term introgression measurement are better addressed with slowly evolving markers such as SNPs or longer sequenced regions. Therefore, for longer-term introgression estimates, coalescent-based methods (Kuhner 2009) may be appropriate. Two widely used programs implementing these methods are MIGRATE (Beerli 2006) and the isolation with migration (IM) suite of programs (Hey 2010). These methods explicitly take into account genealogical relationships between alleles, enabling estimates of both directional introgression rates and other demographic parameters such as divergence time and population size (Kuhner 2009). MIGRATE assumes an equilibrium scenario in which population sizes and migration rates have been stable for a long time relative to initial divergence. In contrast, the IM programs explicitly model a nonequilibrium scenario in which divergence occurred in the relatively recent past, allowing for the disentangling of genetic similarity due to gene flow versus shared retention of ancestral polymorphisms. Consequently, IM is likely more appropriate for many crop-WW systems. However, the current methods do not efficiently estimate the timing of introgression (Strasburg & Rieseberg 2011); additional relevant research would be valuable.

A "genomic clines" method for examining genomic patterns of introgression for both short- and long-term timescales has recently been described (Gompert & Buerkle 2009, 2011). This method compares introgression patterns at individual loci relative to the genomic background for detecting selection that affects introgression rates. Simulations suggest that this method can be informative on timescales as short as five generations, making it applicable to many crop-WW systems.

CROP ALLELE INTROGRESSION INTO WW POPULATIONS

Crops and their WW relatives present both disadvantages and advantages for detecting introgression. Most crops began their evolutionary journey hundreds to thousands of years ago, leaving little time for substantial genome divergence. However, the plant improvement process may have fixed alleles that are nearly absent in WW populations because they are detrimental under natural conditions. Undesirable traits eliminated from crops include seed dormancy, seed shattering in seed crops, and early bolting in vegetable crops (e.g., Hartman et al. 2013a, Weeden 2007). Bottlenecks from strong selection under domestication vary widely in both intensity and duration (Gross & Olsen 2010); stronger bottlenecks generally promote higher levels of differentiation between domesticates and their wild ancestors. Furthermore, plant breeders have occasionally introgressed wild germplasm into a crop, resulting in decreased genetic differentiation between the taxa. Such deliberate wild-to-crop introgression might be difficult to distinguish from introgression in the opposite direction, but methods exist for detecting asymmetric patterns of gene flow (Beerli 2006, Hey 2010). In contrast, transgenesis and mutation breeding create evolutionarily unique single locus (for transgenics) and multiple locus (for mutation breeding) crop-specific markers.

Crops, as part of the study system, provide numerous benefits because of their economic importance. They are well studied; for example, at the moment, substantial genome sequence data are available for 49 plant species, 35 of which are crops. (Michael & Jackson 2013). Dozens of crop species not yet sequenced have been extensively genetically mapped with much transcriptome data available (e.g., Bowers et al. 2012). Crops have the advantage that historical and geographic

information are often available—e.g., cultivars and the duration of cultivation in a given location. Crop genetic resources are readily available, and pure material of specific crop cultivars involved in hybridization can be obtained for screening. Moreover, extensive germplasm collections of old landraces and wild accessions are available for many crops.

Consequently, reports of domesticated allele introgression into WW populations have gradually accumulated. **Table 1** features a nonexhaustive list of crop species for which introgression is known to have occurred or is occurring on the basis of substantial data. We provide some case studies below as examples. These represent an array of crops from different plant families with different uses, breeding systems, and life histories. None of the following crops, except for beet, have genetically engineered cultivars that are commercially available.

CASE STUDIES

Beet, *Beta vulgaris vulgaris*

Beet is mainly wind pollinated and predominantly outcrossing; cross pollination on the order of a kilometer is not unusual (Bartsch 2010). Spontaneous hybridization easily occurs between the crop *B. v. vulgaris* (both root beet and sugar beet) and its wild progenitor, *Beta vulgaris maritima* (sea beet), as documented by both descriptive and experimental studies (reviewed by Bartsch 2010, Ellstrand 2003). Much of the research was first stimulated by the evolution of hybrid-derived annual weed beets in Europe's biennial sugar beet fields. Because these bolting individuals have a woody root, their presence often substantially reduces yields and damages the machinery used in harvesting and processing sugar beets. Several studies using suites of genetically based morphological and molecular markers revealed that the initial weed beets were natural F_1 hybrids between sea beets and sugar beets used as maternal parents for commercial seed production. If weed beets are not eradicated prior to seed dispersal, their descendants evolve quickly via crossing with one another into stably introgressed weed beet populations well adapted to flourish in sugar beet fields (Arnaud et al. 2010, Bartsch 2010). Weed beet seeds can survive in the soil for several years. Their seedlings and young plants are indistinguishable from sugar beet. They tolerate all herbicides suitable for sugar beet cropping. Thus, weed beet is hard to control without long crop rotation intervals.

Introgression in the other direction (from the crop to the WW relatives) attracted attention when concerns arose about the environmental risks associated with transgenic sugar beets. Descriptive studies with molecular markers demonstrated that Italian sea beet populations growing within a few kilometers of beet seed multiplication sites were introgressed with domesticated alleles. Additionally, populations of the more distantly related *Beta macrocarpa* in southern California that were sympatric with sugar beet fields were found to have a low frequency of sugar beet–specific alleles. In both cases, about 100 years of gene flow has increased allelic diversity of the WW populations. Thus, for these taxa, crop-to-weed introgression has already delivered novel crop alleles to WW populations and would be expected to do so for transgenes as well (Bartsch 2010, Ellstrand 2003). Transgenic herbicide-resistant sugar beet is currently commercially grown in the United States (James 2012). Without extraordinary measures to prevent gene flow, transgene introgression in the long run is likely if beets are grown close to compatible WW congeners, such as those found in California's Imperial Valley.

Carrot, *Daucus carota sativus*

Both carrot and its WW progenitor, *Daucus carota carota*, are predominantly outcrossing and capable of insect-pollinated at more than four kilometers (Rong et al. 2010). Spontaneous

hybridization between cultivated carrot and its WW subspecies is well known to carrot industry seed scientists because pollen from WW relatives adjacent to cultivar seed multiplication fields often contaminates commercial seed, resulting in the germination of individuals with intermediate morphologies. Plants with similar hybrid-like morphologies have also been found in the adjacent WW populations (Wijnheimer et al. 1989).

Evidence for crop-WW introgression in this system has been demonstrated only recently. Magnussen & Hauser (2007) used amplified fragment length polymorphisms to genetically analyze 71 *D. c. carota* plants growing within, adjacent to, and quite distant from cultivated fields. Four individuals (5.6%) were found to have hybrid ancestry, but did not have genotypic patterns consistent with F_1 hybrids; they had to be F_2s, backcrosses, or more advanced introgressants. They also found that wild carrot populations in close proximity to cultivated carrot fields were genetically more similar to cultivated carrot than wild carrot populations located far from cultivated carrot fields, suggesting introgression had changed the genetic structure of the WW populations.

In a three-year field experiment, Hauser & Shim (2007) showed that F_1s between cultivated and wild carrots could survive and reproduce outside cultivation, although with somewhat lower lifetime fitness overall. Fitness varied with environment and was not so low as to impede the hybrids from crossing and backcrossing. The fitness of crop-wild offspring beyond first generation hybrids has not been reported yet, and the dynamics of domesticated alleles in wild carrot populations remain unknown.

Sunflower, *Helianthus annuus*

Domesticated in North America, cultivated sunflower is well known to naturally hybridize with some WW relatives (Ellstrand 2003). Wild *Helianthus* species are self-incompatible and insect-pollinated; the crop is self-fertile but not highly selfing. The system is well studied with regard to crop-to-WW hybridization and introgression. Numerous descriptive and experimental studies using a variety of molecular and morphological markers have demonstrated natural hybridization between the crop and WW *H. annuus* as well as *Helianthus petiolaris* (reviewed by Ellstrand 2003). Not only has spontaneous hybridization been documented repeatedly in North America (e.g., Linder et al. 1998), the center of *Helianthus* diversity, but it has also been documented in regions where the crop and its WW relatives have been introduced, such as Argentina (Gutierrez et al. 2010).

Experimental fitness comparisons between wild *H. annuus* and crop-wild hybrids often show an initial F_1 fitness penalty, which varies among genotypes and environments, declining over subsequent generations of introgression (Gutierrez et al. 2011, Mercer et al. 2007, Presotto et al. 2012). Such field-based experiments have also shown that some genetically based crop traits are detrimental, whereas others, such as larger seed size, early competitive advantage, and insect resistance, have a selective advantage that could facilitate their introgression into WW populations (Baack et al. 2008, Dechaine et al. 2009, Snow et al. 2003). Interestingly, Burke & Rieseberg (2003) found that a transgene that confers disease resistance in the crop did not boost the fitness of crop × WW hybrids in the presence of the disease organism.

Several studies have demonstrated that cultivated sunflower alleles have introgressed into WW populations. For example, in a five-year study, Whitton et al. (1997) found crop-specific random amplified polymorphic DNA markers continuing to spread spatially through a WW *H. annuus* population after a single season of natural hybridization with an adjacent crop population. Likewise, interspecific introgression, in particular into *H. petiolaris*, has been documented with suites of molecular markers (e.g., Gutierrez et al. 2010, Rieseberg et al. 1999). With known significant introgression into WW populations, so much fitness data, and the development of resources

(such as the 10,000 locus *Helianthus* genetic map) (Bowers et al. 2012), *Helianthus* is an important model system for studying how introgression proceeds and the role it plays in evolution.

Lettuce, *Lactuca sativa*

Although both cultivated lettuce and the cross-compatible WW *Lactuca serriola* are predominantly selfing, substantial cross-pollination by insects has been observed. Given those data, the fact that both species often grow in close proximity, and their flowering periods overlap for weeks, experimental field studies have been conducted to measure crop-to-WW outcrossing rates. Typical hybridization rates are on the order of a few percent or less, but individual plants at short distance may bear as much as 26% hybrid seed (D'Andrea et al. 2009).

Uwimana et al. (2012a) genotyped a wide range of *Lactuca* germplasm with ten microsatellite loci: almost eight thousand individual samples of European *L. sativa* and *L. serriola* accessions. Genetic admixture analysis by the model-based clustering program *structure* (Pritchard et al. 2000) revealed two clusters and 7% introgressed plants. The Bayesian NewHybrids program (Anderson & Thompson 2002) revealed that some of these introgressants were the result of several generations of selfing from either an F_1 or a BC_1 ancestor, consistent with how introgression should proceed for such highly selfing plants.

Hooftman and coworkers conducted a set of field experiments to assess the effects of hybrid ancestry on fitness in *Lactuca* allowing for natural levels of competition. They found hybrid-derived plants (via both selfing and backcrossing) outperforming the parental species for at least up to four generations (Hooftman et al. 2007). Second generation hybrids, when compared to the wild relative, had twice the relative fitness based on seed-output-per-seed-sown. Differences in germination and survivorship were found for all generations; differences in seed output were found only for the first two generations (Hooftman et al. 2005). The degree of the fitness difference decreased over generations (Hooftman et al. 2007), suggesting heterosis was the underlying cause of the initial fitness increase. However, transgressive segregation occurred as well: Several of the selfing lines continuously outperformed the wild species in a subsequent experiment (Hooftman et al. 2009). Such patterns suggest the possible evolution of adaptive gene combinations. The same pattern was shown independently by Hartman et al. (2013b) using recombinant inbred lines from a cross between a Californian *L. serriola* accession and a *L. sativa* cultivar.

Molecular evidence also supports the hypothesis that selfing has fixed *Lactuca* adaptive transgressive gene combinations. Using those plants that survived until flowering from their multiyear study, Hooftman et al. (2011) demonstrated that various parts of the hybrid-derived genomes were skewed toward one or the other parent species. Quantitative trait loci studies using the best performing lines revealed an adaptive combination of alleles from both parental species at unlinked genomic locations. Similar results were obtained from other recent studies of lines derived from *L. sativa* × *L. serriola* hybrids (Hartman et al. 2012, 2013b; Uwimana et al. 2012b).

Asian Cultivated Rice, *Oryza sativa sativa*

Despite the crop's high selfing rate, natural hybridization occurs between cultivated rice and two of its close WW relatives. One is cultivated rice's ancestor, *Oryza rufipogon*. The other is weedy rice; *Oryza sativa* f. *spontanea*, a noxious weed of cultivated rice that closely mimics the crop but disperses (shatters) its seed before it can be harvested. These taxa are also self-compatible but often have somewhat higher outcrossing rates than cultivated rice (Andersson & de Vicente 2010, Lu & Snow 2005). Like most grasses, wind disperses pollen in domesticated rice and its WW relatives. Support for spontaneous hybridization comes from descriptive studies, showing that hybrid

swarms typically occur when the crop grows in close proximity with its WW relatives. Likewise, experimental field studies have measured crop-to-WW *Oryza* hybridization rates. At multimeter distances, hybridization rates are typically <1% (Lu & Yang 2009). However, hybridization rates between intermixed cultivated rice and *O. rufipogon* can be as high as 18% (Wang et al. 2006).

Studies employing molecular markers have demonstrated crop allele introgression into some sympatric and peripatric populations of both the wild ancestor and the derived weedy rice (dos Reis Goulart et al. 2012, Song et al. 2006, Xia et al. 2011). Furthermore, genetic analysis has confirmed some weedy rice populations are descendants of *O. sativa* × *O. rufipogon* hybrids, whereas others evolved directly from feral cultivated rice (Ellstrand et al. 2010).

Numerous studies have evaluated the relative fitness of *Oryza* hybrids and introgressants, with and without transgenes. Given the polygenic dominance of shattering, direct comparison of hybrids with their WW parents in a common garden experiment is reasonable. F_1s of the crop and wild rice generally have some reduced sexual fitness relative to the pure wild parent, but have increased vegetative tillering (Song et al. 2004); however, F_1s of the crop and weedy rice generally have about the same fitness as pure weedy rice (Lu & Yang 2009).

Lu and colleagues have started conducting experimental studies to examine the fitness consequences of transgenic insect resistance introgressed into WW rice. Their limited research on hybrids between cultivated rice and *O. rufipogon* and their descendants has shown little significant effect of the crop transgene (either *Bt* or *CpTI*, which confer resistance to insects) on seed dormancy (Dong et al. 2011). With regard to hybrids and introgressants involving weedy rice, the results have been complex, depending on both genotype and environment. Transgenic hybrids and progeny through F_3 showed increased fecundity under natural insect pressure compared with their nontransgenic counterparts, but the fecundity boost disappeared under low insect pressure (Yang et al. 2011). Transgenic rice has been deregulated in the United States, Iran, and China but, to our knowledge, it is has not yet been grown commercially anywhere.

Nonetheless, considering the results of the following recent study, favorable transgenes are likely to flow easily from rice into nearby populations of weedy rice, as well as into populations of the wild ancestor, where crop gene flow has already been shown to occur at higher rates than in the weed. Nontransgenic rice with herbicide resistance to imidazolinone has been introduced worldwide. The evolution of imidazolinone resistance in companion weedy rice populations has occurred in several locations. But was the evolution of resistance due to spontaneous mutation or introgression of the resistance allele from the crop? Researchers in Brazil (dos Reis Goulart et al. 2012) addressed this question by genetically analyzing thousands of resistant weedy rice plants from dozens of populations for the resistance mutations known in the cultivars as well as for cultivar-specific microsatellite markers. Almost 99% of the herbicide-resistant plants had evolved via introgression of the crop allele; 1.1% evolved as a result of spontaneous mutation.

Radish, *Raphanus sativus*

Radish and its WW relative *Raphanus raphanistrum* are insect-pollinated and self-incompatible annuals (Snow & Campbell 2005). The two often naturally hybridize where these taxa co-occur, often forming localized hybrid swarms (Stace 1975). Surprisingly, these spontaneous hybrid populations have attracted little attention from researchers (Snow & Campbell 2005).

However, in California, bilateral introgression between the two introduced taxa over the past century or so has been so extensive that the species have coalesced into a widespread set of hybrid-derived populations, as demonstrated by both morphological and molecular analyses (reviewed by Ellstrand et al. 2010). Although nowhere near as problematic as Europe's weed beet, "California

17.12   Ellstrand et al.

wild radish" is classified as both an agricultural weed and as an invasive plant. The molecular evidence supports a history of multiple hybridizations that have contributed to what is now a more-or-less evolutionary cohesive metapopulation.

Despite the relative uniformity of molecular markers, field experiments have shown that these new populations have undergone adaptive evolution (reviewed by Ellstrand et al. 2010). California wild radish has a higher maternal fitness than either progenitor in several environments, producing more seeds per plant. Another set of reciprocal transplant experiments involving plants from northern and southern sites revealed the evolution of local adaptation in probably less than 100 generations.

Snow and colleagues (Campbell et al. 2006, Hovick et al. 2012, Snow et al. 2010) conducted some long-term experiments to study the consequences of crop allele introgression into naturalized populations of *R. raphanistrum*. Under natural field conditions, they created four pure *R. raphanistrum* populations and four populations of crop-WW F_1s, allowing these annual plants to evolve in Michigan year by year. They harvested seeds from both types of populations after four generations, planting them in common gardens to compare fitness. Fitness of the hybrid-derived lineages had recovered in a couple of generations. The advanced generation hybrids had slightly lower lifetime fecundity in than WW *R. raphanistrum*. But when Michigan-evolved hybrids and their WW parent were grown in southern Californian (Campbell et al. 2006) and southeastern Texan (Hovick et al. 2012) common gardens, the fecundity of the hybrid lineage was roughly triple that of the pure parent.

The Michigan field experiment allowed for monitoring of some unlinked crop-specific alleles for several years. Each showed a different evolutionary trajectory over ten years. One crop trait, a dominant white flower color allele, was found to be present 14 years after the experiment's start (Snow et al. 2010).

THEORETICAL STUDIES

Motivated by recent interest in transgene flow, computational and mathematical research on introgression has been steadily increasing. Such theoretical studies on transgene introgression are also usually valid for general crop allele introgression. Typically, theoretical studies concern two topics: identifying factors affecting introgression and transgenic plant risk assessment. Sometimes the same analysis considers both, but it is instructive to address them separately.

Factors Affecting Introgression

The dynamics of introgression depend on evolutionary (e.g., gene flow, positive selection, negative selection, variation in selection), spatial (e.g., metapopulation structure), ecological (e.g., environment-dependent heterosis), genetic (e.g., ploidy differences between hybridizing taxa, linkage of the allele in question to other genes under selection), and many other factors. Space limitations preclude an exhaustive discussion of them here. Theoretical approaches have proven well suited for evaluating the relative importance of what are considered some of the most important factors.

General population genetics theory provides some simple predictions for introgression (summarized by Ellstrand 2003). If a modest amount of gene flow occurs without opposing selection (e.g., one successful immigrant per generation), immigrant alleles can establish in the recipient population. Under opposing selection, the immigrant alleles eventually disappear after a single gene flow incident. But if gene flow reoccurs uniformly over generations with opposing selection, immigrant alleles may be maintained in immigration-selection equilibrium.

Selection pressure and gene flow often vary in space and time. Kuparinen & Schurr (2007) developed a spatially explicit model allowing for spatial and temporal heterogeneities in gene flow and selection. Furthermore, they considered different modes of crop transgene deployment and expression. They found that if crop plants are heterozygous for a transgene enhancing the fitness of WW relatives, introgression levels would be reduced because, compared with homozygous plants, immigration pressure is halved. Using recessive transgenes (as opposed to dominant or additive ones) limits the spread of advantageous genes because positive selection on recessive genes is very weak when they are rare. Therefore, unless gene flow is high and/or recurrent, recessive transgenes are more likely to be lost through genetic drift.

A transgene's location in the genome can also affect introgression. Inserting a transgene next to a domestication gene that is deleterious in WW populations could drastically reduce introgression (Gressel 1999, Stewart et al. 2003). Ghosh et al. (2012a) used both mathematical and simulation approaches to address mitigation by linkage and found it to be a viable mitigation strategy. They defined the hazard rate as the probability per unit time that a permanent transgene lineage is formed, given that one has not previously been formed. Small WW populations yielded high hazard rates owing to the force of repeated hybridization. At intermediate population sizes, hazard rates were lower because low-frequency crop alleles are removed by drift. Selection was only of importance in large populations, leading to an increase in the hazard rate.

A WW population receives most gene flow from a crop when intermixed with or adjacent to the crop field. Such WW populations are rarely isolated, but often form a part of a larger WW metapopulation. Some recent models examined introgression in the context of gene flow between metapopulation demes in concert with their extinction-recolonization. One such study (Meirmans et al. 2008) found that even very low levels of introgression and positive selection can lead to the fixation of a transgene in the metapopulation. Thus, metapopulation dynamics can play an important role in promoting introgression.

Quantitative Risk Assessment

Traditional transgene risk assessment defines risk as a function of exposure and hazard ("an act or phenomenon that has the potential to produce harm or other undesirable consequences to humans or what they value"; Natl. Res. Counc. 1996, p. 215), and transgene introgression involves the "exposure" component of that analysis. (Note the definition of hazard is quite different than the "hazard rate" defined earlier in this article.) Introgression is not a hazard in itself. Introgression has been largely associated with possible environmental hazards of depletion of genetic diversity, evolution of weediness/invasiveness, and extinction. The hazards of introgression are not necessarily restricted to environmental effects (de Jong & Rong 2013). For example, introgression has been associated with various culturally identified hazards, such as the altered genetic makeup of wild plants sacred to indigenous peoples (e.g., Walker & Doerfler 2009). These hazards are not unique to transgene flow but could be associated with gene flow from any crop to a wild relative and have, on rare occasion, occurred (Ellstrand 2003).

Therefore, a complete transgene introgression risk assessment should comprise the probability of introgression as well as its consequences (Hill 2005) and how it differs from nontransgenic introgression. Although natural crop-WW hybridization occurs for most crop-WW relative combinations, it is generally rare (Ellstrand 2003, Stewart et al. 2003). Given a low hybridization rate and the fact that the size and lifespans of recipient WW populations are often limited, the introgression process is inherently strongly stochastic, driven by both selection and genetic drift. Thus, stochastic approaches can supplement the long-used deterministic models to investigate gene flow and introgression in transgene risk assessment.

Classical population genetic models (e.g., Kimura 1962) predict that drift often prevents the fixation of a gene that is present in only a few copies in a large population, even if it is selectively favored. Consequently, if hybridization occurs once and at a low rate, introgression may fail to occur. However, a crop is rarely grown once near a WW population; repeated crop-to-WW gene flow typically occurs. If gene flow is recurrent and the interaction period is indefinitely long, then introgression is inevitable for neutral and favorable alleles.

Inevitable does not necessarily mean immediate, because it might still take a very long time before an allele permanently invades a population. Whether a transgene goes to fixation after hybridization stops at a given time is relevant to risk assessment, because fixation represents a permanent evolutionary change to the population in question. Ghosh & Haccou (2010) quantified the risk of introgression based on probability of fixation, proposing that the hazard rate (defined above) is a suitable measure of the exposure component of risk. Ghosh et al. (2012b) considered effects of periodical or permanent interruption of gene flow on the hazard rate.

Because the introgression process depends on so many factors, a comprehensive model to quantify introgression probability will necessarily be detailed, especially when considering a metapopulation scale. Wilkinson et al. (2003) did this by using a combination of modeling and remote sensing to estimate the number of *Brassica rapa* × *Brassica napus* hybrids formed annually across the entire United Kingdom. DiFazio et al. (2012) used a smaller spatial scale, basing their model on extensive geographic information systems information including topology and land-use information. They parameterized their model using data from experiments with their model system, *Populus trichocarpa*, measuring pollen flow, seed flow, and establishment. With this model, they were able to generate very detailed predictions about transgene escape. However, as a consequence of the size of the model, it could only be run for a limited number of generations and therefore could not be used to quantify the eventual fate of the transgene in the population, that is, its probability of reaching fixation. Such detailed models are large, with long run times. With so many different parameters, each of which has to be estimated from experimental or observational data, quantifying the uncertainty in the estimated risks is challenging. Such "uncertainty about uncertainty" makes interpretation of the results difficult, especially for those policy makers who desire clear-cut answers. Of course, sensitivity analyses can help to distinguish which parameters are important versus those that contribute little to the results. Furthermore, limiting a model to a small set of realistic usage scenarios often makes an exhaustive exploration of the parameter space unnecessary. DiFazio et al. (2012) combined these two approaches and compared two different usage scenarios: a small field trial and a large-scale plantation. Their sensitivity analyses indicated that increasing seed and vegetative dispersal led to a proportionally larger increase in transgene spread in the plantation than in the field trial. Their results suggest that at different scales of employment of the transgenic crop, different factors determine the introgression risk.

CONCLUSIONS AND A LOOK TO THE FUTURE

A few decades ago the idea of crops spontaneously hybridizing with their WW relatives was controversial. That question was resolved with data demonstrating that such spontaneous hybridization occurs for most crops, but with varying frequency depending on many factors. The controversy has since moved on to the extent and nature of crop allele introgression into WW populations and its consequences. As discussed above, techniques for collecting data and analyzing data regarding introgression have made a quantum leap since the last major review of introgression in plants two decades ago (Rieseberg & Wendel 1993).

Consequently, strong evidence of spontaneous crop allele introgression has now been found for several crops (**Table 1**). Our case studies reveal that such evidence varies over crops. For some

cases, simply the fact of introgression is known, in others, it is clear that introgression played a key role in the evolution of a more problematic weed (e.g., transfer of herbicide resistance in rice) or even a new noxious weed (e.g., weed beet). Experimental work on post-F_1 fitness with and without transgenes is becoming more common. Such data will provide information to feed emerging demographic models for predicting the progress of introgression (e.g., see lettuce case study).

Although concern about transgene introgression fueled interest in crop-to-WW introgression research and millions of hectares have been planted to transgenic crops for more than two decades, examples of spontaneous transgenic crop × WW hybridization remain exceedingly few. The discovery of transgenes in WW populations is primarily due to the fact that the vast bulk of transgenic commercial plants are restricted to four crop species (maize, cotton, oilseed rape, and soybean) generally planted far (hundreds of kilometers) from their WW relatives.

Transgenes in WW populations are known from three cases. Evidence for hybridization is clear for those cases but not for introgression.

Transgenic oilseed rape has been commercially grown in Quebec, where it overlaps with the range of the closely related weed, *Brassica rapa*. Transgenic hybrids were identified in two separate locations in Quebec. The oilseed rape crop is no longer grown at either site. The *Brassica* populations have been monitored, and the transgene appears to be decreasing in frequency. One individual was identified as a putative introgressant based on joint cytogenetic data and the herbicide-resistance phenotype of the transgene (Warwick et al. 2008).

In 2003, pollen and seed unintentionally dispersed from field trials of transgenic cultivated creeping bentgrass. Transgenic F_1s with weedy creeping bentgrass were subsequently identified. No post-F_1 transgenic bentgrass has yet been reported, but it is not clear that such plants have yet been sought (Reichman et al. 2006).

Transgenic cotton is grown in Mexico, but generally far from WW cotton populations. However, after removing the lint, cottonseed is transported in open vehicles throughout the country. Wegier et al. (2011) used immunoassays to test for transgenes in cotton seeds collected from apparently wild Mexican *Gossypium hirsutum* plants, that is, those whose morphology showed no evidence of introgression from the crop. F_1s are typically intermediate to the crop and the wild plant. Plants from four populations tested positive for one or more transgenes, suggesting they are not hybrids but introgressants. More study of all three systems would be helpful to establish whether introgression is occurring or the establishment of transgenes is transient.

The study of introgression of crop alleles into WW populations is a healthy field. It has come a long way and is poised for future growth to answer both basic and applied questions at the interface of evolution and ecology. Although interest in crop-WW gene flow is rapidly growing, the same can be said for studies of plant introgression that do not involve crops at all. Space limitations preclude a comprehensive review. But some approaches do not involve expensive and sophisticated genetic analysis. For example, *Helianthus annuus* ssp. *texanus* is a hybrid derivative of wild *H. annuus* ssp. *annuus* and *H. debilis*. By growing these three taxa, synthetic hybrids, and backcrosses in two common gardens and measuring how fitness covaried with herbivore pressure and ecophysiological, phenological, and architectural traits, Whitney and colleagues (2006, 2010) were able to identify the adaptive introgression of specific abiotic and biotic traits from *H. debilis* into *H. annuus*. Obviously, such research can be a model for future work on crop-WW introgression.

DISCLOSURE STATEMENT

ACKNOWLEDGMENTS

LITERATURE CITED

Anderson E. 1949. *Introgressive Hybridization*. New York: Wiley

Anderson EC, Thompson EA. 2002. A model-based method for identifying species hybrids using multilocus data. *Genetics* 160:1217–29

Andersson MS, de Vicente MC. 2010. *Gene Flow between Crops and Their Wild Relatives*. Baltimore: Johns Hopkins Univ. Press

Arnaud JF, Fenart S, Cordellier M, Cuguen J. 2010. Populations of weedy crop-wild hybrid beets show contrasting variation in mating system and population genetic structure. *Evol. Appl.* 3:305–18

Arnold M. 2006. *Evolution through Genetic Exchange*. Oxford, UK: Oxford Univ. Press

Baack EJ, Rieseberg LH. 2007. A genomic view of introgression and hybrid speciation. *Curr. Opin. Genet. Dev.* 17:513–18

Baack EJ, Sapir Y, Chapman MA, Burke JM, Rieseberg LH. 2008. Selection on domestication traits and quantitative trait loci in crop-wild sunflower hybrids. *Mol. Ecol.* 17:666–77

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, et al. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3(10):e3376

Bartsch D. 2010. Gene flow in sugar beet. *Sugar Tech.* 12:201–6

Beaumont MA. 2010. Approximate Bayesian computation in evolution and ecology. *Annu. Rev. Ecol. Evol. Syst.* 41:379–406

Beerli P. 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* 22:341–45

Bowers JE, Bachlava E, Brunick RL, Rieseberg LH, Knapp SJ, Burke JM. 2012. Development of a 10,000 locus genetic map of the sunflower genome based on multiple crosses. *G3* 2:721–29

Buerkle C, Morris R, Asmussen M, Rieseberg L. 2000. The likelihood of homoploid hybrid speciation. *Heredity* 84:441–51

Burke JM, Rieseberg LH. 2003. Fitness effects of transgenic disease resistance in sunflowers. *Science* 300:1250

Campbell LG, Snow AA, Ridley CE. 2006. Weed evolution after crop gene introgression: greater survival and fecundity of hybrids in a new environment. *Ecol. Lett.* 11:1198–209

Corander J, Waldmann P, Marttinen P, Sillanpaa MJ. 2004. BAPS 2: enhanced possibilities for the analysis of genetic population structure. *Bioinformatics* 20:2363–69

D'Andrea L, Felber F, Guadagnuolo R. 2009. Hybridization rates between lettuce (*Lactuca sativa*) and its wild relative (*L. serriola*) under field conditions. *Environ. Biosaf. Res.* 7:61–71

Dechaine JM, Burger JC, Chapman MA, Seiler GJ, Brunick R, et al. 2009. Fitness effects and genetic architecture of plant–herbivore interactions in sunflower crop–wild hybrids. *New Phytol.* 184:828–41

de Jong TJ, Rong J. 2013. Crop to wild gene flow: Does more sophisticated research provide better risk assessment? *Environ. Sci. Policy* 27:135–40

den Nijs HCM, Bartsch D, Sweet J. 2004. *Introgression from Genetically Modified Plants into Wild Relatives*. Wallingford: CABI

deWet JMJ, Harlan JR. 1975. Weeds and domesticates: evolution in the man-made habitat. *Econ. Bot.* 29:99–108

Difazio S, Leonardi S, Slavov G, Garman S, Adams W, Strauss S. 2012. Gene flow and simulation of transgene dispersal from hybrid poplar plantations. *New Phytol.* 193:903–15

Dong SS, Xiao MQ, Rong J, Liao H, Lu B-R, et al. 2011. No effect of transgene and strong wild parent effects on seed dormancy in crop-wild hybrids of rice: implications for transgene persistence in wild populations. *Ann. Appl. Biol.* 159:348–57

dos Reis Goulart ICG, Pacheco MT, Nunes AL, Merotto A. 2012. Identification of origin and analysis of population structure of field-selected imidazolinone-herbicide resistant red rice (*Oryza sativa*). *Euphytica* 187:437–47

Ellstrand NC. 2003. *Dangerous Liaisons? When Cultivated Plants Mate with Their Wild Relatives.* Baltimore: Johns Hopkins Univ. Press

Ellstrand NC, Elam DR. 1993. Population genetic consequences of small population size: implications for plant conservation. *Annu. Rev. Ecol. Syst.* 24:217–42

Ellstrand NC, Heredia SM, Leak-Garcia JA, Heraty JM, Burger JC, et al. 2010. Crops gone wild: evolution of weeds and invasives from domesticated ancestors. *Evol. Appl.* 3:494–504

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, et al. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6(5):e19379

Ghosh A, Haccou P. 2010. Quantifying stochastic introgression processes with hazard rates. *Theor. Popul. Biol.* 77:171–80

Ghosh A, Meirmans P, Haccou P. 2012a. Quantifying introgression risk with realistic population genetics. *Proc. R. Soc. B* 279:4747–54

Ghosh A, Serra MC, Haccou P. 2012b. Quantifying time-inhomogeneous stochastic introgression processes with hazard rates. *Theor. Popul. Biol.* 81:253–63

Gompert Z, Buerkle CA. 2009. A powerful regression-based method for admixture mapping of isolation across the genome of hybrids. *Mol. Ecol.* 18:1207–24

Gompert Z, Buerkle CA. 2011. Bayesian estimation of genomic clines. *Mol. Ecol.* 20:2111–27

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, et al. 2010. A draft sequence of the Neandertal genome. *Science* 328:710–22

Gressel J. 1999. Tandem constructs: preventing the rise of superweeds. *Trends Biotechnol.* 17:361–66

Gross BL, Olsen KM. 2010. Genetic perspectives on crop domestication. *Trends Plant Sci.* 15:529–37

Gutierrez A, Cantamutto M, Poverene M. 2011. Persistence of sunflower crop traits and fitness in *Helianthus petiolaris* populations. *Plant Biol.* 13:821–30

Gutierrez A, Carrera A, Basualdo J, Rodriguez R, Cantamutto M, Poverene M. 2010. Gene flow between cultivated sunflower and *Helianthus petiolaris* (Asteraceae). *Euphytica* 172:67–76

Hajjar R, Hodgkin T. 2007. The use of wild relatives in crop improvement: a survey of developments over the last 20 years. *Euphytica* 156:1–13

Hartman Y, Hooftman DAP, Schranz ME, van Tienderen PH. 2013a. QTL analysis reveals the genetic architecture of domestication traits in Crisphead lettuce. *Genet. Resour. Crop Evol.* 60:1487–500

Hartman Y, Hooftman DAP, Uwimana B, van de Wiel CCM, Smulders MJM, et al. 2012. Genomic regions in crop-wild hybrids of lettuce are affected differently in different environments: implications for crop breeding. *Evol. Appl.* 5:629–40

Hartman Y, Uwimana B, Hooftman DAP, Schranz ME, van de Wiel CCM, et al. 2013b. Genomic and environmental selection patterns in two distinct lettuce crop-wild hybrid crosses. *Evol. Appl.* 6:569–84

Hauser TP, Shim SI. 2007. Survival and flowering of hybrids between cultivated and wild carrots (*Daucus carota*) in Danish grasslands. *Environ. Biosaf. Res.* 6:237–47

Hey J. 2010. Isolation with migration models for more than two populations. *Mol. Biol. Evol.* 27:905–20

Hill R. 2005. Conceptualizing risk assessment methodology for genetically modified organisms. *Environ. Biosaf. Res.* 4:67–70

Hoban S, Bertorelle G, Gaggiotti OE. 2012. Computer simulations: tools for population and evolutionary genetics. *Nat. Rev. Genet.* 13:110–22

17.18   *Ellstrand et al.*

Hooftman DAP, Flavell AJ, Jansen H, den Nijs JCM, Syed NH, et al. 2011. Locus-dependent selection in crop-wild hybrids of lettuce under field conditions and its implication for GM crop development. *Evol. Appl.* 4:648–59

Hooftman DAP, Hartman Y, Oostermeijer JGB, den Nijs JCM. 2009. Existence of vigorous lineages of crop-wild hybrids in lettuce under field conditions. *Environ. Biosaf. Res.* 8:203–17

Hooftman DAP, Jong MJD, Oostermeijer JGB, den Nijs JCM. 2007. Modelling the long-term consequences of crop-wild relative hybridization: a case study using four generations of hybrids. *J. Appl. Ecol.* 44:1035–45

Hooftman DAP, Oostermeijer JGB, Jacobs MMJ, den Nijs JCM. 2005. Demographic vital rates determine the performance advantage of crop-wild hybrids in lettuce. *J. Appl. Ecol.* 42:1086–95

Hovick SM, Campbell LG, Snow AA, Whitney KD. 2012. Hybridization alters early life-history traits and increases plant colonization success in a novel region. *Am. Nat.* 179:192–203

James C. 2012. Global status of commercialized biotech/GM crops: 2011. *ISAAA Brief* No. 43. Ithaca, NY: ISAAA

Jarvis DJ, Hodgkin T. 1999. Wild relatives and crop cultivars: detecting natural introgression and farmer selection of new genetic combinations in agroecosystems. *Mol. Ecol.* 8:S159–73

Kiaer MP, Philipp M, Jørgensen RB, Hauser TP. 2007. Genealogy, morphology and fitness of spontaneous hybrids between wild and cultivated chicory (*Cichorium intybus*). *Heredity* 99:112–20

Kimura M. 1962. On the probability of fixation of mutant genes in a population. *Genetics* 47:713–19

Kovalic D, Garnaat C, Guo L, Yan Y, Groat J, 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* 5:149–63

Kuhner MK. 2009. Coalescent genealogy samplers: windows into population history. *Trends Ecol. Evol.* 24:86–93

Kuparinen A, Schurr FM. 2007. A flexible modelling framework linking the spatio-temporal dynamics of plant genotypes and populations: application to gene flow from transgenic forests. *Ecol. Model.* 202:476–86

Kwit C, Moon HS, Warwick SI, Stewart CN Jr. 2011. Transgene introgression in crop relatives: molecular evidence and mitigation strategies. *Trends Biotechnol.* 29:284–93

Ladizinsky G. 1985. Founder effect in crop-plant evolution. *Econ. Bot.* 39:191–99

Leak-Garcia JA, Holt JS, Kim S-C, Mu L, Mejías JA, Ellstrand NC. 2013. More than multiple introductions: multiple taxa contribute to the genesis of the invasive California's wild artichoke thistle. *J. Syst. Evol.* 51:295–307

Levin DA, Francisco-Ortega J, Jansen RK. 1996. Hybridization and the extinction of rare plant species. *Conserv. Biol.* 10:10–16

Linder CR, Taha I, Seiler GJ, Snow AA, Rieseberg LH. 1998. Long-term introgression of crop genes into wild sunflower populations. *Theor. Appl. Genet.* 96:339–47

Lu B-R, Snow AA. 2005. Gene flow from genetically modified rice and its environmental consequences. *BioScience* 55:669–78

Lu B-R, Yang C. 2009. Gene flow from genetically modified rice to its wild relatives: assessing potential ecological consequences. *Biotechnol. Adv.* 27:1083–91

Magnussen LS, Hauser TP. 2007. Hybrids between cultivated and wild carrots in natural populations in Denmark. *Heredity* 99:185–92

Maughan PJ, Smith SM, Fairbanks, DJ, Jellen EN. 2011. Development, characterization, and linkage mapping of single nucleotide polymorphisms in the grain amaranths (*Amaranthus* sp.). *Plant Genome* 4:92–101

Meirmans PG, Bousquet J, Isabel N. 2008. A metapopulation model for the introgression from genetically modified plants into their wild relatives. *Evol. Appl.* 2:160–71

Mercer KL, Andow DA, Wyse DL, Shaw RG. 2007. Stress and domestication traits increase the relative fitness of crop-wild hybrids in sunflower. *Ecol. Lett.* 10:383–93

Michael TP, Jackson S. 2013. The first 50 plant genomes. *Plant Genome* 6(2): In press (doi: 10.3835/plantgenome2013.03.0001in)

Natl. Res. Counc. 1996. *Understanding Risk: Informing Decisions in a Scientific Society.* Washington, DC: Natl. Acad. Press

Natl. Res. Counc. 2002. *Environmental Effects of Transgenic Plants.* Washington, DC: Natl. Acad. Press

Organ. Econ. Co-op. Dev. 1993. *Safety Considerations for Biotechnology: Scale-up of Crop Plants.* Paris: OECD

Presotto A, Ureta MS, Cantamutto M, Poverene M. 2012. Effects of gene flow from IMI resistant sunflower crop to wild *Helianthus annuus* populations. *Agric. Ecosyst. Environ.* 146:153–61

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–59

Reichman JR, Watrud LS, Lee EH, Burdick CA, Bollman MA, et al. 2006. Establishment of transgenic herbicide-resistant creeping bentgrass (*Agrostis stolonifera* L.) in nonagronomic habitats. *Mol. Ecol.* 15:4243–55

Rieseberg LH, Ellstrand NC. 1993. What can molecular and morphological markers tell us about plant hybridization? *Crit. Rev. Plant Sci.* 12:213–41

Rieseberg LH, Kim MJ, Seiler GJ. 1999. Introgression between the cultivated sunflower and a sympatric wild relative, *Helianthus petiolaris* (Asteraceae). *Int. J. Plant Sci.* 160:102–8

Rieseberg LH, Wendel JF. 1993. Introgression and its consequences in plants. In *Hybrid Zones and the Evolutionary Process*, ed. R Harrison, pp. 70–109. New York: Oxford Univ. Press

Rong J, Janson S, Umehara M, Ono M, Vrieling K. 2010. Historical and contemporary gene dispersal in wild carrot (*Daucus carota* ssp. *carota*) populations. *Ann. Bot.* 106:285–96

Ross-Ibarra J, Tenaillon M, Gaut BS. 2009. Historical divergence and gene flow in the genus *Zea*. *Genetics* 181:1399–413

Schierenbeck KA, Ellstrand NC. 2009. Hybridization and the evolution of invasiveness in plants and other organisms. *Biol. Invasions* 11:1093–105

Snow AA, Campbell LG. 2005. Can feral radishes become weeds? In *Crop Ferality and Volunteerism*, ed. J Gressel, pp. 193–208. Boca Raton, FL: CRC Press

Snow AA, Culley TM, Campbell LG, Sweeney PM, Hegde SG, Ellstrand NC. 2010. Long-persistence of crop alleles in weedy populations of wild radish (*Raphanus raphanistrum*). *New Phytol.* 186:537–48

Snow AA, Pilson D, Rieseberg LH, Paulsen MJ, Pleskac N, et al. 2003. A Bt transgene reduces herbivory and enhances fecundity in wild sunflowers. *Ecol. Appl.* 13:279–86

Song ZP, Lu B-R, Wang B, Chen JK. 2004. Fitness estimation through performance comparison of F$_1$ hybrids with their parental species *Oryza rufipogon* and *O. sativa*. *Ann. Bot.* 93:311–16

Song ZP, Zhu WY, Rong J, Lu B-R. 2006. Evidences of introgression from cultivated rice to *Oryza rufipogon* (Poaceae) populations based on SSR fingerprinting: implications for wild rice differentiation and conservation. *Evol. Ecol.* 20:501–22

Stace CA. 1975. *Hybridization and the Flora of the British Isles*. London: Academic

Stace CA. 2010. *New Flora of the British Isles*. Cambridge, UK: Cambridge Univ. Press. 3rd ed.

Stewart CN Jr, Halfhill MD, Warwick SI. 2003. Transgene introgression from genetically modified crops to their wild relatives. *Nat. Rev. Genet.* 4:806–17

Strasburg JL, Rieseberg LH. 2011. Interpreting the estimated timing of migration events between hybridizing species. *Mol. Ecol.* 20:2353–66

Twyford AD, Ennos RA. 2012. Next-generation hybridization and introgression. *Heredity* 108:179–89

Uwimana B, D'Andrea L, Felber F, Hooftman DAP, den Nijs HCM, et al. 2012a. A Bayesian analysis of gene flow from crops to their wild relatives: cultivated (*Lactuca sativa* L.) and prickly lettuce (*L. serriola* L.), and the recent expansion of *L. serriola* in Europe. *Mol. Ecol.* 21:2640–54

Uwimana B, Smulders MJM, Hooftman DAP, Hartman Y, van Tienderen PH, et al. 2012b. Crop to wild introgression in lettuce: following the fate of crop genome segments in backcross populations. *BMC Plant Biol.* 12:43

Walker RD, Doerfler J. 2009. Wild rice: the Minnesota legislature, a distinctive crop, GMOs, and Ojibwe perspectives. *Hamline Law Rev.* 32:499–527

Wang F, Yuan QH, Shi L, Qian Q, Liu WG, et al. 2006. A large-scale field study of transgene flow from cultivated rice (*Oryza sativa*) to common wild rice (*O. rufipogon*) and barnyard grass (*Echinochloa crusgalli*). *Plant Biotechnol. J.* 4:667–76

Warwick SI, Légère A, Simard MJ, James T. 2008. Do escaped transgenes persist in nature? The case of an herbicide resistance transgene in a weedy *Brassica rapa* population. *Mol. Ecol.* 17:1387–95

17.20   *Ellstrand et al.*

Warwick SI, Stewart CN Jr. 2005. Crops come from wild plants—how domestication, transgenes, and linkage together shape ferality. In *Crop Ferality and Volunteerism*, ed. J Gressel, pp. 9–30. Boca Raton, FL: CRC Press

Weeden NF. 2007. Genetic changes accompanying the domestication of *Pisum sativum*: Is there a common genetic basis to the 'Domestication Syndrome' for legumes? *Ann. Bot.* 100:1017–25

Wegier A, Piñeyro-Nelson A, Alarcón J, Gálvez-Mariscal A, Álvarez-Buylla ER, Piñero D. 2011. Recent long-distance transgene flow into wild populations conforms to historical patterns of gene flow in cotton (*Gossypium hirsutum*) at its centre of origin. *Mol. Ecol.* 20:4182–94

Whitkus R, Doebley J, Wendel JF. 1994. Nuclear DNA markers in systematics and evolution. In *DNA-Based Markers in Plants*, ed. RL Philips, IK Vasil, pp. 116–41. Dordrecht: Kluwer

Whitlock MC, McCauley DE. 1999. Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1)$. *Heredity* 82:117–25

Whitney KD, Randell RA, Rieseberg LH. 2006. Adaptive introgression of herbivore resistance traits in the weedy sunflower *Helianthus annuus*. *Am. Nat.* 167:794–807

Whitney KD, Randell RA, Rieseberg LH. 2010. Adaptive introgression of abiotic tolerance traits in the sunflower *Helianthus annuus*. *New Phytol.* 187:230–39

Whitton J, Wolf DE, Arias DM, Snow AA, Rieseberg LH. 1997. The persistence of cultivar alleles in wild populations of sunflowers five generations after hybridization. *Theor. Appl. Genet.* 95:33–40

Wijnheimer EHM, Brandenburg WA, Ter Borg SJ. 1989. Interactions between wild and cultivated carrots (*Daucus carota* L.) in the Netherlands. *Euphytica* 40:147–54

Wilkinson MJ, Elliott LJ, Allainguillaume J, Shaw MW, Norris C, et al. 2003. Hybridization between *Brassica napus* and *B. rapa* on a national scale in the United Kingdom. *Science* 302:457–59

Wilson GA, Rannala B. 2003. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163:1177–91

Wright S. 1931. Evolution in Mendelian populations. *Genetics* 16:97–159

Xia H-B, Wang W, Xia H, Zhao W, Lu B-R. 2011. Conspecific crop-weed introgression influences evolution of weedy rice (*Oryza sativa* f. *spontanea*) across a geographical range. *PLoS ONE* 6:e16189

Yang X, Xia H, Wang W, Wang F, Su J, et al. 2011. Transgenes for insect resistance reduce herbivory and enhance fecundity in advanced generation of crop-weed hybrids of rice. *Evol. Appl.* 4:672–84

Changes may still occur before final publication online and in print