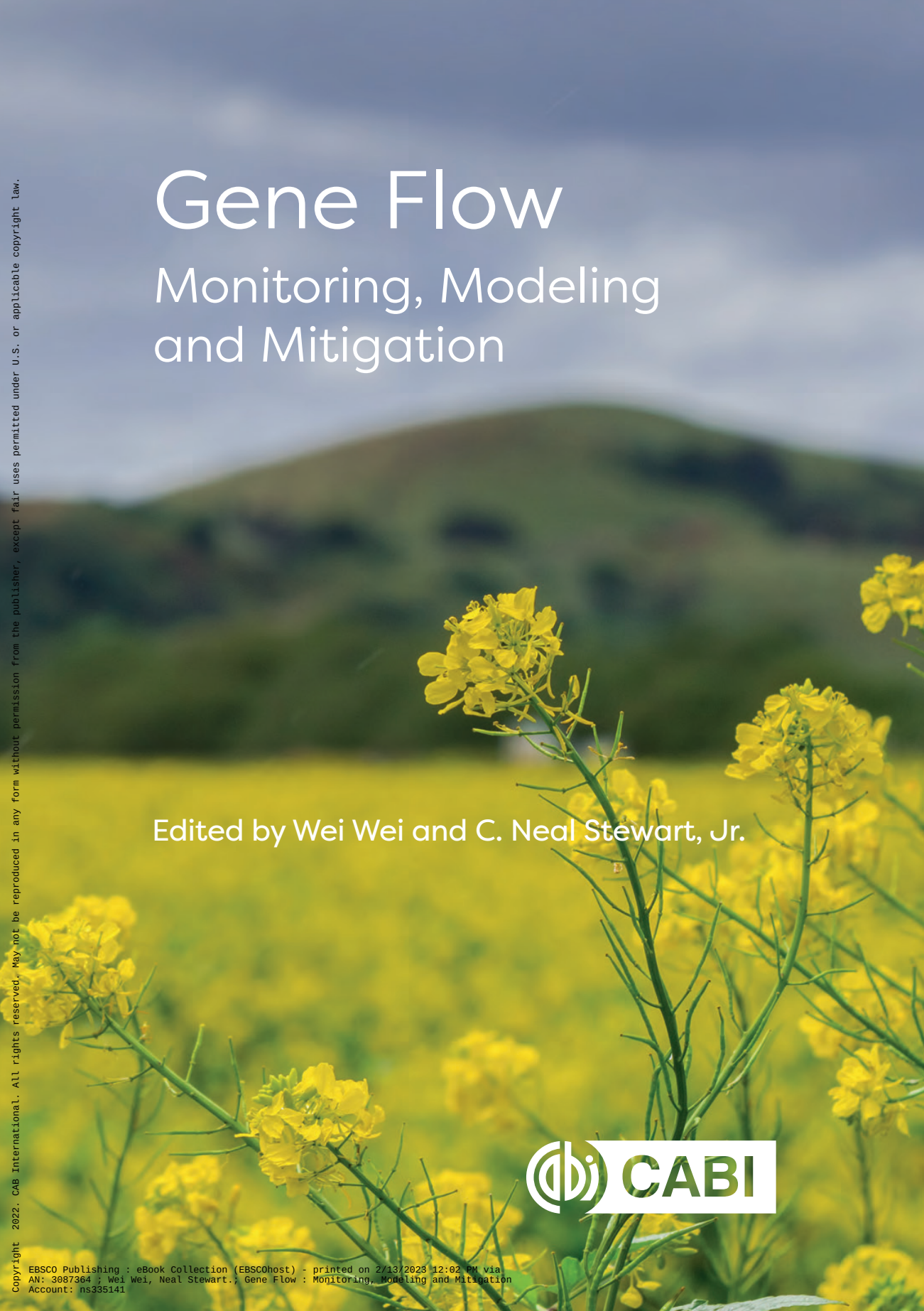
Gene Flow

Monitoring, Modeling and Mitigation



Edited by Wei Wei and C. Neal Stewart, Jr.

CABI

Gene Flow

Monitoring, Modeling and Mitigation

Gene Flow

Monitoring, Modeling and Mitigation

Edited by

Wei Wei and C. Neal Stewart, Jr

CABI is a trading name of CAB International

CABI
Nosworthy Way
Wallingford
Oxfordshire OX10 8DE
UK

Tel: +44 (0)1491 832111
Fax: +44 (0)1491 833508
E-mail: info@cabi.org
Website: www.cabi.org

CABI
WeWork
One Lincoln St
24th Floor
Boston, MA 02111
USA

T: +1 (617)682-9015
E-mail: cabi-nao@cabi.org

A catalogue record for this book is available from the British Library, London, UK.

Library of Congress Cataloging-in-Publication Data

Names: Wei, Wei (Botanist), editor. | Stewart, C. Neal, Jr., editor.
Title: Gene flow : monitoring, modeling and mitigation / edited by Wei Wei and C. Neal Stewart, Jr.
Description: Boston, MA : CAB International, [2022] | Includes bibliographical references and index. | Summary: "This book discusses gene flow of transgenes from genetically modified organisms into nature, with a focus on monitoring, modeling and mitigation. It includes both scientific reviews and perspectives on gene flow and experimental case studies, for example, gene flow of soyabean and poplar"-- Provided by publisher.
Identifiers: LCCN 2021023764 (print) | LCCN 2021023765 (ebook) | ISBN 9781789247480 (hardback) | ISBN 9781789247497 (ebook) | ISBN 9781789247503 (epub)
Subjects: LCSH: Gene flow. | Crops--Genetic engineering. | Transgenic plants. | Biosafety.
Classification: LCC SB123.57 .G4282 2021 (print) | LCC SB123.57 (ebook) | DDC 631.5/233--dc23
LC record available at https://lccn.loc.gov/2021023764
LC ebook record available at https://lccn.loc.gov/2021023765

References to Internet websites (URLs) were accurate at the time of writing.

ISBN-13: 978 1 78924 748 0 (hardback)
978 1 78924 749 7 (ePDF)
978 1 78924 750 3 (ePub)

DOI: 10.1079/9781789247480.0000

Commissioning Editor: David Hemming
Editorial Assistant: Emma McCann
Production Editor: James Bishop

Typeset by Exeter Premedia Services Pvt Ltd, Chennai, India
Printed and bound in the UK by Severn, Gloucester

Contents

Preface

When we started our careers in the 1990s, we both gravitated to performing research on the ecological effects of transgene flow from engineered crops to weedy and wild relatives. In fact, we thought it was so interesting that we continued to work on this topic over the course of our careers. Back then, we were drawn to *Brassica* species because of the interesting relationships among crops and wild plants in the 'Triangle of U' (U, 1935). Both then and now, canola (*Brassica napus*) was the subject of several genetic engineering targets for crop improvement with imminent environmental release. *B. napus* was known to hybridize with several wild relatives within and outside the genus. Thus, the frequency, extent, and consequences of gene flow among the *Brassica*s seemed to be a fruitful system at the time. Indeed, that is so. With some notion of how big and full of uncertainties the world is, we reasoned that gene flow of transgenic DNA would be worth our investment of time and effort. This is true.

It may have surprised our younger selves to have learned that, over two decades later, gene flow research as it pertains to genetically engineered crops is still going strong, even in the face of the absence of ecological disasters in the nearly 30 years of widescale biotech crop commercialization. Nonetheless, ecological timeframes are within the study scope of the sort of research performed to date covered in this book. These studies have greatly informed regulations that govern biotech crops.

The chapters in this book capture various aspects of scientific disciplines that span from organismal studies, to population and community ecology, to molecular biology. We have to understand what we need to manage. While the taxonomic diversity of commercialized engineered plants remains largely focused on highly domesticated row crops, there appears to be a trend to increase both types of crops and traits for engineering (NASEM, 2016). One important driver of these trends has been the ability to easily edit the genomes of organisms using CRISPR, which is increasingly employed as a molecular breeding approach. Therefore, understanding the ecological effects of mainly plant-to-plant gene flow (vertical gene flow) is more important than ever to ensure biosafety.

When we consider long-term evolutionary effects of genetic engineering, the world becomes an even bigger place. Scientists have struggled to capture horizontal gene flow (i.e. that between taxonomically diverse species) experimentally, since the gene flow events are relatively rare. Yet as our understanding increases of cross-kingdom horizontal gene flow, we gain new insights on how we may study transgene effects. One striking recent example is the discovery of plant-to-insect gene flow in evolutionary time in which whitefly obtained a plant gene that allows the insect to resist the effects of phenolic glycoside toxins produced by plants (Xia *et al.*, 2021). This horizontal gene transfer enables the insect to become an agricultural pest. Thus, this first documented case of natural

gene flow from a plant to an insect is strikingly relevant to crop engineering; pest resistance is a primary trait of interest to protect crops from defoliating insects.

And so it goes: humans will manipulate nature and nature will manipulate humans in agriculture, medicine, and other fields. Scientists will continue to gain an understanding of these effects. Ultimately, humans will have to learn to live in harmony with nature. Certainly, gene flow research will remain a constant requirement to inform regulations and assure biosafety for a sustainable future.

References

NASEM (2016) *Genetically Engineered Crops: Experiences and Prospects*. National Academies of Sciences, Engineering, and Medicine (Committee on Genetically Engineered Crops). The National Academies Press, Washington, DC.

U, N. (1935) Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Japanese Journal of Botany* 7, 389–452.

Xia, J.-X., Guo, Z.-J., Yang, Z.-H., Han, H.-L., Wang, S-L. *et al*. (2021) Whitefly hijacks a plant detoxification gene that neutralizes plant toxins. *Cell* 184, 1693–1705. https://doi.org/10.1016/j.cell.2021.02.014

Wei Wei and Neal Stewart
Beijing and Knoxville
April 2021

Contributors

Sarah Zanon Agapito-Tenfen GenØk Centre for Biosafety, Tromsø, Norway

María Eugenia Barrandeguy Department of Genetics, Faculty of Exact, Chemicals and Natural Sciences, National University of Misiones (FCEQyN – UNaM), Posadas, Misiones, Argentina and Institute of Subtropical Biology – Posadas Node (UNaM – CONICET), Posadas, Misiones, Argentina and National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina

Hugh J. Beckie Australian Herbicide Resistance Initiative, School of Agriculture and Environment, The University of Western Australia, Perth, Western Australia, Australia

María Victoria García Department of Genetics, Faculty of Exact, Chemicals and Natural Sciences, National University of Misiones (FCEQyN – UNaM), Posadas, Misiones, Argentina and Institute of Subtropical Biology – Posadas Node (UNaM – CONICET), Posadas, Misiones, Argentina and National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina

Jian-Jun Hu State Key Laboratory of Tree Genetics and Breeding, Research Institute of Forestry, Chinese Academy of Forestry, Beijing, China

Yan-Da Li Institute of Agricultural Engineering, Jiangxi Academy of Agricultural Sciences, Nanchang City, Jiangxi Province, China

Biao Liu Nanjing Institute of Environmental Sciences, Ministry of Ecology and Environment, Nanjing, China

Lai-Pan Liu Nanjing Institute of Environmental Sciences, Ministry of Ecology and Environment, Nanjing, China

Yong-Bo Liu State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing, China

Bao-Rong Lu Ministry of Education Key Laboratory for Biodiversity and Ecological Engineering, Department of Ecology and Evolutionary Biology, Fudan University, Shanghai, China

Sara L. Martin Ottawa Research and Development Centre, Agriculture and Agri-Food Canada, Ottawa, Ontario, Canada

Xiang-Cheng Mi State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing, China

Lei Pei Biofaction KG, Vienna, Austria

Markus Schmidt Biofaction KG, Vienna, Austria

C. Neal Stewart, Jr Department of Plant Sciences, University of Tennessee, Knoxville, Tennessee, USA

Jun-Ming Wang Climate and Atmospheric Science Section, Illinois State Water Survey, Prairie Research Institute, University of Illinois at Urbana-Champaign, Champaign, Illinois, USA

Xin-Yu Wang State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing, China

Wei Wei State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing, China

Fern Wickson North Atlantic Marine Mammal Commission, Tromsø, Norway

Kun Xue College of Life and Environmental Sciences, Minzu University of China, Beijing, China

Lei Zhang State Key Laboratory of Tree Genetics and Breeding, Research Institute of Forestry, Chinese Academy of Forestry, Beijing, China

Yan-Ming Zhu State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing, China

1 Assessing Environmental Impact of Pollen-Mediated Transgene Flow[1]

Bao-Rong Lu*
Ministry of Education, Key Laboratory for Biodiversity and Ecological Engineering, Department of Ecology and Evolutionary Biology, Fudan University, Shanghai, China

Abstract

Potential environmental impact caused by pollen-mediated transgene flow from commercially cultivated genetically engineered (GE) crops to their non-GE crop counterparts and to their wild and weedy relatives has aroused tremendous biosafety concerns worldwide. This chapter provides information on the concept and classification of gene flow, the framework of the environmental biosafety assessment caused by pollen-mediated gene flow, and relevant case studies about transgene flow and its environmental impact. In general, gene flow refers to the movement of genes or genetic materials from a plant population to other populations. Crop-to-crop transgene flow at a considerable frequency may result in transgene 'contamination' of non-GE crops, causing potential food/feed biosafety problems and regional or international trade disputes. Crop-to-wild/weedy transgene flow may bring about environmental impacts, such as creating more invasive weeds, threatening local populations of wild relative species, or affecting genetic diversity of wild relatives, if the incorporated transgene can normally express in the recipient wild/weedy plants and significantly alter the fitness of the wild/weedy plants and populations. It is therefore necessary to establish a proper protocol to assess the potential environmental impacts caused by transgene flow. Three steps are important for assessing potential environment impacts of transgene flow to wild/weedy relatives: (i) to accurately measure the frequencies of transgene flow: (ii) to determine the expression level of a transgene incorporated in wild/weedy populations; and (iii) to estimate the fitness effect (benefit or cost) conferred by expression of a transgene in wild/weedy populations. The recently reported case of non-random allele transmission into GE and non-GE hybrid lineages or experimental populations challenges the traditional method of estimating the fitness effect for the assessment of environmental impacts of transgene flow. Furthermore, case studies of transgenic mitigation (TM) strategies illustrate ways that may reduce the impacts of a transgene on wild/weedy populations if crop-to-wild/weedy transgene flow is not preventable, such as in the case of gene flow from crop rice to its co-occurring weedy rice.

Keywords: biosafety; ecological consequence; fitness; gene flow; introgression; transgene expression

*brlu@fudan.edu.cn

[1] This chapter has been developed from the article 'Transgenic Escape from GM Crops and Biosafety Consequences: An Environmental Perspective' by Bao-Rong Lu in *ICGEB Collection of Biosafety Articles* (2008) 4, 66–141, with permission from ICGEB.

 Gene Flow: Monitoring, Modeling and Mitigation (eds Wei Wei and C. Neal Stewart, Jr)
DOI: 10.1079/9781789247480.0001

1.1 Introduction

The rapid increases in world human population and gradual decreases in global arable land and natural resources, such as water, soil nutrients, and biodiversity, have posed a great challenge to world food security. These changes at the global scale have significantly promoted the research and development of biotechnology, particularly transgenic biotechnology, aiming to meet the world food demands (Lu and Snow, 2005; Lu, 2008). The extensive applications of biotechnologies or genetic bioengineering technologies in agriculture for food production, including the genetic improvement of plant and animal species, have provided great opportunities to meet increasing food demands and to alleviate food security problems (Serageldin, 1999; Lu, 2008). Consequently, many GE plant and animal varieties or lines with improved traits have been successfully produced using transgenic biotechnologies. Some of these transgenic products have been applied commercially (ISAAA, 2019), and some of the other products are in the pipeline, ready for future commercialization.

To date, a large number of GE crop varieties representing different plant species and conveying different transgenic traits, such as disease and insect resistance, herbicide tolerance, high protein content and unique nutritional compounds, salt/drought tolerances, and other improved quality traits, have been commercially cultivated worldwide (ISAAA, 2019). The estimated total area of worldwide commercial cultivation of GE crops is more than 190 million hectares (ISAAA, 2019). The most successful GE crops for commercial application include GE soybean, cotton, oilseed rape, maize, and sugar beet, which have an enormous impact on world crop production and cultivation patterns of agricultural plants (ISAAA, 2019). Undoubtedly, the application of GE technologies in agriculture has contributed significantly to the world's food production and the alleviation of food security problems.

The global commercial cultivation of GE crops with various agronomically beneficial traits has opened a new dimension for meeting world demands and alleviating the great challenge of food security, by enhancing the efficiency of crop production. However, the extensive cultivation and environmental release of GE crop varieties has aroused great biosafety concerns worldwide (Ellstrand, 2001, 2003; Stewart *et al.*, 2003; Lu, 2008, 2016). In some cases, the application of GE crops has stirred up sometimes heated debates and fights in many regions. These biosafety concerns are represented by a wide range of areas such as food and health safety, environmental or ecological safety, labeling of products containing transgenes, legal and regulatory issues, risk assessment systems, and socio-economic and ethical impacts relevant to GE technology and products. In many countries, these biosafety concerns have turned out to be critical constraints for the further development and wider applications of GE biotechnologies and products. Therefore, it is extremely important to address various biosafety issues and to close the 'knowledge gap' for effective biosafety assessment through providing relevant results based on solid scientific research.

Among the biosafety issues concerned, the potential environmental or ecological impacts caused by commercial cultivation of various GE crops on a very extensive scale become the most worrisome biosafety issues (Stewart *et al.*, 2003; Lu and Snow, 2005; Lu, 2008, 2016; Ellstrand *et al.*, 2013). Therefore, it is necessary to help the public to understand the environmental biosafety issues and the current status regarding gene flow, particularly pollen-mediated transgene flow, in addition to fundamental knowledge of the biosafety assessment of potential environmental consequences caused by transgene flow.

1.2 The Importance of Assessing Environmental Impacts of Transgene Flow

It is a great challenge to evaluate environmentally or ecologically related biosafety issues. This is owing to the situation of unpredictable and complex environmental issues, which makes an accurate assessment of the long-term environmental impacts of GE crops extremely difficult. In general, the most concerning environmental or ecological biosafety issues include: (i) potential environmental impacts of transgene flow from GE crops to their non-GE crop varieties and wild relative species (Lu and Snow, 2005; Rong *et al.*, 2007; Lu and Yang, 2009; Wang *et al.*, 2014; Yan *et al.*, 2017); (ii) direct and indirect effects

of toxic transgene products (e.g. insect and disease resistance genes) on non-target organisms such as natural enemies, symbionts, and predators (Oliveira *et al.*, 2007; Lu *et al.*, 2012); (iii) interactions and influences of transgenes and GE crops on biodiversity, ecosystem functions, and soil microbes (Oliveira *et al.*, 2007); (iv) potential risks associated with evolution and development of resistance to biotic (e.g. insect and disease) resistance transgenes in target organisms (Dalecky *et al.*, 2007; Li *et al.*, 2007); and (v) development of more invasive and noxious weeds directly from GE crops (such as oilseed rape) through competition or through transgene introgression into conspecific weeds (the same biological species of crops, such as weedy rice, weedy oilseed rapes, and weedy sugar beets) and crop wild relative species (such as wild *Oryza* species, wild *Brassica* species) (Hall *et al.*, 2000). In this chapter, we only focus on the issues regarding the potential environmental impacts caused by pollen-mediated transgene flow.

Gene flow, as an important evolutionary driving force, is a well-known phenomenon that commonly occurs in natural habitats (Lu and Snow, 2005; Lu, 2008; Ellstrand *et al.*, 2013). Spontaneous transgene flow will likely occur from a GE crop variety to its neighboring non-GE counterparts and to the populations of its conspecific weeds and wild relative species distributed in the vicinity. Environmental impacts caused by transgene flow become the most challenging biosafety issue of GE crop commercialization (Lu and Snow, 2005; Lu, 2008; Lu and Yang, 2009; Ellstrand *et al.*, 2013). This is because the movement of a transgene from a GE crop to its weedy and wild relatives through gene flow cannot be circumvented in some regions, which will possibly cause environmental consequences if significant frequencies of transgenes are introgressed into non-GE crops and weedy/wild relative species. This is particularly true when specific transgenes can introduce evolutionary selective advantages (fitness benefits) or disadvantages (fitness costs) to the non-GE crop varieties or wild/weedy populations. Knowledge on the potential environmental impacts caused by transgene flow is fairly limited, but a considerable number of studies concerning gene flow, transgene flow, and the assessment of potential environmental consequences of transgene flow have been reported.

It is very important to answer the relevant questions associated with transgene flow and its potential consequences for a better understanding of the environmental impacts. Increased knowledge on transgene flow and its environmental impacts will facilitate the effective biosafety assessment of GE crops, and consequently guarantee the safe and sustainable development of GE technology and applications of GE products.

1.3 The Concept and Categories of Gene Flow

1.3.1 The concept of gene flow

Gene flow is a natural phenomenon and has occurred for millions of years. It is an important driving force that significantly influences the evolutionary processes of living organisms (Ellstrand *et al.*, 2013). As a scientific process, gene flow never attracted much public attention until the issues of biosafety associated with the cultivation of GE crops emerged. The public started to be concerned about the potential adverse environmental and socio-economic impacts in terms of 'superweeds' and transgene 'contamination' when they became aware of the possibility of transgene 'escape' into the environment through gene flow. Obviously, the impacts of gene flow have been largely exaggerated. Understanding the concept and types of gene flow, in addition to the fate of a transgene that has transferred into a non-GE crops and recipient populations of wild and weedy relatives through gene flow, will facilitate the effective assessment of environmental impacts caused by transgene flow. This increased knowledge may also relieve the public's tension and concern over transgene flow and its environmental and socio-economic impacts.

By a simple definition, **gene flow** refers to the movement of one or more genes from one organism to another organism. In the terminology of population genetics or evolutionary biology, gene flow (also referred to as gene migration) indicates the transfer of alleles or genes from one individual or population to another (Hartl and Clark, 1989). Usually, there are two types of gene flow: **vertical gene flow** and **horizontal gene flow** (commonly referred to as **horizontal**

gene transfer) (Lu, 2008). Vertical gene flow (commonly called gene flow) refers to the movement of genes between the same species or between closely related species through sexual intercrosses, where genes flow from parents to their descendants vertically. Horizontal gene transfer occurs between unrelated species, such as plants and microorganisms, and between different microorganisms (Thomson, 2001), with an extremely low frequency. Therefore, horizontal gene transfer is not included for discussion in this chapter.

1.3.2 The categories of gene flow

Usually, there are three main avenues for gene flow: **pollen-mediated gene flow**, **seed-mediated gene flow**, and **vegetative propagule-mediated gene flow** (Lu, 2008). The three different types of gene flow will be discussed separately.

1.3.2.1 Seed-mediated gene flow

Seed-mediated gene flow occurs through the natural or human-influenced dispersal of seeds from one population to another by vectors such as animals, wind, water, or humans. For crop species, human activity such as long-distance transportation can move the seeds intentionally within or between geographical regions through seed exchange and national/international trade. Therefore, human activity can extensively promote seed-mediated gene flow with significant frequencies.

1.3.2.2 Vegetative propagule-mediated gene flow

Vegetative propagule-mediated gene flow occurs through the movement or dispersal of vegetative organs (e.g. tillers, stems, roots, and tubers) of plant species vectored by animals, wind, and water. Human activity can also extensively promote vegetative propagule-mediated gene flow at a significant level, for example the long-distance transportation of agricultural products such as potatoes, sugar cane, and forage grasses within or between regions and countries.

1.3.2.3 Pollen-mediated gene flow

Pollen-mediated gene flow occurs when pollen grains travel or flow from one plant individual or population (pollen donor) to another individual or population (pollen recipient), eventually resulting in hybridization and sexually produced hybrids (Fig. 1.1). Pollen-mediated gene flow can happen between individuals within the same biological species or between different but phylogenetically related species. The most common vectors for pollination are wind and animals (e.g. insects and birds).

It is important to point out that pollen-mediated gene flow can cause **sexual hybridization and genetic recombination** between a GE crop and its wild or weedy relatives, incorporating a transgene into the wild or weedy plants or populations. Through consecutive backcrosses between the GE hybrids and wild/weedy plants, the transgene can further introgress into populations of wild or weedy relatives. This type of gene flow (causing sexual hybridization) can lead to genetic recombination between the GE crop and wild/weedy genomes, as well as transmission or spread of a transgene in the wild and weedy populations, resulting in long-term potential ecological and environmental impacts. This chapter will focus on the discussion of potential ecological and environmental impacts from pollen-mediated transgene flow.

1.3.3 Crop-to-crop and crop-to-wild/weedy gene flow

In general, there are different types of pollen recipients – a crop, wild, and weedy species – in relation to transgene flow and its environmental impacts. A transgene can move from a GE crop to these recipients, including a non-GE crop counterpart and population of wild or weedy relatives of the crop in the vicinity, through pollen-mediated gene flow. According to the types of pollen recipients of a GE crop, pollen-mediated gene flow can be further categorized into **crop-to-crop gene flow**, **crop-to-wild gene flow**, and **crop-to-weed gene flow** (Lu, 2008; Lu and Yang, 2009).

Fig. 1.1. Possibility of spontaneous gene flow (or natural hybridization) between a crop and its wild relative species using the rice genus (*Oryza* L.) as an example. The plant on the right is cultivated rice (*O. sativa* L.), the plant on the left is the wild ancestor (*O. rufipogon* Griff.), and plants in the middle are the resulting interspecific hybrid descendants that usually become weedy rice (*O. sativa* f. *spontanea* Rosh.) and can further hybridize with cultivated rice.

1.3.3.1 Crop-to-crop gene flow

Crop-to-crop gene flow refers to the movement of a gene from one crop variety to another crop variety (Fig. 1.2). Gene flow from one crop field to other adjacent fields planted with non-GE crop varieties of the same species can easily happen. The frequencies of transgene movement mediated by pollination between GE and non-GE crops depend essentially on the breeding (mating) systems and pollen quantity of the crops. Relatively high gene flow frequencies will be expected in outcrossing crops at the same spatial dimension from a pollen source under the same climate condition compared with inbreeding crop species where low gene flow frequencies will be expected.

For practical purposes, understanding the frequency of crop-to-crop gene flow for a particular crop species through pollination is very useful, if different growers or countries want to separate GE crops from their non-GE varieties for marketing or regulatory reasons. This will help to determine the extent of consequences caused by crop-to-crop gene flow in different crop species. For example, cultivated rice is characterized by its self-pollination and very little cross-pollination between adjacent plants or fields (typically < 1.0%). Experiments in Italy showed that pollen-mediated gene flow from a transgenic, herbicide-resistant rice variety to adjacent plants of a non-transgenic counterpart was 0.05–0.53% (Messeguer *et al.*, 2001). Likewise, in China, the average frequency of transgene flow from insect-resistant GE rice varieties and their non-GE counterparts was 0.02–0.80% when the plants were grown at close spacing (Rong *et al.*, 2005).

A similar study based on molecular fingerprints (simple sequence repeat, SSR) also indicated very low gene flow frequencies between hybrid rice and traditional landraces grown next to each other. Interestingly, the measured gene flow frequencies of landrace-to-hybrid (~0.1%) and hybrid-to-landrace (~0.04%) were significantly different (Rong *et al.*, 2004). This asymmetric pattern in rice suggests that the frequency of gene flow is essentially determined by the outcrossing ratios of pollen recipients, given the same amount of pollen load. A further study has shown that gene flow frequency dramatically reduced with the increase of spatial isolation distances from the GE rice pollen donors by only a few meters (Rong *et al.*, 2007).

These findings are consistent with the small isolation distances that are recommended for maintaining the purity of cultivated rice grown in seed nurseries. In the USA, for instance, rice

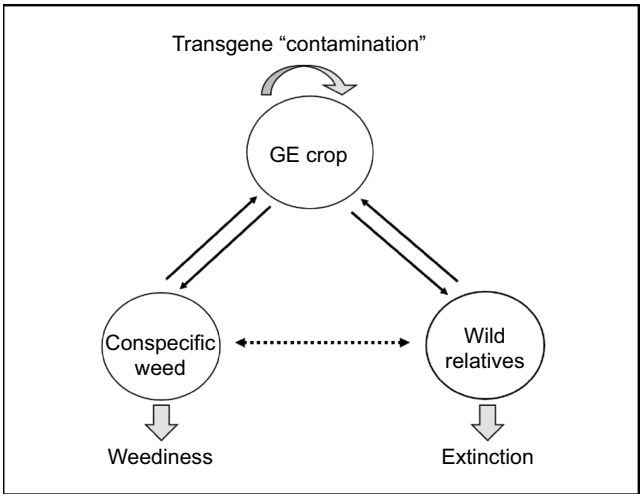


Fig. 1.2. Crop-to-crop (top), crop-to-wild relative (right), and crop-to-weed (left) gene flow mediated by pollination. Arrows with solid lines indicate the direction of gene flow; an arrow with an empty line indicates the potential gene flow if wild and conspecific weed populations co-occur; shaded arrowheads indicate the potential environmental impacts.

plants that are grown for certified seed to be sold to farmers must be isolated from other rice varieties by only 6 m or less (Gealy *et al.*, 2003). Consequences caused by crop-to-crop gene flow in a cross-pollinating species such as maize would be much more serious. This concern can be reflected by the world debates caused by the 'contamination' of traditional maize varieties in Oaxaca, Mexico (Quist and Chapela, 2001; Ortiz-García *et al.*, 2005).

1.3.3.2 Crop-to-wild/weedy gene flow

Crop-to-wild gene flow refers to the movement of a gene or genetic materials from a crop variety to its wild relative species that belong to different species but have a certain degree of genetic affinity with the crop species (Fig. 1.2). **Crop-to-weed gene flow** refers to the movement of a gene or genetic materials from a crop variety to its conspecific weeds that belong to the same biological species as the crop species (Fig. 1.2). As mentioned earlier, gene flow mediated by pollination may allow permanent introgression and spread of a transgene into wild/weedy populations only in the cases of recurrent hybridization, which is relatively rare in crops (Stewart *et al.*, 2003). Nonetheless, introgression may cause evolutionary changes of wild/weedy populations resulting in potential environmental impacts, depending on the trait.

Many studies have shown that crops are viable in natural ecosystems and can interbreed with their wild relatives (Hall *et al.*, 2000; Stewart *et al.*, 2003; Chen *et al.*, 2004; Lu and Snow, 2005; Lu, 2008; Lu and Yang, 2009; Ellstrand *et al.*, 2013). The most publicized environmental concern is the creation of more invasive weeds if GE crops modified to tolerate biotic and abiotic stresses transfer their transgenes into wild or weedy relatives through gene flow. Crops can also be modified with traits that allow them to reproduce more (for example, by enhancing seed production), and grow in new types of habitats (for example, by enhancing drought and cold tolerance). In principle, the potential environmental impacts caused by crop-to-wild/weedy transgene flow can be effectively determined by the frequency of transgenes that have outflowed to the wild and weedy populations, and by the characteristics of the transgenic traits that have or do not have evolutionary advantages under natural selection.

When wild/weedy populations incorporate a transgenic trait likely to confer a selective advantage and are then exposed to a relevant selective pressure (e.g. herbicides, pest attacks or drought/salinity stresses), these populations

may exhibit an enhanced performance (Linder and Schmitt, 1994; Ellstrand, 2003; Song *et al.*, 2004b; Lu and Snow, 2005; Mercer *et al.*, 2007; Lu and Yang, 2009; Li *et al.*, 2016; Yan *et al.*, 2017), leading to unwanted environmental consequences. It is necessary to point out that crop-to-wild/weedy gene flow can recur over time, because plants of wild/weedy species generally persist in their habitats, or their seeds remain in the local soil seedbank. The frequency of transgene flow can increase through recurrent gene flow over different years or seasons from GE crops cultivated in surrounding areas. This is different from the case of crop-to-crop transgene flow, where crops are harvested at the end of the season. If the crops are consumed or used by industry/manufacturing, the transgenes do not accumulate in the crop populations. However, if the crops are used as seeds, the transgene-contaminated seeds may be propagated and disseminated to different regions.

1.4 Assessing Impacts from Pollen-Mediated Transgene Flow

1.4.1 A framework

To effectively assess environmental biosafety impacts created by transgene flow from GE crops to their wild/weedy relatives through pollination, it is essential to attain the main knowledge (e.g. baseline data of crops and their wild relatives) that is relevant to the specific risk assessment. Particularly, it is important to determine knowledge gaps for which it is essential to address the relevant scientific questions, stringently following the principle of risk assessment. Based on the well-known risk assessment framework, a risk is derived as a function of hazard and exposure:

$$\text{Risk} = \text{hazard} \times \text{exposure} \qquad \text{(Eqn. 1.1)}$$

where **risk** represents the uncertainty about deviation from the normal consequence, **hazard** represents a transgenic plant or transgene product with potential adverse or harmful effects, and **exposure** represents a quantitative measurement of the extent to which a given hazard is present in the environment or ecosystem.

Risk, in the context of environmental biosafety and of transgene flow in particular, indicates the probability of adverse effect from a transgene escaping into the environment through gene flow. Accordingly, the key knowledge gaps for environmental impacts from transgene flow should include the following aspects.

1. What is the frequency for a transgene to transfer from a GE crop to its non-GE crop and wild/weedy relatives through pollen-mediated gene flow (exposure)?
2. What is the expression level of a transgene that has transferred into individuals of wild/weedy relatives (hazard)?
3. Does a transgene incorporated into a wild/weedy relative significantly enhance invasiveness or reduce the survival of the wild/weedy populations, posing environmental impacts (hazard)?
4. How can the environmental impacts from transgene flow be correctly assessed (methodology)?

Answers to these questions will be essential to close the knowledge gaps, and therefore facilitate the establishment of a standard protocol to assess the environmental impacts caused by transgene flow.

The assessment of environmental impacts from transgene flow is a procedure that helps to determine the likelihood and magnitude of expected risks, which essentially depends on the success of transgene flow and establishing/spreading of the transgene in a wild/weedy population. To meet the objective of a risk assessment, it is necessary to establish a general framework and protocol for determining whether or not environmental risks associated with transgene flow will occur, and how serious the risks will be at the various steps. Through logical reasoning and analysis of potential environmental consequences from transgene flow, we understand that there are three major steps or procedures closely associated with the accurate assessment of transgene flow and its environmental impacts:

1. To measure the frequencies (exposure) of a transgene flow from a GE crop to its crop counterparts and wild/weedy relatives.

2. To estimate expression and inheritance (hazard) of an introgressed transgene in the hybrids and advanced progenies between a GE crop and wild/weedy relatives.
3. To analyze the level or magnitude of changes and fitness effect (hazard) conveyed by an introgressed transgene in wild/weedy recipient individuals/populations (Lu, 2008).

In addition, important principles, such as the science-based principle, case-by-case principle, and step-by-step principle, should be strictly followed in the environmental risk assessment. These principles serve as a guideline for effectively undertaking the risk assessment of environmental impacts from transgene flow. Given that the consequences and impacts of transgene flow to a non-GE crop counterpart and wild/weedy populations are different, the potential consequences and impacts from crop-to-crop and crop-to-wild/weedy transgene flow will be discussed separately.

1.4.2 Potential consequences from crop-to-crop transgene flow

Depending on species, transgenes can easily move from a GE crop variety to other non-GE crop varieties of the same species grown in the adjacent fields through pollen-mediated gene flow. The frequencies of transgene movement between GE and non-GE crops depend essentially on the breeding (mating) systems and pollen quantity of the crops, particularly the pollen donors. Relatively high gene flow frequencies will be expected in outbreeding crops at the same spatial dimension from a pollen source under the same climate condition compared with inbreeding crop species where low gene flow frequencies will be expected. On a very practical level, the understanding of crop-to-crop gene flow through pollination is useful if different growers or countries want to separate GE crops from their non-GE crop varieties for marketing or regulatory reasons. Such knowledge will help to determine the extent of consequences caused by crop-to-crop gene flow in different crops species.

Commonly, the greatest concern about a transgene transferring from a GE crop to its non-GE crop varieties through pollen-mediated gene flow is the transgene 'contamination' of non-GE crop varieties that have commercial values (Fig. 1.3). The consequences of crop-to-crop transgene flow may not necessarily be related to environmental issues. Instead, crop-to-crop transgene flow may cause problems of regional or international trading or legal disputes, if the regions or countries require transgenic labels legally for their products at a certain level (threshold), such as > 0.9% in EU countries. In addition, the movement of a transgene from a GE crop variety to local varieties or landraces through crop-to-crop gene flow may result in the gradual losses of genetic diversity of the local varieties or landraces, if the transgene encodes traits with strong natural selective advantages (Fig. 1.3). Therefore, to reduce the potential consequences from crop-to-crop transgene flow, the measurement and control of gene flow frequencies is the key for the biosafety assessment.

1.4.3 Potential environmental impacts from crop-to-wild/weed transgene flow

Transgenes can move from a GE crop variety to the populations of its wild or weedy relative species occurring in the vicinity (wild or weed) and co-occurring in fields (weed) through pollen-mediated gene flow. Following the risk assessment framework, the potential environmental or ecological impacts of crop-to-wild/weedy transgene flow are determined by three components: (i) the frequency of transgene flow; (ii) the expression level of a transgene; and (iii) fitness effect of a transgene. The three components or steps are important for the assessment of potential risks caused by crop-to-wild/weedy transgene flow, in which the frequency represents the exposure, whereas expression and fitness of the transgene represent the hazard. Understanding the exposure (frequency) and hazard (fitness) of transgene flow will facilitate the assessment of its risks (Fig. 1.4). To evaluate the potential impacts of crop-to-wild/weedy transgene flow, the measurement and control of gene flow frequencies is the key for the biosafety assessment.

The direct and immediate consequence of crop-to-wild/weedy gene flow may be changes in the genetic integrity/structure of wild or weedy

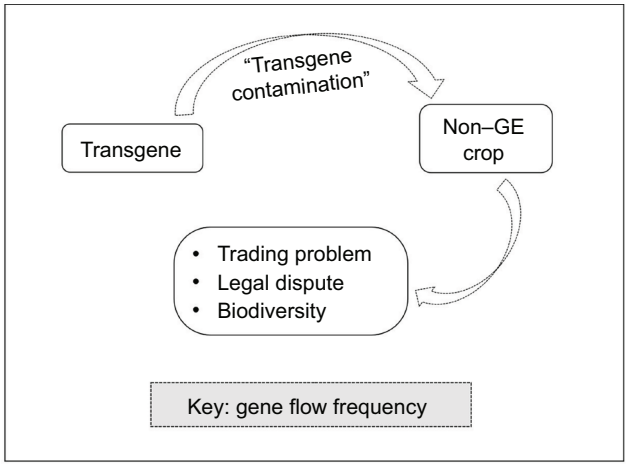


Fig. 1.3. Potential consequences caused by crop-to-crop transgene flow mediated by pollination. The major consequence is 'transgene contamination' to non-GE crops. Proper measurement of transgene flow frequencies is the key determinant to reduce the risks.

species by recurrently 'pumping' transgenes into wild populations. If the transgene conveys a selectively advantageous trait(s) to the host, the movement of such a transgene into wild relatives may change the fitness of wild or weedy relatives, resulting in demographic alterations (either increase or decrease) of the wild/weedy populations, or rapid accumulation and spread of the transgenes in the populations by speeding up introgression of the transgene into wild/weedy populations. These changes may lead to diverse environmental impacts.

There are different possibilities for the fate of a wild/weedy population that contains a transgene (Fig. 1.5). If the transgene can enhance the fitness of wild relatives through the expression of a favorable trait such as pest resistance, drought tolerance, or enhanced growth ability, the outflowed transgene may persist and spread in the population in the absence of negative pleiotropic effects (Fig. 1.5, upper pathway). Individuals that express the transgene may outcompete other individuals without the transgene in the population under natural selection. This process may promote a rapid increase of transgenic individuals and enhance their invasiveness, causing different degrees of new weed problems by enabling the wild/weedy populations to expand into new territories. With the advent of well-known herbicide-resistant crops, such as soybean and oilseed rape, there is strong public concern about the production of 'superweeds' that are resistant to multiple herbicides (Hall *et al.*, 2000).

On the other hand, if the transgene reduces the fitness of receiving wild relatives, the frequencies of individuals that contain the disadvantageous transgene may decrease gradually by negative selection (Fig. 1.5, lower pathway), but only if the transgene becomes fixed in the population. This process may be accelerated by recurrent gene flow and introgression from the nearby GE crop, possibly leading to the extinction of local populations by the so-called swarm effect (Ellstrand and Elam, 1993). In many parts of the world, such swarm effects have already happened in the absence of GE crops through crop-to-wild gene flow (Kiang *et al.*, 1979), where populations of wild relatives are surrounded by crop fields in agricultural ecosystems and inhabit bordering areas between agricultural lands and natural habitats (Ellstrand *et al.*, 1999, 2013; Ellstrand, 2003; Lu, 2008; Lu and Yang, 2009). If the transgene is selectively neutral and does not alter the fitness of the wild relatives, such as those genes encoding nutritional compounds and quality traits that are only favorable to human taste or health conditions, the outflow of such a transgene may not have considerable influences

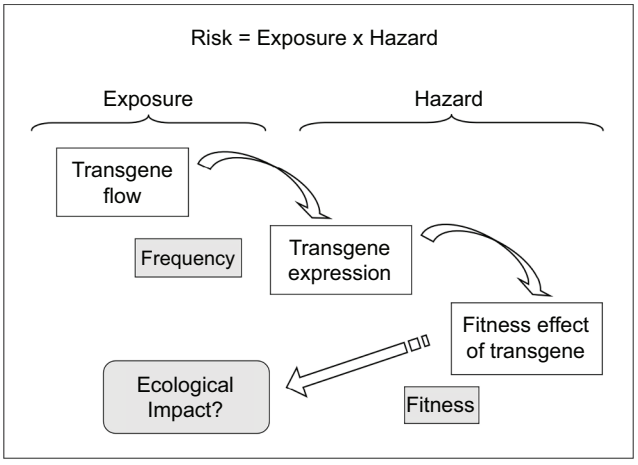


Fig. 1.4. Potential impacts caused by crop-to-wild and crop-to-weed transgene flow mediated by pollination, in which three important steps are needed: (i) measuring transgene flow; (ii) analyzing transgene expression in wild/weedy hybrid populations; and (iii) assessing fitness effect in wild/weedy hybrid populations. Relevant estimates of both transgene flow frequencies and fitness effects in appropriate environments are important for risk assessment.

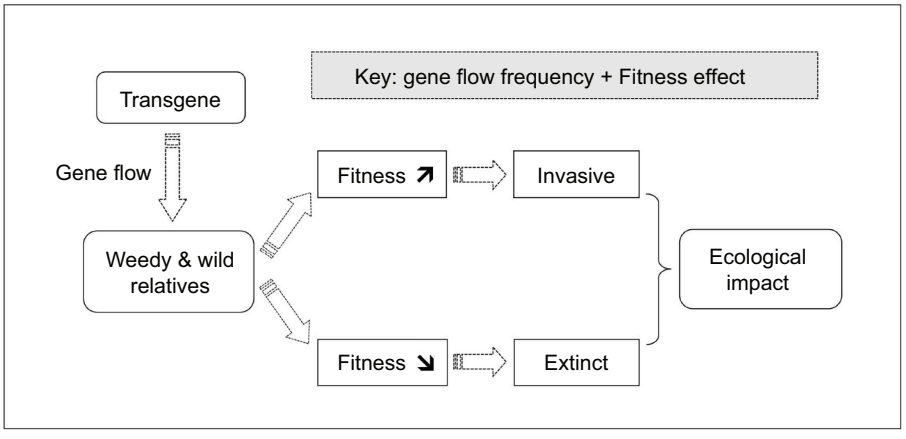


Fig. 1.5. An illustration to show the hypothetical reasoning of potential environmental impacts caused by crop-to-wild and crop-to-weed transgene flow. Accurate measurement of fitness effect becomes the key determinant for risk assessment at this step.

on the population dynamics of the wild relatives. In this case, the likelihood of environmental impacts of such gene flow should be low. It should be noted that neither effects of increased invasiveness nor extinction have been observed in transgenic systems, but they have been observed as part of nature.

In summary, the potential environmental impacts caused by crop-to-wild/weedy transgene flow can be effectively determined by: (i) frequencies of transgenes that have moved to the wild/weedy populations; and (ii) by the transgenic traits that bring about evolutionary advantages/disadvantages or fitness

benefit/cost under natural selection. When wild/weedy populations incorporate a transgenic trait likely to confer a considerable selective advantage or disadvantage and are exposed to a relevant selective pressure, the populations could conceivably exhibit an enhanced or reduced performance, leading to unwanted environmental impacts (Fig. 1.5). Importantly, crop-to-wild/weedy transgene flow may recur over time, because of the long-term persistence of wild/weedy plants in their habitats, or wild/weedy seeds in the soil seedbanks. Therefore, the environmental impacts caused by crop-to-wild/weedy gene flow may theoretically be significant and need to be determined in the long term, as short-term experiments may warrant. In addition, there are still many biological mechanisms underlying the process of gene flow and fitness change to be understood. Science-based studies should be conducted to test whether in reality the predicted impacts exist under a case-by-case situation, and to measure the magnitude of such impacts.

1.5 Methods to Assess Impacts Caused by Pollen-Mediated Transgene Flow

1.5.1 Measuring pollen flow

For measuring pollen-mediated transgene flow, it is necessary to know the dynamics of pollen flow of a specific crop (pollen donor), which is the essential environmental risk assessment of a transgenic crop (Song *et al.*, 2004a; Lu, 2008). This is because pollen acts as a vehicle that disseminates transgenes. A considerable amount of pollen flow between a crop and its wild/weedy relatives will result in outcrossing (gene flow) and, consequently, alter the gene pool of wild/weedy and cultivated species (Ellstrand *et al.*, 1999). In addition, it is necessary to determine pollen 'contamination' and avoided 'contamination' such as in conventional crop breeding, especially for the production of hybrid crops. An effective method for measuring pollen flow of a crop is important, which will provide valuable information to estimate pollen-mediated transgene flow and the minimum isolation distance as the buffering area between fields of a GE crop and its non-GE crop and wild relatives.

Pollen flow assessment usually includes the measurement of pollen spreading rates, the maximum distance of pollen dispersal horizontally, spatial dynamics of pollen movement vertically, and patterns of pollen dispersal of a particular crop (Song *et al.*, 2004a; Lu, 2008). The relevant literature indicates that airborne pollen flow can follow an exponential leptokurtic pattern and can be greatly influenced by meteorological factors, particularly microclimatic conditions, such as wind speed and direction, ambient temperature, and relative air humidity (Rong *et al.*, 2010; Wang *et al.*, 2016). It is generally understood that pollen dispersal is linked to height, namely that herbaceous species have more intensive pollen content at lower heights (Alcazar and Comtois, 2000), and that airborne pollen flow and source size are more significant than density, as is the case in anthophilous pollen flow (Rognli *et al.*, 2000).

However, pollen flow in different plant species mediated by various vectors differs significantly in dispersal patterns (Alcazar *et al.*, 1998). It is therefore important to carry out a detailed investigation of pollen flow of a particular species, where environmental risk assessment requires an accurate estimation of pollen flow including both distance and intensity of pollen dispersal. Song *et al.* (2004a) conducted a study of pollen flow for cultivated rice under experimental conditions, where the authors used a cultivated rice variety, 'Minghui-63', as a pollen donor to analyze the pollen flow pattern at horizontal and vertical levels. Data obtained from pollen traps for six designed populations (pollen sources) at different intervals showed that the dispersal of rice pollen decreased with the increase of distance from pollen sources and that the rice pollen flow was significantly influenced by weather conditions, particularly by wind direction and speed. The observed distance of rice pollen dispersal was 38.4 m in a downwind direction, suggesting a relatively small range of rice pollen flow, although the fitted curve of the maximum distance of rice pollen flow was up to 110 m, using regression analysis. In addition, more pollen grains were detected at the height of 1.0–1.5 m, indicating the relatively low heights of rice pollen dispersal (Song *et al.*, 2004a).

Information regarding pollen flow dynamics serves as a useful basis for understanding pollen-mediated transgene flow from a transgenic crop to its non-transgenic crop and wild/weedy relatives.

1.5.2 Measuring pollen-mediated transgene flow

Estimating frequencies of pollen-mediated gene flow is a key component of and the first step for risk assessment. It will answer questions relating to 'exposure' in the risk assessment process. Measuring the frequencies at which gene flow occurs is a challenging task in many crop-to-crop and crop-to-wild/weedy complexes, because gene flow frequencies can vary significantly between plant species with different mating systems and modes of pollination (for example, wind pollination versus insect pollination), in addition to the complicity of environmental factors. Obviously, pollen-mediated gene flow is influenced by many biological factors, such as flowering habits, outcrossing ratios, the amount of pollen produced by donors, the duration of pollen viability, and the population sizes of pollen donors and recipients. Also, pollen-mediated gene flow can be determined by many non-biological factors, such as the distance between pollen donors and recipients, wind speed or insect pollinator activity, humidity, and other climate conditions.

The measurement of gene flow frequencies for different crop–wild species at different locations should strictly follow the case-by-case principle (Ellstrand *et al.*, 1999; Lu and Snow, 2005; Lu and Yang, 2009; Ellstrand *et al.*, 2013). In other words, gene flow data obtained for one type of crop (e.g. wind- and self-pollinated) cannot be used for the risk assessment of another type of crop (e.g. insect-pollinated outbreeder). A number of experimental and empirical approaches have been developed to estimate the relative frequencies of pollen dispersal and pollen-mediated gene flow rates of plant species (Ellstrand, 2003; Song *et al.*, 2003, 2004b; Chen *et al.*, 2004; Lu and Snow, 2005; Wang *et al.*, 2006; Koopman *et al.*, 2007), which can be used for measuring transgene escape from GE crops. Prior to obtaining crop-to-wild gene flow frequencies, it is very helpful to study the crossability and compatibility between the crop species to which the GE trait will be, or has already been, transferred and the potential wild relative species that occur in the areas where the GE crop is expected to be cultivated. This assists in determining the various possibilities in which the transgene could move from the GE crop to its wild relative species. If the crop has a high crossability with its wild species under natural conditions, then transgene outflow to wild species/populations will be high, and vice versa.

Studies of the crossability between a crop and its wild relative species can be conducted in the greenhouse or field by hand pollination, including reciprocal crosses (using wild relatives as both maternal and paternal parents). The number of individuals used during test pollinations should be sufficient (> 30 individuals for each parent ideally) to ensure that the resulting data is truly representative of the crossability between the crop and wild species. If the crop species is not compatible with the wild relatives or the ratio of crossability is extremely low, no further biosafety assessment for transgene escape to wild relatives through pollen-mediated gene flow is needed.

Otherwise, the biosafety assessment should proceed to the next tier/step. The crop–wild crossability can be estimated from the ratio of seed set (R_s) between the crop and wild species after hand pollination, which can be calculated from the formula:

$$R_S\,(\%) = Nh / Tf \times 100\% \qquad \text{(Eqn. 1.2)}$$

where R_s represents ratio of seed set, Nh represents number of hybrid seeds obtained, and Tf represents the total number of flowers pollinated.

It is also important to examine the fertility of the artificial crop–wild hybrids, including pollen (male) fertility and seed (male and female) fertility. Pollen fertility, ascertained by staining hybrid pollen grains with an iodine–potassium iodide (I–KI) solution, can be calculated from the formula:

$$F_{po}\,(\%) = P_s / P_t \times 100\% \qquad \text{(Eqn. 1.3)}$$

where F_{po} represents pollen fertility, P_s represents number of stainable pollen grains, and P_t represents the total amount of pollen examined.

Seed fertility can be estimated using the following formula:

$$F_{pa}(\%) = S_g / S_t \times 100\% \quad \text{(Eqn. 1.4)}$$

where F_{pa} represents seed fertility of hybrids, S_g represents number of good (fertile) seeds and S_t represents the total number of seeds examined.

An accurate measurement of the frequency of gene flow mediated by pollination is vital in preliminary estimations of the extent of risks caused by transgene escape to wild relative species, given that the risk is determined by the number of transgenes that move to wild populations and the adverse effect of the transgene to wild populations. However, the frequency of gene flow can be significantly different among plant species with diverse mating systems, as well as diverse climate conditions. Therefore, gene flow frequencies measured from one species cannot be used for another species, and even the measured gene flow frequencies from one variety of the same species cannot be used completely for other varieties under different environmental conditions. Thus, the 'case-by-case' principle should be strictly followed in biosafety assessment.

Information on individual outcrossing rates and variation in flowering times is useful for evaluating the potential for hybridization. Small-scale experiments involving plants with distinct genetic markers can be used to measure gene flow between adjacent plants in a given location and year, but they may not reflect large-scale or long-term processes. Although these types of information are undoubtedly incomplete, they can be used to assess the potential for transgene escape and to develop strategies to minimize the escape of certain types of transgenes through pollination.

Seed-mediated gene flow can also be very effective as a means of transgene dispersal, especially when seeds are traded within and between countries. Usually, rice seeds have their husks removed before commercial shipments and exports for food consumption. In this case, the seeds are not viable, because their embryos are damaged during the milling process. However, sometimes rice seeds are transported without dehusking, including those that are intended for domestic seed sales. Also, viable seed can be dispersed when the grain is threshed and dried in the open air, and when it is handled, sorted, and transported for milling. A few studies have attempted to quantify the extent of gene flow by means of seed dispersal in rice or other cultivated species (Saji *et al.*, 2005). However, in the following sections, pollen-mediated gene flow from cultivated rice (*Oryza sativa*) to other rice crops, weedy rice (*O. sativa* f. *spontanea*), and wild rice species (other *Oryza* species) will be the main focus.

1.5.3 Measuring transgene expression in wild or weedy relatives

A large number of examples demonstrate that genes can move from a crop to its wild or weedy populations through gene flow, the frequencies of which in some plant species can be relatively high (Ellstrand, 2003; Song *et al.*, 2003, 2004b; Chen *et al.*, 2004; Wang *et al.*, 2006). This means that in many cases transgene escape from GE crops to their wild relatives is unavoidable. However, whether or not the escape of the transgene will have ecological or environmental consequences depends essentially on the extent to which the function of the transgene will be maintained (or changed) in wild relative receiving species. If the escaped transgene cannot express normally (i.e. with much lower levels of its products) in wild relative species following outflowing, the transgene may not alter the traits or fitness of the wild relatives. As a result, the transgene escape would not introduce any ecological consequences. On the other hand, if the transgene can express normally, or even more strongly, than in the parental GE crops, after its introgression into wild relative species, then the escaped transgene might provide a fitness advantage to the populations of recipient wild relatives, resulting in unwanted ecological consequences.

To facilitate the biosafety assessment of transgene escape to populations of wild relative species, it is important to conduct scientific research for a proper estimation of the expression level of a particular transgene in wild individuals, as well as the inheritance of the transgene in wild populations under different environments. This is particularly relevant for transgenes that have obvious selective advantages for biotic

(such as insect and disease) and abiotic (such as drought and salinity) stresses, if pollen-mediated transgene flow to wild relative species cannot be circumvented. In this case, questions relating to transgene expression and inheritance following introgression into the wild relatives are more important.

There are a number of methodologies to estimate transgene expression levels, including by enzyme-linked immunosorbent assay (ELISA) (Bashir *et al.*, 2005), reverse transcriptase polymerase chain reaction (RT-PCR) (Wang *et al.*, 2014), and real-time PCR (Wang *et al.*, 2014). Commonly, the principle of estimating transgene expression is to measure the amount or level of transgene products (e.g. the Bt toxic protein) that can be detected in individuals or populations of wild relatives, in comparison with the parental GE crops. For example, to determine whether an insect-resistance transgene (*Bt*) introgressed into wild populations will have significant environmental impacts, under the hypothesis that the target lepidopteron species occur in association with the wild relative populations and regulate the populations' dynamics, estimations of transgenic expression levels in wild individuals will help to predict the possibility of changes to the composition of wild populations. Artificial hybrids and their progenies between a GE (e.g. *Bt* insect resistance) crop and wild relatives can be produced by artificial crosses and backcrosses in order to estimate the transgene expression in wild species. The Bt protein content can then be measured in the GE crop, F_1 hybrid and advanced-hybrid populations. If the Bt protein content in the hybrid progenies of wild relatives is similar to that in the GE crop, then it is assumed that the transgene will be able to kill the target lepidopteran species in the new host wild populations. On the other hand, if the Bt protein content is dramatically reduced or undetectable in the hybrid descendants of wild relatives, it is assumed that the transgene will not cause further ecological consequences (see Fig. 1.4), due to 'loss of function' of the inserted transgene in the wild populations.

Similarly, the inheritance of a transgene in the wild population can also be estimated by the production of artificial populations of F_1 hybrids and advanced progenies through crosses between a GE crop and wild relatives, and subsequent backcrosses and self-pollination. If the transgene is normally expressed in crop–wild/weedy hybrids and descendants, as well as inherited between different generations, further biosafety assessments of possible environmental impacts will be necessary.

1.5.4 Measuring fitness effect brought about by a transgene in wild/weedy populations

For the estimation of long-term persistence and spread of transgenes in crop wild populations in relation to the fitness change, several key factors should be taken into consideration, including: (i) genetic mechanisms (e.g. genetic relationships and compatibility) that allow the transfer of transgenes into wild populations; (ii) the degree of gene flow to which transgenes are incorporated into wild/weedy populations; (iii) the fitness of early hybrids relative to their wild/weedy parents; and (iv) possible fitness costs or benefits that are associated with a particular transgene (Jenczewski *et al.*, 2003).

If a transgene can move from a GE crop to its weedy or wild relatives at a certain level, and at the same time the incorporated transgene can be normally expressed and inherited in the wild/weedy relatives, it is then very important to continue the risk assessment to understand whether or not the transgene will change fitness of the recipient populations of wild/weedy relatives, resulting in environmental impacts. This is because if expression of the incorporated transgene causes changes to the fitness of recipient wild/weedy relatives, the pattern of persistence and spread of the transgene in a wild population may vary significantly.

Expression of an incorporated transgene may considerably alter the populations of wild/weedy relatives in terms of their survival, competition, and/or reproduction. These changes may affect the persistence and spread of transgenes in wild/weedy populations in a spatial or temporal dimension. It is very important to establish experimental populations including crop–wild/weedy hybrids and descendants for measuring fitness effects of the transgene that transferred to crop–wild/weedy relatives, for example to produce F_1 hybrids and their F_2, F_3 descendants by self-pollination, and backcross (BC_1) descendants by backcrossing with

or without a transgene to meet the objectives (Fig. 1.6). Such experimental populations can be established through artificial crosses and successive backcrosses between a GE crop and its wild/weedy relatives, which will facilitate data generation for any fitness analysis required during the biosafety assessment.

However, it remains challenging to measure any fitness changes appropriately (sometimes the change can be very minor) of wild plants brought about by the expression of a transgene, and to incorporate the collected fitness data properly into the biosafety assessment system. A well-designed fitness study can take a very long time to complete, especially when data may be required from multiple generations. It is important to point out that fitness is a measurement of the successful survival and reproduction of wild plants, which can be affected by many components throughout the life cycle in the given environment, such as seed dormancy and germination, seedling establishment and vegetative growth, individual viability and fecundity. Even though the determination of an adverse effect and the eventual success of a risk assessment largely relies on the prediction of any fitness change caused by transgenes, it is not straightforward to identify the components crucial for accurately predicting the fitness of recipient wild plants. Therefore, a few aspects need to be taken into consideration when studying fitness for the biosafety assessment.

The usual way to estimate any fitness change is to examine vegetative and reproductive productivity of crop–wild hybrids (mostly the early generations of hybrids), because morphological and reproductive traits appear to be more directly related to the number of offspring that an individual can potentially produce (Song *et al.*, 2004b; Mercer *et al.*, 2007; Xia *et al.*, 2016; Yan *et al.*, 2017). However, the direct measurement of fitness change in F_1 hybrids poses some concerns on the correct judgement of a transgene in crop–wild hybrids, because the role of lifetime fitness, competitive advantage at the specific growth stages, and trade-offs between different components of fitness is still unclear (Jenczewski *et al.*, 2003). It is recognized that crop–wild hybrids will not be genetically uniform, due to variation in wild populations

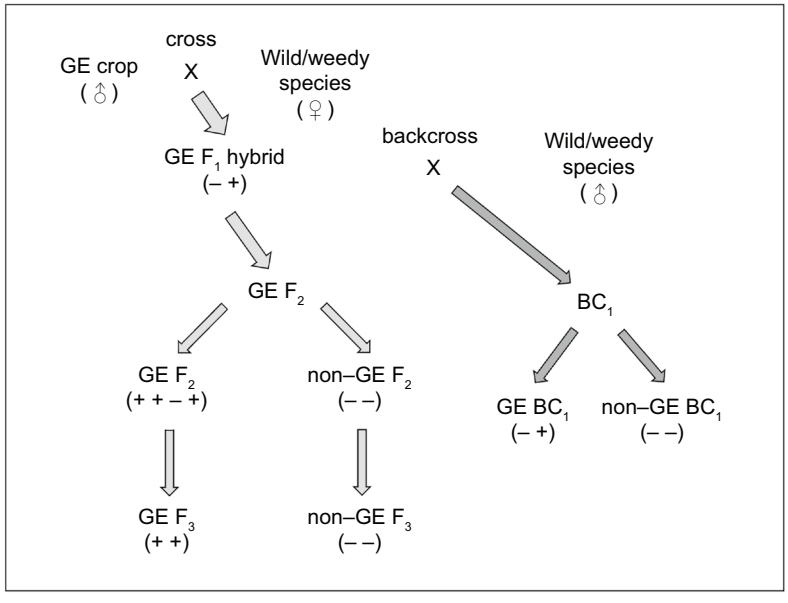


Fig. 1.6. Procedure to create GE (+) and non-GE (–) hybrid lineages (F_1–F_3), and GE and non-GE backcross (BC_1) lineages derived from an artificial cross between a GE crop and wild/weedy relative species for the assessment of fitness effect of wild/weedy populations acquiring a transgene caused by gene flow.

(Linder and Schmitt, 1994; Xia *et al.*, 2016), leading to difficulties in generalizing the results from studies of only F_1 hybrids. According to Linder *et al.* (1998) and Li *et al.* (2016), estimates of early generations of crop–wild hybrid fitness may be of little predictive value for the assessment of transgene establishment.

Sometimes, the F_1 crop–wild hybrids demonstrate enhanced vegetative vigor that contributes to their total fitness, but this may not be of much use in predicting any ecological consequences over a long period of time. It is also important to measure any long-term persistence and spread of crop genes (transgenes) in wild species after crop–wild hybridization and introgression occurs (Campbell and Snow, 2007). The direct measurement of crop gene establishment in early hybrids raises questions about distinguishing introgressive markers from ones jointly inherited from a common ancestor. Studies have therefore concentrated more on analyzing the successive steps in the process of transgene establishment. When designing a study to estimate any fitness change resulting from transgene flow, it is necessary to establish experimental populations of different generations of crop–wild hybrid progenies.

Evaluating the long-term spread and persistence of a transgene in wild/weedy populations also requires an understanding of whether possession of the transgene imposes a cost (fitness penalty) on the wild plants in the absence of a selection pressure on the transgene. It is also important to know whether the transgene confers a trait with selective advantage or if it is selectively neutral. Transgenes with different selective values (advantageous, neutral, or disadvantageous) are expected to spread in wild populations at quite different rates. In addition, apart from the balance between the fitness costs and benefits of a transgene brought to wild plants under natural selection, the persistence and spread of a transgene in wild populations will depend on the strength and frequency of transgene flow. Strong and recurrent transgene flow can be sufficient to establish transgenes in populations, even though the transgenes may contribute to slight fitness disadvantages. The frequency of transgenes can be expected to increase in the wild populations if crop-to-wild transgene flow is significantly strong and frequent. Taken together, all these factors will significantly affect the dynamics of a wild/weedy population that has acquired a transgene or transgenes through pollen-mediated gene flow, significantly influencing the risk assessment procedure.

When all the steps of the risk assessment procedure are completed, through published literature consultation and data collection from actually designed experiments, it should be possible to make a conclusion with a high degree of confidence concerning the environmental impacts caused by the transgene flow from a GE crop to wild/weedy relative species. The risk assessment exercise not only provides us with a tool to determine the possibility of transgene flow from a GE crop to the wild relatives, but also allows the appropriate measurement of adverse effects caused by a particular transgene that has been incorporated into a population of wild relatives. This will help to facilitate decision making concerning an application for environmental release and commercial production of a GE crop in a particular region under unique environmental conditions.

1.6 Case Studies of Crop-to-Wild/Weedy Transgene Flow

1.6.1 Insect-resistant transgenic plants

The insect-resistance transgene (*Bt*, *Bacillus thuringiensis*) is one of the earliest transgenes that are transferred into crop species. Insect-resistant GE crops are also one of the most extensively commercially cultivated in the world (Lu *et al.*, 2012; ISAAA, 2019). Therefore, understanding the balance between fitness benefits and possible underlying costs that are associated with transgenic cultivars should be very useful for evaluating crop performance and inferring the fitness effect of crop–wild/weedy hybrid progeny under the scenario of transgene flow. However, only a few researchers have tested for such benefits and costs under rigorous experimental conditions.

Chen *et al.* (2006) examined changes in net benefits and costs of insect-resistance transgenes in cultivated rice (*Oryza sativa*) by using two levels of insect pressure (low versus moderate) and two types of competition (pure lines versus mixed

lines). The authors compared the growth and fecundity of potted rice plants from three transgenic lines: *Bt* (*cryIAc*), *CpTI* (cowpea trypsin inhibitor) with a single transgene, and *Bt*/*CpTI* with tightly linked two genes, relative to isogenic control plants at outdoor locations in Fuzhou, China. Results from the experiments showed that, under moderate insect pressure, *Bt* and *Bt*/*CpTI* plants produced 36–65% more seeds than the non-transgenic controls. In contrast, under low insect pressure, *Bt* and *Bt*/*CpTI* plants produced 16% fewer tillers, 6% smaller seeds, and 30% fewer seeds than their non-transgenic control plants. These results in general indicated that the insect-resistance transgenes brought about a certain degree of fitness benefits to the rice plants under moderate insect pressure, particularly for the double-transgene *Bt*/*CpTI* lines, but a low degree of costs under low insect pressure. In a similar study of insect-resistant transgenic rice (*Bt*, *cryIAc*) at a more extensive scale of 2-year field experiments, Xia *et al.* (2016) also reported the detectable yield benefit and underlying cost of the rice lines containing the insect-resistance transgene under different insect pressure.

Transgene flow from GE crops allows novel traits to spread to sexually compatible wild/weedy species. Insect-resistance transgenes may enhance the fitness of wild/weedy populations. To test for the fitness effects under natural field conditions, Cao *et al.* (2009) and Yang *et al.* (2011) created F_1, F_2, and F_3 crop–weedy rice hybrid lineages of GE rice, using lines with two transgene constructs (*CpTI* and *Bt*/*CpTI*). Comparison between three weedy rice populations and their F_1 hybrid lines containing the two transgenes indicated an enhanced relative performance of the crop–weed hybrids, with taller plants and more tillers, panicles, and spikelets per plant, as well as higher 1000-seed weight in the GE hybrid lines, although the hybrids produced fewer filled seeds (Cao *et al.*, 2009).

Further common-garden experiments, including the F_2 and F_3 crop–weedy hybrid lineages with or without the insect-resistance transgenes (*CpTI* and *Bt*/*CpTI*), showed variable fitness effects to the hybrid descendants. The *CpTI* transgene alone did not show significant effects on fecundity, although this transgene could reduce herbivory to some extent. In contrast, under certain conditions, the *Bt*/*CpTI* transgenic hybrid lineages conferred up to ~79% less insect damage and ~47% greater fecundity relative to non-transgenic controls, in addition to a ~44% increase in fecundity relative to the weedy parent. A small fitness cost was detected in F_3 lineages with the *Bt*/*CpTI* transgenes when the hybrid plants were grown under low insect pressure and direct competition with transgene-negative controls. Based on these results, the authors concluded that the fitness effect of transgenes in crop–weedy hybrids was considerably different, and the *Bt*/*CpTI* transgenes could introgress into co-occurring weedy rice populations and contribute to greater seed production when target insects are abundant. However, the fitness benefits that were associated with the *Bt*/*CpTI* transgenes could be ephemeral if insect pressure is lacking, for example because of widespread planting of *Bt* cultivars that suppress target insect populations (Yang *et al.*, 2011).

It is clear that different types of insect-resistance transgenes (e.g. *Bt* and *CpTI*) introgressed from a GE crop into populations of weedy or wild relatives can provide natural selective advantages under insect pressure. Apart from natural selection in variable environments, the advantages may likely be associated with genetic background of transgene recipients, meaning different genotypes of the weedy or wild populations. To explore the role of the environment and background of transgene recipients in affecting the advantages, Xia *et al.* (2016) estimated the fitness of F_1 and F_2 crop–weed hybrid lineages derived from crosses between insect-resistant transgenic (*Bt*/*CpTI*) rice with five weedy rice populations from different geographical locations under varied insect pressure. Results indicated significant effects of both transgenes and weedy rice genotypes (populations) on the performance of crop–weed hybrid lineages in the high-insect environment. Increased fecundity was detected in most transgene-present F_1 and F_2 hybrid lineages under high insect pressure, but varied among crop–weed hybrid lineages with different weedy rice parents. Increased fecundity of transgenic crop–weed hybrid lineages was associated with the environmental insect pressure and genotypes of their weedy rice parents. The findings suggest that the fitness effects of an insect-resistant transgene introgressed into weedy populations are not

uniform across different environments and genotypes/populations of the recipient plants that have acquired the transgenes (Xia *et al.*, 2016). Therefore, factors such as transgenes, environmental conditions, and genotypes of weedy (and likely wild) populations should be considered when assessing the environmental impacts of transgene flow to weedy or wild rice relatives.

A similar pattern of fitness effects was also found in crop–wild rice hybrid descendants with or without the insect-resistance transgenes when a common-garden experiment was conducted to examine the environmental impacts from transgene flow from GE rice lines to wild rice populations (Li *et al.*, 2016). In the experiment, the authors produced F_1 and F_2 hybrid descendants from crosses of two insect-resistant GE rice lines (*Bt*, *Bt*/*CpTI*) and their non-GE rice parent with a wild rice (*Oryza rufipogon*) population to estimate transgenic fitness. Insect damages and life-cycle fitness of transgenic and non-transgenic crop–wild hybrid descendants, as well as their wild parent, were examined in the experiment. No significant differences in insect damages were observed between the wild rice parent and transgenic hybrid descendants under high insect pressure, suggesting a comparably high level of insect resistance of the wild rice population and transgenic hybrids. The wild rice parent showed significantly greater relative survival–regeneration ratios than its transgenic and non-transgenic hybrid descendants under both high and low insect pressure. However, more seeds were produced in transgenic hybrid descendants than their non-transgenic counterparts under high insect pressure (Li *et al.*, 2016). Based on the fact that the introduction of *Bt* and *Bt*/*CpTI* transgenes did not provide greater insect resistance to the crop–wild hybrid descendants than their wild parent, the authors concluded that transgene flow from GE insect-resistant rice to wild rice populations may not cause considerable environmental impacts.

In summary, the transfer of insect-resistance transgenes from a GE rice variety to the populations of its wild and weedy rice relatives through pollen-mediated gene flow may not result in considerable ecological or environmental impacts. This prediction is based on two facts generated from the science-based experiment. Firstly, an insect-resistance transgene does not confer considerable net fitness benefits to the wild/weedy rice populations under low insect pressure. In fact, insect pressure is generally low in wild rice populations in natural habitats with much less population density, compared with the situation in the intensively cultivated agricultural fields. In addition, if insect-resistant GE rice will be extensively cultivated, insect pressure will be low in the agroecosystem. As a consequence, weedy rice plants that have acquired an insect-resistance transgene and that co-occur with GE insect-resistant rice varieties may not have obvious fitness benefit. Secondly, an insect-resistance transgene may not provide additional insect resistance to wild rice populations, because the wild plants are commonly more resistant to insect attacks than their cultivated descendants.

1.6.2 Herbicide tolerance transgenic plants

It is clearly shown, based on recent studies, that the extent of environmental impacts of transgene flow will vary greatly, depending on the types of transgenes. Whether or not other transgenes (e.g. herbicide tolerance) transferred from crops to cross-compatible wild/weedy relatives will result in negative environmental impacts needs further investigation. Again, it is important for the environmental biosafety assessment to measure the fitness change in wild/weedy populations with an introgressed transgene. Crops containing a herbicide-tolerance transgene are the most extensively applied and cultivated GE crops in the world. Therefore, understanding the evolutionary and ecological consequences of herbicide-tolerance transgenes acquired by wild/weedy populations through pollen-mediated gene flow is necessary.

In rice, transgenic herbicide resistance is expected to spread to conspecific weedy rice through hybridization or gene flow. To measure the environmental impacts of a herbicide-resistance transgene, Wang *et al.* (2014) studied fitness effects of a herbicide resistance that overexpressed an endogenous 5-enolpyruvoylshikimate-3-phosphate synthase (*EPSPS*) transgene developed to confer glyphosate herbicide resistance in rice. Controlling for genetic background, the authors examined physiological traits and field

performance of the crop–weed hybrid lineages that segregated for the presence or absence of this *EPSPS* transgene. Surprisingly, the authors found that the transgenic F_2 crop–weed hybrid lineages produced 48–125% more seeds per plant than the non-transgenic controls (F_2) in monoculture- and mixed-planting designs without glyphosate application. Transgenic plants also had greater EPSPS protein levels, tryptophan concentrations, photosynthetic rates, and percentage seed germination compared with the non-transgenic controls. These findings suggest that over-expression of a modified endogenous rice *EPSPS* gene can lead to significant fitness advantages, even without exposure to the glyphosate herbicide (Wang *et al.*, 2014).

Further studies involving the same transgenic weedy rice populations indicated that overexpressing the *EPSPS* transgene can significantly increase the protoplast contents and the expression level of some protoplast-synthesis related genes (Ma *et al.*, 2020). In addition, Wu *et al.* (2020) found that overexpressing the *EPSPS* transgene can increase the total content of lignin and the expression level of some lignin-synthesis related genes in transgenic weedy rice hybrid lineages. The authors therefore predicted that the transfer of a herbicide-tolerant transgene by the mechanism of overexpressing the *EPSPS* gene to weedy rice may cause considerable negative environmental impacts (Wang *et al.*, 2014; Ma *et al.*, 2020; Wu *et al.*, 2020). This is because the substantially enhanced fitness benefits will speed up the spread of the *EPSPS* transgene in weedy rice populations following transgene introgression.

Similarly, Yan *et al.* (2017) found substantial alteration of phenology and increases in fitness-related traits in F_1–F_3 crop–wild hybrid descendants derived from crosses between an *EPSPS* GE rice line and two endangered wild rice (*Oryza rufipogon*) populations, based on the common-garden field experiments. Under the glyphosate-free condition, transgenic hybrid lineages showed significantly earlier tillering and flowering time than their non-transgenic hybrid controls. Also, the transgenic hybrid lineages showed increased fecundity and overwintering survival/regeneration abilities, compared with the non-transgenic hybrid controls. Furthermore, a negative correlation was detected between the content of the endogenous EPSPS protein of wild, weedy, and cultivated rice parents and fitness differences caused by the incorporation of the *EPSPS* transgene. In other words, a lower level of the endogenous EPSPS protein in the transgene-recipient populations displayed a more pronounced enhancement in fitness. The altered phenology and enhanced fitness of crop–wild hybrid offspring by the *EPSPS* transgene may cause unwanted environmental impacts when this type of glyphosate-resistant transgene introgresses into wild rice populations through gene flow.

In summary, unlike the insect-resistance transgenes, our understanding of the environmental impacts caused by herbicide-tolerant transgene flow is very limited. This is because of the general expectation that herbicides will not be applied to natural habitats where wild relatives occur, and without selection pressure the introgressed herbicide-tolerant transgenes may not cause any problems in natural habitats. However, recent studies have indicated that the transfer of a herbicide-tolerant transgene, particularly by the engineered overexpressing *EPSPS* genes, from a GE rice variety to the populations of its wild and weedy rice relatives through pollen-mediated gene flow may cause considerable environmental impacts by the unexpected fitness alteration, even when there is no herbicide application.

1.7 Transgene Inheritance and Non-random Transmission of Parental Alleles in Hybrid Lineages

As mentioned previously, it is necessary to assess environmental impacts caused by transgene flow from any GE crops to their wild/weedy relatives before the commercialization of GE crops. As a consequence, many studies are conducted to determine the environmental impacts of transgene flow by measuring the fitness effects of a transgene (Lu and Snow, 2005; Lu, 2008; Lu *et al.*, 2016). In these studies, it was found that a transgene usually transmitted into the hybrid descendants following Mendel's law of genetic inheritance. Based on such a pattern of transgene inheritance, the performance of fitness-related trials (such as vegetative growth and seed production) in comparative transgenic

and non-transgenic hybrid lineages (see Figs 1.5 and 1.6) is measured as key for the risk assessment of environmental impacts from transgene flow.

Usually, the transgenic and non-transgenic hybrid lineages used in the risk assessment are sampled from the same interspecific crosses between a target GE transgenic crop and its wild/weedy relative, and the resulting hybrids are self-pollinated or backcrossed to obtain advanced hybrid lineages. In the comparative measurement of fitness effects, segregated isogenic hybrid lineages or subpopulations with or without a transgene in the same generation (e.g. F_2 or F_3) of hybrid lineages or subpopulations are included for analysis (see Fig. 1.6). The analysis of fitness effect is based on the general recognition that the transgene-positive and transgene-negative hybrid lineages or subpopulations used in the fitness analysis should share the same genetic background, and they only differ in having or not having the transgene (Fig. 1.6).

Unexpectedly, Wang *et al*. (2017) detected substantial differences in morphological characteristics (phenotypes) between transgene-present and transgene-absent hybrid lineages (F_2–F_3), sampled from the same hybrid offspring populations (Fig. 1.7). These hybrid descendants (F_2–F_3) were developed from an artificial cross of a GE rice line containing a glyphosate-tolerant transgene (5-enolpyruvoylshikimate-3-phosphate synthase, *EPSPS*) with a wild (*Oryza rufipogon*) plant (Wang *et al*., 2017; Yan *et al*., 2017). Further studies indicated that such phenotypical differentiation is attributed to the increased frequencies of crop-parent alleles in transgenic hybrid lineages at multiple loci across the genome, as estimated by the neutral SSR molecular markers. This means that the alleles from the two parents are not randomly transmitted to their transgene-present and transgene-absent hybrid lineages used for transgenic fitness estimate.

Such preferential transmission of parental alleles was also found in equally sampled crop–wild/weed hybrid lineages with or without a particular neutral SSR identifier used to establish the identifier-present and identifier-absent hybrid lineages. The authors concluded that selecting either a transgene or neutral marker as an identifier to create hybrid lineages will lead to differences in phenotypes and genomic backgrounds between the hybrid lineages due to non-random transmission of parental alleles (Wang *et al*., 2017). Non-random allele transmission may misrepresent the outcomes of transgenic fitness effects for the assessment of environmental impacts caused by transgene flow. It is thus necessary to seek other means to evaluate fitness effects of transgenes for assessing environmental impacts caused by crop-to-wild/weed gene flow.

1.8 Mitigating Environmental Impacts Caused by Transgene Flow

It is well recognized that completely restricting transgene flow from a GE crop to its wild/weedy relatives is not always a desirable and straightforward strategy. This is particularly true for crop species such as rice, sugar beet, and sunflowers that have conspecific weeds, as well as for those that have closely related weedy species, such as oilseed rape, sorghum, wheat, barley, and maize that co-occur with their crops in the same fields or in the close vicinity (Lu, 2008). Although there are a number of confinement methods that have been developed or proposed for the purpose of minimizing the outflow of transgenes into wild/weedy relatives, a certain level of transgene flow (leakage) is always inevitable (Lu, 2008). Actually, the confinement of pollen-mediated crop-to-wild/weedy transgene flow is nearly impossible in practice. Accordingly, a strategy to mitigate the impacts of transgene escape has been proposed, in situations where completely restricting gene flow is impossible (Gressel, 1999; Lu, 2008).

Transgenic mitigation (TM) has received increasing attention as an approach for minimizing the persistance and spread of a transgene in wild/weedy populations by compromising the fitness of wild/weedy plants that receive positive survival traits from crop genes through introgression (Gressel, 1999; Hani et al., 2005; Al-Ahmad and Gressel, 2006; Al-Ahmad *et al*., 2006). In this concept, the so-called mitigating (mitigator) transgenes are introduced into a GE crop and tandemly linked to the primary desired transgene(s). The ‘mitigator’ transgenes would specifically reduce the fitness of any hybrids and

Fig. 1.7. Phenotypic variation of seedlings in F_3 transgene-present (dark arrow) and transgene-absent (white arrow) hybrid lineages derived from the same artificial cross between an *EPSPS* transgenic rice line and wild rice (*Oryza rufipogon*).

their progenies resulting from pollen-mediated transgene flow, considerably reducing any negative environmental consequences. According to Gressel (1999; 2002), the TM approach is based on the premises that: (i) tandem constructs act as tightly linked transgenes with exceedingly rare segregation from each other; (ii) the selected TM traits are neutral or favorable to crops, but deleterious to non-crop progeny, such as wild/weedy hybrid descendants; and (iii) individuals bearing even mildly harmful TM traits will remain at very low frequencies in weed/wild populations because weeds typically have very high seed production and strongly compete among themselves, eliminating even marginally unfit individuals.

Therefore, if the target transgene providing the agricultural advantage is flanked in a tandem construct by mitigator transgenes such as dwarfing, uniform seed ripening, seed persisting (non-shattering), anti-secondary dormancy, or non-bolting genes, the overall effect would be deleterious after introgression into wild or weedy relatives, as the TM genes will reduce the fitness and competitive ability of the transgenic hybrids/descendants. As a consequence, these hybrids or hybrid descendants will compete poorly with normal wild plants, and therefore the transgenes can persist in only low frequencies in agricultural ecosystems. Some studies are conducted applying the TM technology to reduce general fitness of crop–wild/weedy hybrid descendants in crops of tobacco, oilseed rape, and rice, involving the unfit dwarf and non-shattering gene as the mitigator (Al-Ahmad and Gressel, 2006; Al-Ahmad *et al.*, 2006; Yan *et al.*, 2017).

There are, however, some concerns remaining over the use of transgenic mitigation. For example, the technology cannot solve the problems of massive amounts of transgenes moving into wild or weedy populations through recurrent transgene flow. The destiny and long-term consequences of the mitigator genes in crop and weedy or wild populations are still unpredictable (Lu, 2008; Yan *et al.*, 2017). In addition, establishing tandem constructs with tightly linked genes will require considerable efforts for multigene engineering. Future attempts to transfer multiple genes with different traits into one crop

variety may be a challenge unless transgenes are genetically linked within the same construct. The long-term ecological impacts resulting from these transgenes as a whole package in the environment is as yet unknown and will be difficult to predict.

Although still being discussed and argued, TM brings new insight for effective management of transgene flow and its environmental consequences by mitigating the risks to a minimum level. Probably, there is no single approach that can be very effective to confine transgene escape to wild relatives and to mitigate the consequences from such an escape. Also, it is not necessary to put the same effort to confine all the transgenes from all GE crops under all environmental conditions where GE crops will be released for cultivation. This is because many transgenes that do not provide a selective advantage to the host plant in nature may not pose any environmental consequences, and many crops that have extremely low gene flow frequencies (e.g. some legume species) already have a low risk of transgene escape. In addition, for some geographical locations where wild relatives or conspecific weeds of the GE crops are absent, transgene escape through pollen-mediated gene flow would not be an issue. A strategic combination of transgene confinement from gene flow and mitigation to minimize its impacts in a particular circumstance should provide an effective strategy to manage any environmental consequences caused by transgene escape.

1.9 Future Perspectives

Environmental impacts caused by pollen-mediated transgene flow from a GE crop to its wild/weedy relatives remains a concern for global commercial cultivation of GE crops. While gene flow is a natural process that has existed for millions of years since the presence of higher plants, human activities, particularly agricultural practices, have significantly influenced gene flow between domesticated and wild/weedy plants for thousands of years. Gene flow, including transgene flow to non-GE crops and populations of wild/weedy relatives, may not be preventable in some species, as long as crops are cultivated in agroecosystems. It is also necessary to point out that gene flow per se is not a risk. Therefore, identifying whether or not transgene flow can bring about potential environmental impacts is the key. The fundamental scientific questions that are closely associated with environmental impacts of transgene flow should be identified and addressed following science-based, case-by-case, and step-by-step principles.

Based on the framework of risk assessment for transgene flow and its potential environmental impacts, some key scientific questions and their underlying mechanisms have been revealed and addressed. Among these, knowledge has been generated regarding the frequencies of transgene flow for the world's major crops and transgene expression in wild/weedy relatives. However, accurate knowledge associated with fitness effects of different types of transgenes on their populations and species of wild/weedy relatives under the complex environmental conditions still needs to be generated by more investigations. Importantly, the fitness effects conferred by a transgene will greatly influence the survival and reproduction of wild/weedy relative plants, which can directly affect the assessment of environment impacts from transgene flow.

It is necessary to emphasize the implementation of in-depth research on the transgenic fitness and its associated factors, such as transgenes and transgenic events, and genetic background of recipient wild/weedy populations or species, in addition to environmental conditions and the interaction of all the above factors. Understanding the role of each of these factors and their interactions in the change of fitness and evolutionary potential by incorporating a transgene will facilitate the appropriate risk assessment of the transgene that transferred to wild/weedy populations through gene flow, particularly for the long-term environmental impacts.

References

Al-Ahmad, H. and Gressel, J. (2006) Mitigation using a tandem construct containing a selectively unfit gene precludes establishment of *Brassica napus* transgenes in hybrids and backcrosses with weedy *Brassica rapa*. *Plant Biotechnology Journal* 4(1), 23–33. DOI: 10.1111/j.1467-7652.2005.00153.x.

Al-Ahmad, H., Dwyer, J., Moloney, M. and Gressel, J. (2006) Mitigation of establishment of *Brassica napus* transgenes in volunteers using a tandem construct containing a selectively unfit gene. *Plant Biotechnology Journal* 4, 7–21. DOI: 10.1111/j.1467-7682.2005.00152.x.

Alcazar, P. and Comtois, P. (2000) The influence of sampler height and orientation on airborne Ambrosia pollen counts in Montreal. *Grana* 39, 303–307.

Alcazar, P., Galan, C., Carinanos, P. and Dominguez-Vilches, E. (1998) Vertical variation in Urticaeae airborne pollen concentration. *Aerobiologia* 14, 131–134.

Bashir, K., Husnain, T., Fatima, T., Riaz, N., Makhdoom, R. *et al*. (2005) Novel indica Basmati line (B-370) expressing two unrelated genes of *Bacillus thuringiensis* is highly resistant to two lepidopteran insects in the field. *Crop Protection* 24, 870–879.

Campbell, L.G. and Snow, A.A. (2007) Competition alters life history and increases the relative fecundity of crop–wild radish hybrids (*Raphanus spp*.). *New Phytologist* 173, 648–660.

Cao, Q.-J., Yang, X., Yang, X. and Lu, B.-R. (2009) Performance of hybrids between weedy rice and insect-resistant transgenic rice under field experiments: implication for environmental biosafety assessment. *Journal of Integrative Plant Biology* 51(12), 1138–1148. DOI: 10.1111/j.1744-7909.2009.00877.x.

Chen, L.-Y., Snow, A.A., Wang, F. and Lu, B.-R. (2006) Effects of insect-resistance transgenes on fecundity in rice (*Oryza sativa*, Poaceae): a test for underlying costs. *American Journal of Botany* 93, 94–101.

Chen, L.-J., Lee, D.S., Song, Z.-P., Suh, H.-S. and Lu, B.-R. (2004) Gene flow from cultivated rice (*Oryza sativa*) to its weedy and wild relatives. *Annals of Botany* 93(1), 67–73. DOI: 10.1093/aob/mch006.

Dalecky, A., Bourguet, D. and Ponsard, S. (2007) Does the European corn borer disperse enough for a sustainable control of resistance to *Bt* maize via the High Dose/Refuge strategy? *Agricultures* 16, 171–176.

Ellstrand, N.C. (2001) When transgenes wander, should we worry? *Plant Physiology* 125(4), 1543–1545. DOI: 10.1104/pp.125.4.1543.

Ellstrand, N.C. (2003) Current knowledge of gene flow in plants: implications for transgene flow. *Philosophical Transactions of the Royal Society B: Biological Sciences* 358(1434), 1163–1170. DOI: 10.1098/rstb.2003.1299.

Ellstrand, N.C. and Elam, D. (1993) Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology Evolution and Systematics* 24, 217–242.

Ellstrand, N.C., Prentice, H.C. and Hancock, J.F. (1999) Gene flow and introgression from domesticated plants into their wild relatives. *Annual Review of Ecology and Systematics* 30, 539–563.

Ellstrand, N.C., Meirmans, P., Rong, J., Bartsch, D. and Ghosh, A. (2013) Introgression of crop alleles into wild or weedy populations. *Annual Review of Ecology, Evolution, and Systematics* 44, 325–345.

Gealy, D.R., Mitten, D.H. and Rutger, J.N. (2003) Gene flow between red rice (*Oryza sativa*) and herbicide-resistant rice (*O. sativa*): implications for weed management. *Weed Technology* 17, 627–645.

Gressel, J. (1999) Tandem constructs: Preventing the rise of superweeds. *Trends in Biotechnology* 17(9), 361–366. DOI: 10.1016/s0167-7799(99)01340-2.

Gressel, J. (2002) *Molecular Biology of Weed Control*. Taylor & Francis, London.

Hall, L., Topinka, K., Huffman, J., Davis, L. and Good, A. (2000) Pollen flow between herbicide-resistant *Brassica napus* is the cause of multiple-resistant *B. napus* volunteers. *Weed Science* 48, 688–694.

Hani, A.A., Galili, S. and Gressel, J. (2005) Poor competitive fitness of transgenically mitigated tobacco in competition with the wild type in a replacement series. *Planta* 222(2), 372–385. DOI: 10.1007/s00425-005-1540-6.

Hartl, D.L. and Clark, A.G. (1989) *Principle of Population Genetics*, 2nd edn. Sinauer Associates Inc, Sunderland, Massachusetts.

ISAAA (2019) *Global Status of Commercialized Biotech/GM Crops in 2019*, ISAAA Brief No. 55. International Service for the Acquisition of Agri-biotech Applications, Ithaca, New York.

Jenczewski, E., Ronfort, J. and Chèvre, A.-M. (2003) Crop-to-wild gene flow, introgression and possible fitness effects of transgenes. *Environment Biosafety Research* 2(1), 9–24. DOI: 10.1051/ebr:2003001.

Kiang, Y.-T., Antonvics, J. and Wu, L. (1979) The extinction of wild rice (*Oryza perennis formosa*) in Taiwan. *Journal of Asian Ecology* 1, 1–9.

Koopman, W.J.M., Li, Y.-H., Coart, E., van de Weg, W.E., Vosman, B. *et al*. (2007) Linked vs. unlinked markers: multilocus microsatellite haplotype-sharing as a tool to estimate gene flow and introgression. *Molecular Ecology* 16(2), 243–256. DOI: 10.1111/j.1365-294X.2006.03137.x.

Li, G.P., Wu, K.M., Gould, F., Wang, J.K. and Miaoi, J. (2007) Increasing tolerance to Cry1Ac cotton from cotton bollworm, *Helicoverpa armigera*, was confirmed in *Bt* cotton farming area of China. *Ecological Entomology* 32, 366–375.

Li, L., Yang, X., Wang, L., Yan, H. and Su, J. (2016) Limited ecological risk of insect-resistance transgene flow from cultivated rice to its wild ancestor based on life-cycle fitness assessment. *Science Bulletin* 61, 1440–1450.

Linder, C.R. and Schmitt, J. (1994) Assessing the risks of transgene escape through time and crop–wild hybrid persistence. *Molecular Ecology* 3, 23–30.

Linder, C.R., Taha, I., Rieseberg, L.H., Seiler, G.J. and Snow, A.A. (1998) Long-term introgression of crop genes into wild sunflower populations. *Theoretical and Applied Genetics* 96(3-4), 339–347. DOI: 10.1007/s001220050746.

Lu, B.-R. (2008) Transgene escape from GM crops and potential biosafety consequences: an environmental perspective. *International Centre for Genetic Engineering and Biotechnology (ICGEB), Collection of Biosafety Reviews* 4, 66–141.

Lu, B.-R. (2016) Challenges of transgenic crop commercialization in China. *Nature Plants* 2(6), 1–2. DOI: 10.1038/nplants.2016.77.

Lu, B.-R. and Snow, A.A. (2005) Gene flow from genetically modified rice and its environmental consequences. *BioScience* 55, 669–678.

Lu, B.-R. and Yang, C. (2009) Gene flow from genetically modified rice to its wild relatives: assessing potential ecological consequences. *Biotechnology Advances* 27, 1083–1091.

Lu, B.-R., Yang, X. and Ellstrand, N.C. (2016) Fitness correlates of crop transgene flow into weedy populations: a case study of weedy rice in China and other examples. *Evolutionary Applications* 9(7), 857–870. DOI: 10.1111/eva.12377.

Lu, Y.-H., Wu, K.-M., Jiang, Y.-Y., Guo, Y.-Y. and Desneux, N. (2012) Widespread adoption of *Bt* cotton and insecticide decrease promotes biocontrol services. *Nature* 487(7407), 362–365. DOI: 10.1038/nature11153.

Ma, T.-J., Yuan, Y., Ling, L.-Q., Nan, P. and Lu, B.-R. (2020) Overexpression of *EPSPS* transgene enhances chlorophyll synthesis in hybrid progeny between cultivated rice and weedy rice. *Fudan University Journal of Natural Sciences* 59(2), 185–193. [In Chinese with English abstract.]

Mercer, K.L., Andow, D.A., Wyse, D.L. and Shaw, R.G. (2007) Stress and domestication traits increase the relative fitness of crop-wild hybrids in sunflower. *Ecology Letters* 10(5), 383–393. DOI: 10.1111/j.1461-0248.2007.01029.x.

Messeguer, J., Fogher, C., Guiderdoni, E., Marfa, V. and Catala, M.M. (2001) Field assessment of gene flow from transgenic to cultivated rices (*Oryza sativa* L.) using an herbicide resistance gene as tracer marker. *Theoretical and Applied Genetics* 103, 1151–1159.

Oliveira, A.R., Castro, T.R., Capalbo, D.M.F. and Delalibera, I. (2007) Toxicological evaluation of genetically modified cotton (Bollgard®) and Dipel ® WP on the non-target soil mite *Scheloribates praeincisus* (Acari: Oribatida). *Experimental and Applied Acarology* 41(3), 191–201. DOI: 10.1007/s10493-007-9059-0.

Ortiz-Garcia, S., Ezcurra, E., Schoel, B., Acevedo, F., Soberon, J. *et al*. (2005) Absence of detectable transgenes in local landraces of maize in Oaxaca, Mexico (2003–2004). *Proceedings of the National Academy of Sciences of the United States of America* 102(35), 12338–12343. DOI: 10.1073/pnas.0503356102.

Quist, D. and Chapela, I.H. (2001) Transgenic DNA introgressed into traditional maize landraces in Oaxaca, Mexico. *Nature* 414(6863), 541–543. DOI: 10.1038/35107068.

Rognli, O.A., Nilsson, N.O. and Nurminiemi, M. (2000) Effects of distance and pollen competition on gene flow in the wind-pollinated grass *Festuca pratensis* Huds. *Heredity* 85(Pt 6), 550–560. DOI: 10.1046/j.1365-2540.2000.00789.x.

Rong, J., Xia, H., Zhu, Y.-Y., Wang, Y.-Y. and Lu, B.-R. (2004) Asymmetric gene flow between traditional and hybrid rice varieties (*Oryza sativa*) estimated by nuclear SSRs and its implication in germplasm conservation. *New Phytologist* 163, 439–445.

Rong, J., Song, Z.-P., Su, J., Xia, H., Lu, B.-R. *et al*. (2005) Low frequency of transgene flow from *Bt/CpTI* rice to its non-transgenic counterparts planted at close spacing. *New Phytologist* 168(3), 559–566. DOI: 10.1111/j.1469-8137.2005.01539.x.

Rong, J., Lu, B.-R., Song, Z.-P., Su, J. and Snow, A.A. (2007) Dramatic reduction of crop-to-crop gene flow within a short distance from transgenic rice fields. *New Phytologist* 173, 346–353.

Rong, J., Song, Z.-P., de Jong, T.J., Zhang, X.-S., Sun, S.-G. *et al*. (2010) Modelling pollen-mediated gene flow in rice: risk assessment and management of transgene escape. *Plant Biotechnology Journal* 8(4), 452–464. DOI: 10.1111/j.1467-7652.2009.00488.x.

Saji, H., Nakajima, N., Aono, M., Tamaoki, M., Kubo, A. *et al*. (2005) Monitoring the escape of transgenic oilseed rape around Japanese ports and roadsides. *Environmental Biosafety Research* 4, 217–222. DOI: 10.1051/ebr:2006003.

Serageldin, I. (1999) Biotechnology and food security in the 21st century. *Science* 285(5426), 387–389. DOI: 10.1126/science.285.5426.387.

Song, Z.-P., Lu, B.-R., Zhu, Y.G. and Chen, J.K. (2003) Gene flow from cultivated rice to the wild species *Oryza rufipogon* under experimental field conditions. *New Phytologist* 157(3), 657–665. DOI: 10.1046/j.1469-8137.2003.00699.x.

Song, Z.-P., Lu, B.-R. and Chen, J.-K. (2004a) Pollen flow of cultivated rice measured under experimental conditions. *Biodiversity and Conservation* 13, 579–590.

Song, Z.-P., Lu, B.-R., Wang, B. and Chen, J.-K. (2004b) Fitness estimation through performance comparison of F_1 hybrids with their parental species *Oryza rufipogon* and *O. sativa*. *Annals of Botany* 93(3), 311–316. DOI: 10.1093/aob/mch036.

Stewart, C.N., Halfhill, M.D. and Warwick, S.I. (2003) Transgene introgression from genetically modified crops to their wild relatives. *Nature Reviews Genetics* 4(10), 806–817. DOI: 10.1038/nrg1179.

Thomson, J.A. (2001) Horizontal transfer of DNA from GM crops to bacteria and to mammalian cells. *Journal of Food Science* 66, 188–193.

Wang, F., Yuan, Q.-H., Shi, L., Qian, Q., Liu, W.-G. *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 Biotechnology Journal* 4(6), 667–676. DOI: 10.1111/j.1467-7652.2006.00210.x.

Wang, L., Haccou, P. and Lu, B.-R. (2016) High-resolution gene flow model for assessing environmental impacts of transgene escape based on biological parameters and wind speed. *PLoS ONE* 11(3), e0149563. DOI: 10.1371/journal.pone.0149563.

Wang, W., Xia, H., Yang, X., Xu, T. and Si, H.-J. (2014) A novel 5-enolpyruvoylshikimate-3-phosphate (EPSP) synthase transgene for glyphosate resistance stimulates growth and fecundity in weedy rice (*Oryza sativa*) without herbicide. *New Phytologist* 202, 679–688.

Wang, Z., Wang, L., Wang, Z. and Lu, B.-R. (2017) Non-random transmission of parental alleles into crop-wild and crop-weed hybrid lineages separated by a transgene and neutral identifiers in rice. *Scientific Report* 7(1), 10436. DOI: 10.1038/s41598-017-10596-4.

Wu, J., Fang, J., Cai, X.-X. and Lu, B.-R. (2020) Overexpressing 5-enolpyruvylshikimate-3-phosphate synthase gene increases lignin content of transgenic progeny derived from hybrids of *EPSPS* transgenic rice with weedy and wild rice. *Fudan University Journal of Natural Sciences* 59(6), 666–676. [In Chinese with English abstract.]

Xia, H., Zhang, H., Wang, W., Yang, X., Wang, F. *et al*. (2016) Ambient insect pressure and recipient genotypes determine fecundity of transgenic crop–weed rice hybrid progeny: implications for environmental biosafety assessment. *Evolutionary Applications* 9(7), 847–856. DOI: 10.1111/eva.12369.

Yan, H.-X., Li, L., Liu, P., Jiang, X.-Q., Wang, L. *et al*. (2017) Reduced weed seed shattering by silencing a cultivated rice gene: strategic mitigation for escaped transgenes. *Transgenic Research* 26(4), 465–475. DOI: 10.1007/s11248-017-0016-3.

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 weed-crop hybrids of rice. *Evolutionary Applications* 4(5), 672–684. DOI: 10.1111/j.1752-4571.2011.00190.x.

2 Indirect Methods for Monitoring and Modeling Gene Flow in Natural Plant Populations

María Eugenia Barrandeguy[1,2,3]* and María Victoria García[1,2,3]
[1]Department of Genetics, Faculty of Exact, Chemicals and Natural Sciences., National University of Misiones (FCEQyN – UNaM), Posadas, Misiones, Argentina; [2]Institute of Subtropical Biology – Posadas Node (UNaM – CONICET), Posadas, Misiones, Argentina; [3]National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina

Abstract

Gene flow is a micro-evolutionary process that maintains the allelic exchange among local populations, increasing population genetic diversity. Because of the immobility of plants, pollen plays a major role in connecting extant populations while seeds are necessary to establish and maintain populations of plants.

Despite the widespread use of next-generation sequencing platforms, co-dominant molecular markers, such as microsatellites, are still useful and informative tools in molecular ecology and conservation genetic studies. These markers are currently the most frequently used tools for population genetic studies in plants. These molecular markers reveal information for gene flow estimation by indirect methods and also for comparing the role of gene flow by pollen versus gene flow by seeds in the determination of population genetic structures. Approximate Bayesian computation methods are often used to determine the most probable model of genetic admixture among populations.

Keywords: allelic exchange; gene flow by pollen; gene flow by seed; genetic diversity; microevolutionary process; molecular markers

2.1 Introduction

Gene flow is a broad term that includes all the mechanisms that result in the movement of alleles from one population to another (Slatkin, 1985a). Despite the immobility of plants, their genetic information can move via two specialized vehicles: before fertilization, by means of the male gametophyte (gene flow by pollen); and after fertilization through the young sporophyte (gene flow by seeds). In this way, pollen plays a major role in connecting extant populations, while seeds are necessary to establish plant populations (Petit and Vendramin, 2007). Figure 2.1 shows the allelic exchange between two hypothetical populations of a hermaphrodite tree species. In this case, each individual was genotyped by means of one nuclear microsatellite locus and one chloroplast microsatellite

*Corresponding author: euge_barra@hotmail.com

 Gene Flow: Monitoring, Modeling and Mitigation (eds Wei Wei and C. Neal Stewart, Jr)
DOI: 10.1079/9781789247480.0002

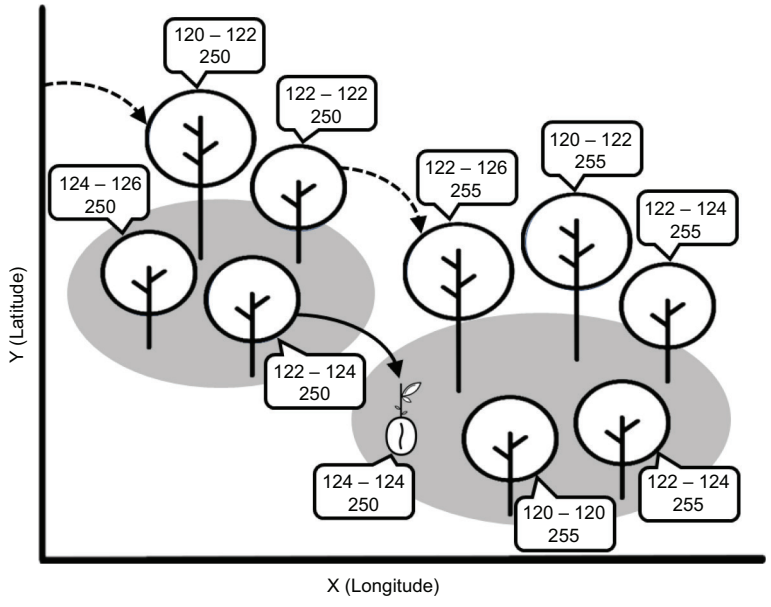


Fig. 2.1. Gene flow by pollen and by seeds between two populations of a hermaphrodite tree species. Dotted lines denote gene flow by pollen and solid lines denote gene flow by seeds. Flags represent the SSR genotype of each individual: the first line shows the nuSSRs genotype and the second line shows the cpSSRs allele.

locus, while their spatial position can be known from the latitude and longitude data.

2.2 Indirect Estimations of Gene Flow

Empirical estimation of gene flow in natural populations may be obtained indirectly from the degree of population genetic structure. Indirect methods for estimating gene flow require allele frequencies from individuals sampled in different populations (Slatkin, 1985a) and are based on the analysis of allele distributions in populations to infer levels or patterns of gene flow among populations (Planter, 2007). Hence, using codominant genetic markers, i.e. those markers that allow the identification of heterozygous genotypes, it is possible to estimate the effective number of migrants from the differences in genotypic frequencies using population genetic models (Broquet and Petit, 2009). Microsatellites or simple sequence repeat markers (SSRs) are one of the most prominent codominant genetic markers. They are monotonous tandem repeats of short nucleotide motifs from 1 bp to 6 bp (Tautz and Renz, 1984). Microsatellites are considered by many authors as selectively neutral markers with a random or quasi-random distribution in the genome (Li *et al.*, 2002). In view of selectively neutral markers, when gene flow is disrupted, genetic drift is a more effective microevolutionary process than mutation for effecting genetic differentiation between populations (Slatkin and Barton, 1989).

Gene flow can be quantified by the parameter m, which describes the movement of each gamete or individual independently of population size (Slatkin and Barton, 1989). However, because gene flow requires both movement and reproduction, m is not simply a measure of the dispersal of individuals or gametes among local populations; conversely, m represents a complex interaction between dispersal patterns and mating system of the species (Templeton, 2006). Taking into account the Wright's islands model, which assumes that a large number of subpopulations of equal size constantly exchange individuals or gametes with other

subpopulations (Wright, 1931, 1951), m indicates the probability of each gamete being an immigrant (Slatkin and Barton, 1989).

The evolutionary consequences of gene flow can be quantified by the effective number of migrants, N_em (N_em = effective population size × migration rate) (Wright, 1931). The effective number of migrants determines if the action of genetic drift can alone produce substantial genetic differentiation between populations. Where N_em is > 1, the gene flow is sufficient to prevent substantial differentiation by genetic drift (Slatkin and Barton, 1989).

Slatkin and Barton (1989) proposed different methods to estimate the relative role of gene flow and genetic drift using the allele frequency distribution. These methods include those based on Wright's fixation statistic (F_{ST}) and another based on rare alleles (Slatkin and Barton, 1989). Given the importance of understanding population dynamics and evolution, these two simple methods for estimating N_em have made significant contributions in ecology and evolution, especially since the 1990s (Yamamichi and Innan, 2012).

2.2.1 Indirect estimation of gene flow from F_{ST} index

The fixation index F_{ST} (Wright, 1951) is one of the most commonly used population genetic structure measures in evolutionary genetics studies. This index calculates the relationship between gene flow and genetic drift and allows knowledge of how this relationship defines the population genetic structure (Templeton, 2006). Wright (1951) showed that, under the island model:

$$F_{ST} \approx \frac{1}{1+4N_em} \text{ if } m<<1 \qquad \text{(Eqn. 2.1)}$$

From this equation,

$$N_em = \frac{1}{4}\left(\frac{1}{F_{ST}} - 1\right) \qquad \text{(Eqn. 2.2)}$$

Hence, the first step for indirect estimation of gene flow is to calculate the F_{ST} index using, for example, analysis of molecular variance (AMOVA) (Excoffier *et al.*, 1992). AMOVA allows estimating F_{ST} together with its statistical significance by nonparametric permutation methods. A prerequisite for the estimation of N_em indirectly from F_{ST} is that this index must be statistically significant.

Indirect estimation of gene flow levels from population genetic structure metrics should be cautious, mainly because these methods use assumptions about populations that are unlikely to be true in nature. Whitlock and McCauley (1999) criticized the estimation of N_em indirectly from F_{ST}, although these authors considered that this indirect estimation may be useful under some assumptions, i.e. that the spatial scale is small, the migration rate is relatively high, and the sample size and the number of loci are high. They considered that, in this case, the estimation of gene flow levels from N_em is the real biological question.

2.2.2 Indirect estimation of gene flow using the rare alleles method

Slatkin (1985b) developed a method to estimate average N_em values from the spatial distribution of rare alleles. Barton and Slatkin (1986) developed an analytical theory to demonstrate that the frequency of alleles present in a single population $\left[\bar{p}_{(1)}\right]$ is analogous to F_{ST}, therefore this would be a measure of dispersion in the distribution of gene frequencies.

Most of the alleles belonging to a single population have been present long enough to achieve a quasi-equilibrium between migration and genetic drift (Barton and Slatkin, 1986). Using simulation studies under the island model, Slatkin (1985b) described the relationship $\left[\log_{10}\left(\bar{p}_1\right)\right]$ approximately linear with $\left[\log_{10}\left(Nm\right)\right]$, in this way:

$$\left[\log_{10}\left(\bar{p}_1\right)\right] \approx a\left[\log_{10}\left(Nm\right)\right] + b \qquad \text{(Eqn. 2.3)}$$

where a and b depend on the number of individuals sampled in each population.

Hence, the presence of rare alleles is a prerequisite, because the estimation of gene flow is based on the frequency of them in the analyzed populations.

2.2.3 Quantification of recent gene flow

The identification of putative immigrants or descendants of recent immigrants can be developed by the Bayesian inference method with

a priori information about the geographical location of the individuals to infer the population genetic structure. This analysis assumes that each individual shows high probability to be native from the geographic region where they were sampled, although it allows for individuals to be immigrants or descendants of recent immigrant individuals, albeit with low probability (Pritchard *et al.*, 2000). To implement this method the migration rate must be previously known, otherwise different values of migration rate must be tested to calculate the probability that an individual is an immigrant. It is suggested that values within the 0.001–0.1 range should be used (Pritchard *et al.*, 2000).

This Bayesian approach allows for the identification of migrant individuals from multilocus genotypes without population genetics models that include assumptions of genetic migration–drift equilibrium. This is because it provides information about recent migration, based on the instability observed in the multi-locus genotypes of migrants or recent descendants of migrants.

The identification of putative immigrants or descendants of recent immigrants by the Bayesian inference method includes a few assumptions, such as that population must be in Hardy–Weinberg equilibrium and about linkage disequilibrium between the considered loci. However, it is only informative about recent migration patterns in opposition to indirect estimates of historical gene flow, both approaches being complementary since they provide information about gene flow at different time scales (Wilson and Rannala, 2003).

2.3 Relative Roles of Pollen- versus Seed-Mediated Gene Flow

Gene flow among plant populations can occur in two different ways; hence a complete description of gene flow should include evaluation of the relative role of pollen and seeds as intermediaries of gene flow (Ennos, 1994). In this way, molecular markers with different inheritance modes are required for distinguishing the effects of seed and pollen movement and make inferences about pollen- and seed-mediated gene flow levels among populations (Oddou-Muratorio *et al.*, 2001). Microsatellites are the most suitable molecular markers for gene flow studies because they are detectable in all three plant genomes: nucleus, chloroplast, and mitochondria (Tautz and Renz, 1984), and also these genomes have different modes of inheritance in plants. In most angiosperms the nucleus is inherited in a biparental manner, dispersed by pollen and seeds, whereas the chloroplast and mitochondria are inherited in a uniparental way, dispersed only by seeds (Corriveau and Coleman, 1988). Because of its small size, conserved gene content and simple structure, the chloroplast genome has generally been the subject for population genetic studies rather than the mitochondrial genome (Raubeson and Jansen, 2005). In view of this, comparative analyses by means of nuclear and chloroplast microsatellite markers have been widely used for population genetic studies in plants because these markers provide complementary and often contrasting information on population genetic structure, genetic differentiation, and gene flow within and between populations (Pakkad *et al.*, 2008).

2.3.1 Estimation of relative roles of pollen- versus seed-mediated gene flow under island model

The theoretical framework for estimating the relative contribution of pollen and seeds to gene flow is based on the assumption that the equilibrium F_{ST} index values for the chloroplast DNA polymorphisms may be dramatically higher than the equilibrium F_{ST} index values for nuclear DNA polymorphisms, as long as the rate of pollen dispersal exceeds the rate of seed dispersal and the inheritance of the chloroplast DNA is strictly maternal (McCauley, 1995). Ennos (1994) demonstrated that the ratio of gene flow by pollen to gene flow by seeds is a simple function of F_{ST} index values for nuclear and chloroplast markers. In angiosperms, biparental and maternal inheritance can result in contrasting F_{ST} index values when pollen movement occurs among populations (McCauley, 1995). Therefore, it is expected that genetic differentiation will vary between uniparental and biparental inheritance markers for the same set of populations (Ennos, 1994).

Populations that show fixed haplotypes in the cytoplasmic genome do not necessarily show low diversity in nuclear markers (Tollefsrud *et al.*, 2009). In migration–genetic drift equilibrium the magnitude of the fixation index for the maternal inherited genome (F_{STm}) is expected to exceed the magnitude of the fixation index for the biparental inherited genome (F_{STb}) because maternally inherited markers only disperse through seeds (Ennos, 1994). Where the levels of seed-mediated gene flow between populations is lower than the pollen-mediated gene flow, the relative rate of pollen-mediated and seed-mediated gene flow (r) from the estimated levels of population genetic structure from biparental inheritance markers (nuSSRs) and from maternal inheritance markers (cpSSRs) can be estimated using the following equation:

$$r = \frac{\left[\left(\frac{1}{F_{STb}-1}\right)-2\times\left(\frac{1}{F_{STm}-1}\right)\right]}{\left(\frac{1}{F_{STm}-1}\right)} \qquad \text{(Eqn. 2.4)}$$

where F_{STb} and F_{STm} indicate the estimated fixation index from nuclear and chloroplast markers, respectively (Ennos, 1994).

When using that equation, the same populations must be genotyped with both cpSSRs and nuSSRs markers, although the genotyping of the same individuals with both markers is not an imperative requisite.

2.3.2 Estimation of relative roles of pollen- versus seed-mediated gene flow under isolation by distance model

Although the Ennos estimator is one of the most popular indices for assessing the contribution of pollen versus seed gene flow, an alternative estimator has been suggested for hermaphroditic plant species under isolation by the distance model. The alternative r_{IBD} index arises from pairwise F_{ST} estimates between populations rather than a single F_{ST} statistic for a complete population (Barrandeguy and García, 2018).

The r_{IBD} index is estimated based on the slope values from the regression between $F_{ST}/(1\text{-}F_{ST})$ and the logarithm of the distance between pairs of populations for nuclear and chloroplast markers using the following equation:

$$r_{IBD} = \frac{\left(\frac{1}{b_b}\right)-2\times\left(\frac{1}{b_m}\right)}{\left(\frac{1}{b_m}\right)} \qquad \text{(Eqn. 2.5)}$$

where b_b and b_m are the slope values from regressions for nuclear and chloroplast markers, respectively. A prerequisite for using the proposed r_{IBD} index is a significant pattern of isolation by distance for both genomes (Barrandeguy and García, 2018).

2.4 Determination of the Most Probable Model of Genetic Admixture Among Populations by Approximate Bayesian Computation Methods

Approaches for making population genetic inferences have changed during the past decades as a consequence of increases in computer processing speed and available genetic data. The widespread availability of different molecular markers and increased computer power that allow the use of simulation methods has promoted the development of sophisticated statistical methods to study complex population history in order to estimate phylogeographic parameters such as historical population sizes, divergence times, and migration rates (r_a) given the stochastic timing of coalescent events (Drummond *et al.*, 2005; Hickerson *et al.*, 2010). The coalescence theory formalized a powerful and elegant method for modeling gene genealogies backwards in time from a sample of alleles under virtually any complex demographic history. In this way, coalescence theory describes the role of population genetic processes for determining the shape of the genealogies (Drummond *et al.*, 2005) and allows description of the relationship between the genealogies and the demographic history of the studied population (Kingman, 1982).

Coalescent simulations are useful for testing if simulated data tend to fit the observed data under a specific model. The most confident model is that which shows fewer differences between observed (real) data and simulated data. This basic concept has made coalescent simulations one of the most important computational tools in modern population genetics (Nielsen and Slatkin, 2013).

Approximate Bayesian computation (ABC) methodology allows for a statistical approach for selecting the optimal model. This method also allows the inference of parameters under complicated scenarios (Beaumont, 2010). The ABC method uses summary statistics, i.e. parameters calculated from the data to represent the maximum amount of information in the simplest possible form, and simulations (Csilléry *et al.*, 2010).

The ABC method is buoyed by the idea that models that show high posterior probability will produce summary statistics close to those estimated from observed data (Barrandeguy *et al.*, 2017). The rejection algorithm is generated from simulations of large datasets under a hypothesized evolutionary scenario. Then this new dataset generated by simulations is reduced to summary statistics while the distance between the simulated and the observed summary statistics allows acceptance or rejection of the sampled parameters. The generated sub-sample of accepted values contains the fitted parameter values, and allows evaluation of uncertainty on parameters given the observed statistics. The parameters of the evolutionary scenario are sampled from a probability distribution (Csilléry *et al.*, 2010).

The three main steps of the ABC method can be summarized as: (i) formulate the model; (ii) apply the model to data and improve the model by checking its fit; and (iii) compare it with other models. These steps are strongly inter-dependent and should be considered as a unified approach, with a possibility of cycling through the three stages (Cornuet *et al.*, 2010; Csilléry *et al.*, 2010).

However, the ABC method has at least two potential limitations: (i) its inference is limited to a finite set of phylogeographic scenarios; and (ii) it establishes an arbitrary choice of summary statistics. In consequence, conclusions from ABC could be influenced by subjective inclusion of evolutionary scenarios and implicit model assumptions unexpected by the modeler (Csilléry *et al.*, 2010). The universe of potential historic scenarios is huge but only a limited subset of these putative scenarios is considered. Finally, the creativity of the individual researcher, not simply their computational prowess, determines the insights gained from a phylogeographic analysis (Knowles, 2009).

A comparison of all methods previously discussed is shown in the Table 2.1. In addition, this table shows the computational free access software that can be used to estimate the parameters.

2.5 Case Studies

In nature, populations are typically structured into subpopulations; gene flow is responsible for maintenance of genetic variation among and within subpopulations (Yamamichi and Innan, 2012). Also, gene flow enhances adaptation to local environments, because some alleles could be adaptive in some environments and not adaptive in others. Thus, to understand the evolutionary dynamics of a population is critical to quantifying the level of gene flow (Yamamichi and Innan, 2012). We review selected case studies in which gene flow estimators were applied to answer several biological questions.

The dragonsblood tree, *Pterocarpus officinalis*, is distributed in the Caribbean, South America, and Central America. To help better define a conservation strategy for this species, Muller *et al.* (2009) analyzed 202 individuals from nine subpopulations by means of three cpSSRs loci and six nuSSRs loci to quantify the genetic variation within Caribbean islands and to assess the pattern of differentiation and infer levels of gene flow. As established by the authors, 'the rate of gene flow was assessed according to two complementary methods; the first is indirect and based on the hypothesis of (mid to long term) equilibrium between drift and migration, whereas the other is direct and estimates "instantaneous" migration', i.e. they implemented indirect estimation of gene flow from the F_{ST} index and quantified recent gene flow, respectively. Both implemented methods produced robust results that confirmed that there was low level of gene flow per generation among the analyzed populations.

However, the authors indicated that migration can occur over very long distances and these results could be expected in the context of populations separated by sea-water.

The epiphytic orchid *Epidendrum firmum* occurs in humid premontane and lower montane wet forests where neotropical mountain

Table 2.1. Comparison among gene flow parameters estimations.

Estimated parameter	Description of the parameter	Method	Reference	Software	Reference of the software	Required input data	GPS data (Yes/No)	Expected result values	Theoretical interpretation of results	Drawbacks
N_em	Effective number of migrants	Indirect estimation from F_{ST}	Wright (1951)	Manual calculation using the F_{ST} estimated from an AMOVA developed with **Arlequín** or **Genalex**	Excoffier and Lischer (2010) Peakall and Smouse (2012)	Multilocus genotypes from different populations	No	From 0 to > 1	N_em >1 the gene flow is sufficient to prevent substantial differentiation by genetic drift	Uses assumptions about populations that are unlikely in nature
		Indirect estimation from rare alleles	Slatkin (1985b)	**GenePop GenePon on the web**	Rousset (2008)	Multilocus genotypes from different populations	No	From 0 to > 1	N_em >1 the gene flow is sufficient to prevent substantial differentiation by genetic drift	Uses assumptions about populations that are unlikely in nature. Requires the presence of rare alleles
Recent gene flow	Identification of possible immigrants or descendants of recent immigrants	Probability of one individual to be a migrant or descendant of recent migrants	Pritchard *et al*. (2000)	**Structure**	Pritchard *et al*. (2000)	Multilocus genotypes with location of the individuals	No	Bayesian analysis barplot of genetic structure highlighted with recent migrant individuals	Provides information about recent migration based on the instability observed in the multi-locus genotypes of migrants or recent descendants of migrants	Assumes Hardy–Weinberg equilibrium within clusters and linkage disequilibrium between the considered loci. Immigration rate *v* must be previously known, otherwise different values of *v* must be tested

Continued

Table 2.1. Continued

Estimated parameter	Description of the parameter	Method	Reference	Software	Reference of the software	Required input data	GPS data (Yes/No)	Expected result values	Theoretical interpretation of results	Drawbacks
r	Relative roles of pollen- versus seed-mediated gene flow	Indirect estimation from F_{ST} estimated from nuclear and chloroplast markers	Ennos (1994)	Manual calculation from F_{ST} for nuclear and chloroplast genomes estimated using **Arlequín** or **Genalex**	Excoffier and Lischer (2010) Peakall and Smouse (2012)	Multilocus nuclear and chloroplast genotypes with location of the individuals	No	The magnitude that pollen gene flow levels exceed the seed gene flow levels under the island model	Relative importance of pollen and seeds as intermediaries of gene flow	Uses assumptions about populations that are unlikely in nature
r_{IBD}	Relative roles of pollen- versus seed-mediated gene flow	Indirect estimation from regression between $F_{ST}/(1\text{-}F_{ST})$ and the logarithm of distance between pairs of population	Barrandeguy and García (2018)	Manual calculation from the slopes of regression between $F_{ST}/(1\text{-}F_{ST})$ and the logarithm of distance between pairs of populations using nuclear and chloroplast markers developed with **Genepop**	Rousset (2008)	Multilocus nuclear and chloroplast genotypes with location of the individuals	Yes	The magnitude that the pollen gene flow levels exceed the seed gene flow levels under isolation by distance model	Relative importance of pollen and seeds as intermediaries of gene flow	Developed for hermaphrodite plant species

Continued

Table 2.1. Continued

Estimated parameter	Description of the parameter	Method	Reference	Software	Reference of the software	Required input data	GPS data (Yes/No)	Expected result values	Theoretical interpretation of results	Drawbacks
r_a	Magnitude of gene flow during admixture	Estimation of the posterior value of the parameter implementing Approximate Bayesian Computation method (ABC) considering scenarios with admixture between populations	Csilléry *et al*. (2010)	**DIYABC**	Cornuet *et al*. (2014)	Multilocus SSRs genotypes, SNPs or sequence data	No	From 0 to 0.99	Posterior distribution of parameter values by simulating data with parameters drawn from specified prior distributions and retaining values that produce datasets similar to the observed data	Inference limited to a finite set of phylogeographic scenarios and arbitrary choice of summary statistics

habitat has narrow elevation bands that could constitute barriers that maintain isolated subpopulations of many species. Hence, genetic connectivity among subpopulations depends on gene flow. Kartzinel *et al*. (2013) studied populations of *E. firmum* in order to know if significant genetic structure reflects the isolation of subpopulations in separate mountain ranges and how this influences pollen and seed flow among subpopulations within a narrow band of mountain habitat. These authors analyzed 225 individuals from 12 subpopulations across Costa Rica by means of five nuSSRs loci and two noncoding chloroplast loci. As established by the authors: 'We estimated the relative contribution of pollen and seeds to gene flow at multiple spatial scales by calculating pollen to seed migration ratios.' Results of the estimation of relative roles of pollen- versus seed-mediated gene flow under the island model showed that pollen contributed substantially more to gene flow among populations than seed (mp/ms = 46). This result was consistent with the contrasting patterns of genetic structure that could be interpreted as the occurrence of significant cpDNA barriers within and among mountain ranges, but not significant nuDNA barriers after accounting for geographic distance. Pollinator-mediated extensive gene flow erodes nuDNA colonization footprints, while seed flow was comparatively limited, possibly because of the direction of prevailing winds across linearly distributed populations. The authors concluded that dispersal traits alone may not accurately inform predictions about gene flow or genetic structure, supporting the need for research into the potentially crucial role of pollinators and landscape context in gene flow among isolated populations.

Scots pine, *Pinus sylvestris*, is a widely distributed conifer of the boreal regions. It is hypothesized that climate change shaped the modern-day genetic structure and phylogeographic pattern of the species. During the Quaternary, climate fluctuations caused radical changes in distribution of tree species and resulted in large-scale range shifts, population contractions and expansions. Tóth *et al*. (2019) analyzed 331 individuals from 16 subpopulations located in the Pannonian Basin and Western Carpathians. The study was performed by means of eight nuSSRs loci. Approximate Bayesian computation methodology and maximum entropy distribution modeling with temperature- and precipitation-related bioclimatic data were implemented in order to describe demography and past distribution patterns of Scots pine populations from the highly fragmented landscape. The ABC analysis showed that two genetic lineages have diverged from an ancestral Scots pine. The divergence dating was in the Mid-Pleistocene; due to the favorable climatic conditions the species underwent population expansion leading to an admixture event.

Cucurbita pepo is a cultivated crop with economic value. The crop's progenitors are *C. pepo* ssp. *pepo*, and two wild taxa (*C. pepo* ssp. *fraterna* and *C. pepo* ssp. *ovifera*). The origin of pumpkins (*Cucurbita pepo* ssp. *pepo*) in Mexico can be analyzed by means of gene flow estimation studies. Castellanos-Morales *et al*. (2019) analyzed 374 individuals from 13 locations located in Mexico by means of two chloroplast regions and nine nuSSRs loci to assess the levels of genetic variation and genetic structure. Four hypotheses regarding the origin of *C. pepo* ssp. *pepo*'s ancestor were tested using ABC. The study concluded that the three subspecies are clearly differentiated. *Cucurbita pepo* ssp. *fraterna* was probably *C. pepo* ssp. *pepo*'s wild ancestor, but subsequent hybridization between taxa make it hard to define *C. pepo* ssp. *pepo*'s ancestor because the estimated migration rate (r_a = 0.619) suggests that *C. pepo* ssp. *pepo* has similar ancestry to that of *C. pepo* ssp. *ovifera* and *C. pepo* ssp. *fraterna*.

2.6 Conclusions

Gene flow is a key population genetics parameter. It may be estimated and viewed in different ways, such as the effective number of migrants, recent colonization, ratio of pollen- versus seed-mediated gene flow, or as a migration rate using coalescent methods. Gene flow estimates using molecular methods are most often inferred using neutral genetic markers. Several variables are important, including the magnitude of genetic drift on allele frequencies in a population and the magnitude of genetic exchange among populations. Genetic

drift could be understood as a pressure, since it causes the loss of the genetic diversity in natural populations. Genetic diversity encompasses all genetic variability within a population and is the basis for local adaptation. Current climate change and fragmentation of natural populations as a consequence of anthropic impacts call for urgent gene flow research. Especially important is documenting gene flow levels among natural populations of endangered species and species of economic importance; they are reservoirs of biodiversity and they will ensure the availability of natural genetic resources. Despite the methodological problems with the estimators described in this chapter, these estimators are useful and irreplaceable for evaluating the gene flow role by means of molecular genetic data in natural populations. Also, these estimators can be used in cultivated species to compare agricultural populations with natural populations. These comparisons enable documentation of gene pool mixtures as well as the effects of genetic engineering, which are helpful to understand *in situ* or *ex situ* germplasm collections and diversity.

References

Barrandeguy, M.E. and García, M.V. (2018) A regression approach to estimate the relative roles of pollen- versus seed-mediated gene flow under an isolation by distance model. *Silvae Genetica* 67(1), 41–50. DOI: 10.2478/sg-2018-0006.

Barrandeguy, M.E., Sanabria, D.J. and García, M.V. (2017) The revitalization of population genetics: from coalescence theory to phylogeography. In: *Population Genetics*. SM Group Open Access eBooks, Dover, Delaware.

Barton, N.H. and Slatkin, M. (1986) A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity* 56, 409–415. DOI: 10.1038/hdy.1986.63.

Beaumont, M.A. (2010) Approximate Bayesian Computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics* 41(1), 379–406. DOI: 10.1146/annurev-ecolsys-102209-144621.

Broquet, T. and Petit, E.J. (2009) Molecular estimation of dispersal for ecology and population genetics. *Annual Review of Ecology, Evolution, and Systematics* 40(1), 193–216. DOI: 10.1146/annurev.ecolsys.110308.120324.

Castellanos-Morales, G., Ruiz-Mondragón, K.Y., Hernández-Rosales, H.S., Sánchez-de la Vega, G., Gámez, N. *et al*. (2019) Tracing back the origin of pumpkins (*Cucurbita pepo* ssp. *pepo* L.) in Mexico. *Proceedings of the Royal Society B: Biological Sciences* 286(1908), 1440. DOI: 10.1098/rspb.2019.1440.

Cornuet, J.-M., Ravigné, V. and Estoup, A. (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics* 11, 401. DOI: 10.1186/1471-2105-11-401.

Cornuet, J.-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M. *et al*. (2014) DIYABC v2.0: A software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics Applications Note* 30(8), 1187–1189. DOI: 10.1093/bioinformatics/btt763.

Corriveau, J.L. and Coleman, A.W. (1988) Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperm species. *American Journal of Botany* 75(75), 1443–1458. Available at: https://www.jstor.org/stable/2444695 (accessed 5 May 2021). DOI: 10.1002/j.1537-2197.1988.tb11219.x.

Csilléry, K., Blum, M.G.B., Gaggiotti, O.E. and Francois, O. (2010) Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology and Evolution* 25(7), 410–418. DOI: 10.1016/j.tree.2010.04.001.

Drummond, A.J., Rambaut, A., Shapiro, B. and Pybus, O.G. (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* 22(5), 1185–1192. DOI: 10.1093/molbev/msi103.

Ennos, R.A. (1994) Estimating the relative rates of pollen and seed migration among plant populations. *Heredity* 72(3), 250–259. DOI: 10.1038/hdy.1994.35.

Excoffier, L. and Lischer, H. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology* 10(3), 564–567. DOI: 10.1111/j.1755-0998.2010.02847.x.

Excoffier, L., Smouse, P. and Quattro, J. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131(2), 479–491. DOI: 10.1093/genetics/131.2.479.

Hickerson, M.J., Carstens, B.C., Cavender-Bares, J., Crandall, K.A., Graham, C.H. *et al*. (2010) Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution* 54(1), 291–301. DOI: 10.1016/j.ympev.2009.09.016.

Kartzinel, T.R., Shefferson, R.P. and Trapnell, D.W. (2013) Relative importance of pollen and seed dispersal across a Neotropical mountain landscape for an epiphytic orchid. *Molecular Ecology* 22(24), 6048–6059. DOI: 10.1111/mec.12551.

Kingman, J.F.C. (1982) The coalescent. *Stochastic Processes and their Applications* 13(3), 235–248. DOI: 10.1016/0304-4149(82)90011-4.

Knowles, L.L. (2009) Statistical phylogeography. *Annual Review of Ecology, Evolution, and Systematics* 40(1), 593–612. DOI: 10.1146/annurev.ecolsys.38.091206.095702.

Li, Y., Korol, A.B., Fahima, T., Beiles, A. and Nevo, E. (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology* 11(12), 2453–2465. DOI: 10.1046/j.1365-294X.2002.01643.x.

McCauley, D.E. (1995) The use of chloroplast DNA polymorphism in studies of gene flow in plants. *Trends in Ecology & Evolution* 10(5), 198–202. Available at: https://www.jstor.org/stable/2445849 (accessed accessed 5 May 2021). DOI: 10.1016/S0169-5347(00)89052-7.

Muller, F., Voccia, M., Ba, A. and Bouvet, J.M. (2009) Genetic diversity and gene flow in a Caribbean tree *Pterocarpus officinalis* Jacq.: a study based on chloroplast and nuclear microsatellites. *Genetica* 135(2), 185–198. DOI: 10.1007/s10709-008-9268-4.

Nielsen, R. and Slatkin, M. (2013) *An Introduction to Population Genetics. Theory and Applications*. Sinauer Associates, Sunderland, Massachusetts.

Oddou-Muratorio, S., Petit, R.J., Le Guerroue, B., Guesnet, D. and Demesure, B. (2001) Pollen versus seed mediated gene flow in a scattered forest tree species. *Evolution* 55(6), 1123–1135. DOI: 10.1111/j.0014-3820.2001.tb00632.x.

Pakkad, G., Ueno, S. and Yoshimaru, H. (2008) Genetic differentiation of *Quercus semiserrata* Roxb. in northern Thailand revealed by nuclear and chloroplast microsatellite markers. *Forest Ecology and Management* 255(3-4), 1067–1077. DOI: 10.1016/j.foreco.2007.10.021.

Peakall, R. and Smouse, P.E. (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28(19), 2537–2539. DOI: 10.1093/bioinformatics/bts460.

Petit, R.J. and Vendramin, G.G. (2007) Plant phylogeography based on organelle genes: an introduction. In: Weiss, S. and Ferrand, N. (eds) *Phylogeography of Southern European Refugia*. Springer, Dordrecht, Netherlands, pp. 23–97.

Planter, E.A. (2007) Flujo génico: métodos para estimarlo y marcadores moleculares. In: Eguiarte, L., Souza, V. and Aguirre, X. (eds) *Ecología molecular*. Instituto Nacional de Ecologia, Mexico, pp. 49–62.

Pritchard, J.K., Stephens, M. and Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics* 155(2), 945–959. DOI: 10.1093/genetics/155.2.945.

Raubeson, L.A. and Jansen, R.K. (2005) Chloroplast genomes of plants. In: Henry, R. (ed.) *Diversity and Evolution of Plants – Genotypic and Phenotypic Variation in Higher Plants*. CABI Publishing, Wallingford, UK, pp. 45–68.

Rousset, F. (2008) GENEPOP '007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8(1), 103–106. DOI: 10.1111/j.1471-8286.2007.01931.x.

Slatkin, M. (1985a) Gene flow in natural populations. *Annual Review of Ecology and Systematics* 16(1), 393–430. DOI: 10.1146/annurev.es.16.110185.002141.

Slatkin, M. (1985b) Rare alleles as indicators of gene flow. *Evolution* 39(1), 53–65. DOI: 10.1111/j.1558-5646.1985.tb04079.x.

Slatkin, M. and Barton, N.H. (1989) A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* 43(7), 1349–1368. DOI: 10.1111/j.1558-5646.1989.tb02587.x.

Tautz, D. and Renz, M. (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research* 12(10), 4127–4138. DOI: 10.1093/nar/12.10.4127.

Templeton, A.R. (2006) *Population Genetics and Microevolutionary Theory*. Wiley-Liss, Hoboken, New Jersey.

Tollefsrud, M.M., Sønstebø, J.H., Brochmann, C., Johnsen, Ø., Skrøppa, T. *et al*. (2009) Combined analysis of nuclear and mitochondrial markers provide new insight into the genetic structure of North European *Picea abies*. *Heredity* 102(6), 549–562. DOI: 10.1038/hdy.2009.16.

Tóth, E.G., Bede-Fazekas, A., Vendramin, G.G., Bagnoli, F. and Höhn, M. (2019) Mid-Pleistocene and Holocene demographic fluctuation of Scots pine (*Pinus sylvestris* L.) in the Carpathian Mountains and the Pannonian Basin: signs of historical expansions and contractions. *Quaternary International* 504(3), 202–213. DOI: 10.1016/j.quaint.2017.11.024.

Whitlock, M.C. and McCauley, D.E. (1999) Indirect measures of gene flow and migration: F_{ST} 1/(4*Nm* + 1). *Heredity* 82, 117–125. DOI: 10.1038/sj.hdy.6884960.

Wilson, G.A. and Rannala, B. (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163(3), 1177–1191. DOI: 10.1093/genetics/163.3.1177.

Wright, S. (1931) Evolution in Mendelian populations. *Genetics* 16(2), 97–159. DOI: 10.1093/genetics/16.2.97.

Wright, S. (1951) The genetical structure of populations. *Annals of Eugenetics* 15(4), 323–354. DOI: 10.1111/j.1469-1809.1949.tb02451.x.

Yamamichi, M. and Innan, H. (2012) Estimating the migration rate from genetic variation data. *Heredity* 108(4), 362–363. DOI: 10.1038/hdy.2011.83.

3 Challenges for Monitoring (Trans)Gene-flow in the Environment[1]

Sarah Zanon Agapito-Tenfen[1]* and Fern Wickson[1,2]
[1]GenØk Centre for Biosafety, Tromsø, Norway; [2]North Atlantic Marine Mammal Commission, Tromsø, Norway

Abstract

Monitoring the presence of transgenes in the environment depends on analytical detection methods and their measurement uncertainties. In this chapter, we aim to identify key methodological aspects and pinpoint the research bottlenecks and needs for building the capacity to effectively monitor transgene escape from genetically modified (GM) crops to wild relatives or landraces. We reviewed the iconic debate concerning the presence of transgenes in landraces of Mexican maize (with both positive and negative results for GM contamination shown over the years) to examine the impact of using different approaches to monitoring and detecting transgenes in landraces. Despite the lack of clear international guidelines that are specific for sampling and testing heterogeneous and dispersed landraces and wild relatives, transgene detection methods have developed significantly over the past decades. There is now an immensely valuable set of tools, approaches, and harmonized protocols that continue to develop and provide a way to evaluate transgene presence. To support this ongoing development and to steer it in directions that are particularly useful for addressing the challenges associated with detection in landraces and wild relatives, we offer lessons from our review for future work in this area.

Keywords: agrobiodiversity; biodiversity; biotechnology; environmental risk; gene flow; genetically modified organism (GMO); genetic testing; maize

3.1 Introduction

Wild relatives are the ancestors of domesticated crop plants that continue to persist in the wild, while landrace varieties are those domesticated crops that have been traditionally bred and adapted over time to meet the needs of certain natural and cultural environments. Genetically modified (GM) crops can be spread into locations occupied by landraces and wild relatives via seed flow through formal and informal seed systems. Once GM seeds are present in an area, transgenes can then be introduced into the genomes of landraces or wild relatives via pollen flow, since pollen from GM crops has the potential to fertilize and further hybridize with female

*Corresponding author: sarah.agapito@uit.no

 Gene Flow: Monitoring, Modeling and Mitigation (eds Wei Wei and C. Neal Stewart, Jr)
DOI: 10.1079/9781789247480.0003

gametes in any sexually compatible species. Transgene introgression takes place if repeated backcrosses occur and the transgene becomes stabilized in the new host genome (Stewart *et al.*, 2003).

The potential for transgene flow into landraces and wild relatives is an important biosafety issue widely recognized around the world and therefore an important component of the regulatory risk assessments performed on GM crops prior to their approval for cultivation. The issue of transgene flow into landraces of traditional crop varieties has particular significance for countries that are centers of origin and diversity, such as Mexico is for maize. Transgene flow really hit the headlines as an important biosafety issue in scientific, public, and policy circles when it was first reported as detected in Mexican maize in late 2001 (Quist and Chapela, 2001).

In 2001, when transgene flow was first reported in Mexican maize, the import of GM maize for food and feed was permitted in Mexico and the vast majority of these imports came from the USA, where GM maize was deregulated and unlabeled. The cultivation of GM maize was not permitted in Mexico at this time. The Mexican government had implemented a moratorium on cultivation in 1998 due to concerns over potential impacts on traditional maize biodiversity. While the government granted approval for field trials of GM maize in 2009, in 2013 this became subject to a class-action lawsuit from a coalition of activist groups claiming that this threatened traditional biodiversity, with the judge in the case ruling that all planting be suspended until a final decision was made (Vargas-Parada, 2014). Indeed, regulatory approvals for GM maize cultivation in Mexico are still (at the time of writing) subject to a high-level legal dispute that continues to block their commercial use (Garcia, 2017; Garcia Ruiz *et al.*, 2018).

The case of transgene flow into landraces of maize in Mexico has now also become an iconic debate. This is not only because maize is an important commodity crop and one of the five staple cereals that we currently depend on for global food security. It was also the first crop to be investigated for the presence of transgenes in its center of origin and diversity. The reported spread of transgenes into Mexican maize then developed into a highly technical debate extending beyond the scientific community to have significant relevance for and engagement from a range of other actors, including policy makers, social scientists, civil society organizations, indigenous communities, and farmer collectives. Indeed, a significant and active civil society movement against GM maize and in defense of native maize biodiversity has developed in Mexico after the initial reports of transgene detection (Baker, 2013; Carro-Ripalda *et al.*, 2015) and, as stated above, court cases continue over the question of approvals for cultivation. This widespread interest in the issue is particularly due to the socio-economic and cultural importance of maize in Mexico (Carro-Ripalda *et al.*, 2015) and the socio-political nature of the controversy surrounding GM organisms (GMOs) (Sarewitz, 2004; McAfee, 2008; Mercer and Wainwright, 2008; Delborne, 2008). Translating this interest in the issue to actually controlling transgene flow in Mexico is made particularly challenging by its geographic proximity to the USA and the high level of import that has been encouraged through the North Atlantic Free Trade Agreement (NAFTA) (Bello, 2009). Furthermore, prioritizing this issue for action is constrained by the availability of the necessary economic resources and governmental will, and the issue's embeddedness in the broader context of threats facing landrace maize biodiversity, including climate change, low profitability, declining subsidies, and emigration away from rural communities (Wainwright and Mercer, 2011; Mercer *et al.*, 2012; Rivera López *et al.*, 2020).

The first study to report the presence of transgenes in maize landraces in Mexico (Quist and Chapela, 2001) was performed in a remote area of Oaxaca. It detected transgenic DNA sequences in landrace varieties using the methods of polymerase chain reaction (PCR) and nucleotide sequencing. These findings were published in the prestigious journal *Nature* but became hotly contested upon their release, with several critiques challenging the accuracy and legitimacy of the results subsequently being published (Christou, 2002; Kaplinsky *et al.*, 2002; Metz and Fütterer, 2002). These critiques pointed to a potential misinterpretation of the specific methodology used to claim that the transgene sequences were permanently

incorporated (or 'introgressed') into the maize landrace genome. These critiques were followed by a response from Quist and Chapela (2002), in which they provided additional data, using a different method, to support their original findings. This high-level exchange sparked widespread scientific debate about the robustness of both the research methods and the results, ultimately leading to an unusual response from *Nature* in which they published an editorial note (appended to Quist and Chapela, 2002) claiming that there was insufficient evidence to justify the original publication, without going so far as to retract the paper.

After the initial study on transgene flow into landraces of maize not only became the subject of significant scientific debate but also demonstrated its broader socio-political significance, others attempted to follow up and provide further evidence to address the contested question of whether transgene flow had taken place into landrace maize at its center of origin (Ortiz-García *et al.*, 2005a; Serratos-Hernández *et al.*, 2007; Dyer *et al.*, 2009; Piñeyro-Nelson *et al.*, 2009a). However, each follow-up study that was performed to examine this issue used a different method and came to varying conclusions. The studies were also often subject to published critiques in which the methods used and interpretations applied were challenged (e.g. Cleveland *et al.*, 2005; Schoel and Fagan, 2009). There were additional studies confirming the presence of transgene flow conducted by both governmental authorities and non-governmental organizations that were never peer-reviewed and published, meaning that the details of their methods remained obscure (Mercer and Wainwright, 2008). The debate over transgene flow into landraces of maize in Mexico has therefore not only generated a broad socio-political debate on issues such as the industrialization of agriculture, neoliberal trade policies, relations with the USA, and threats to indigenous culture and livelihoods (Carro-Ripalda and Astier, 2014); it has also been fraught with scientific and technical disputes over what is the best or most appropriate and reliable method to use when seeking to monitor transgenes in landraces and wild relatives and how to interpret the variance that can be seen in the results.

Transgene detection methodologies and techniques for the purpose of monitoring have certainly evolved since the controversial study in 2001. Although there is still no scientific agreement on the extent of transgene flow into landraces of Mexican maize (Gilbert, 2013), it is now widely assumed that it is likely to be present, but just at low levels that make detection particularly challenging. There is currently no internationally agreed and standardized approach to testing that is specific for the unique challenges associated with landraces and wild relatives and the low-level presence that is likely.

In a previous work (Agapito-Tenfen and Wickson, 2018), we examined the methodological debate that has taken place over the past 15 years to extract lessons from the Mexican maize case for future work on transgene monitoring in landraces and wild relatives. The review was motivated by the need to find an appropriate method for performing our own original empirical research on transgene presence in Mexican maize and finding it difficult to navigate the varying results and debates in the literature. In this chapter, we discuss the major methodological issues and constraints that have emerged in the research over the years and indicate how they still impede robust and reliable transgene detection and monitoring in landraces and wild relatives. Highlighting what has been learned on this topic over the past 15 years of scientific work and methodological development, we then conclude by proposing what still needs to be done if we are to create robust and reliable test methods and monitoring for transgene flow into landraces and wild relatives.

3.2 Key Methodological Challenges

As indicated in our previous review (Agapito-Tenfen and Wickson, 2018) and by others commenting on the case of transgene detection in landrace maize in Mexico (Dalton, 2009; Quist and Catacora-Vargas, 2011; Gilbert, 2013), different sampling and analytical methods have been used by research groups and commercial laboratories and have produced varying results and interpretations. The different results

indicating both transgene absence and presence have significant implications for the political, legal and regulatory arenas (Christophe Bonneuil *et al.*, 2014), as well as for the indigenous farming communities growing native maize (and from which the samples were typically collected). Although the debate is amplified by the social and political importance of the findings, there are legitimate differences in scientific method that allow the lack of consensus to persist. These different approaches in scientific method arguably remain debated because of the unique challenges facing transgene monitoring in landraces and wild relatives and the uncertainty that remains around how to handle detection of low-level presence.

Successful and sound detection of transgenes in landraces or wild relatives faces a number of unique methodological challenges that complicate the task of monitoring, particularly for the type of low-level presence that is likely to occur in the field in places like Mexico where commercial cultivation is not approved. For example, transgenic DNA sequences or transgenic proteins must be intact and/or expressing to be detectable by current methods. Furthermore, methodological controls are typically based on endogenous genes or proteins and, since these might differ in landraces and wild relatives, detection is dependent on developing relevant and adapted controls. However, all available detection methods also have intrinsic problems that become particularly relevant for working with landraces or wild relatives, such as quantitative limits for detecting small amounts of a substance. Only comprehensive and transparent detection and monitoring methodologies can provide an estimation of the quantitative uncertainties involved (Box 3.1). In what follows below, we summarize what we see as the three key challenge areas: (i) environmental sampling; (ii) DNA isolation in heterogeneous samples; and (iii) the potential for false positives and negatives in PCR.

3.2.1 Environmental sampling

Sampling strategies can vary depending on the biological question and hypothesis under investigation. Sampling to detect and monitor transgene flow into landraces and wild relatives has two major challenges. First, sampling the environment is often limited by available time and funds to perform the analysis and so difficult choices have to be made about the scope of sampling and how many samples to test. The main difference between environmental sampling and sampling food or a shipment of seeds or grain is that the boundary of the material to be tested is somewhat more clearly defined and limited in the latter cases. This helps in the estimation of presence percentage and transgene frequency. However, when it comes to landraces and wild relatives in the field, there is rarely an established boundary. Knowing the

Box 3.1. Measurement of quantitative uncertainty in transgene detection methods

Every measurement is subject to some degree of uncertainty. Therefore, a measurement result is only complete within scientific studies if a statement of its uncertainty accompanies it. Measurement uncertainties can come from different sources, such as the measuring instrument, the item being measured, the environment, the operator, etc. These uncertainties can be estimated using statistical analysis of a set of measurements or using other kinds of information about the measurement process (Bell, 2001).

The most applied approach to calculate measurement uncertainty (MU) in the field of GMO detection is a top-down approach, in which data from collaborative trials (including all the factors influencing the MU during the analytical procedure) are used as a source for the estimation of MU. This approach is described in detail in the guidance document on measurement uncertainty for GMO testing laboratories produced by the European Network of GMO Laboratories (ENGL) working group on measurement uncertainty (Žel *et al*., 2012).

When the outcome of the measurement of uncertainty is not adequate (the performance characteristics do not satisfy all criteria set prior to the measurement), the method might be considered to be unreliable for its purpose.

Fig. 3.1. Maize cultivation fields showing the different cultivation stages and plot sizes in the Oaxaca region in Mexico. (Photograph by Sarah Z. Agapito-Tenfen, July 2015.)

cultivation area (for example, through empirical or geostatistical methods) can help in calculating the detection power of different sampling approaches. However, although a farmer's plot or a determined region/area may be considered a limit to determine a cultivation area of interest, it is important to remember that these almost never represent an actual barrier to transgene flow (Serratos-Hernández *et al.*, 2004). For instance, in the Oaxaca region in Mexico the average producer is farming on plots of less than 2.5 ha (Fig. 3.1). It has been suggested that any sampling effort in landraces or wild relatives is likely to underestimate the presence and/or frequency of transgenes in the sampled population (Quist, 2007) and not detecting a transgene in a sample population therefore offers no guarantee that the population is in fact transgene free. The intrinsic uncertainty this creates means that the results are only really valid for that specific sample and it is difficult to calculate frequency for a wider area or community.

Sampling design also directly affects the statistical analysis (Dyer *et al.*, 2009; van Heerwaarden *et al.*, 2012). In our previous review, the majority of authors established the farmer's plot as the sampling unit, but each of the papers used a different sampling strategy and thus reached different conclusions about transgene frequency. The statistical method used in Ortiz-García *et al.* (2005a) to estimate sample representativeness and transgene frequency in a population was challenged and the dispute over the most suitable statistical approach revealed problems with the paper's conclusions (Cleveland *et al.*, 2005; Ortiz-García *et al.*, 2005b). This was particularly in terms of what size sampling area would be sufficiently representative to allow conclusions about the entire state of Oaxaca in Mexico to be drawn and how the estimation of effective population size is performed to make claims regarding transgene frequency.

If the objective of sampling is to determine presence or absence of transgenes in an area, landrace populations should be identified with the highest probability of containing them (Cleveland *et al.*, 2005). In those cases, reporting

on the growing or environmental context and seed management practices is essential. On the other hand, if the objective is determining the frequency of transgenes in the landrace population of a given area, it is necessary to use a sampling strategy that maximizes the probability of finding rare alleles in the reference population and is also representative of that population (Cleveland *et al.*, 2005). In this case, it is important to take them from the maximum number of sampling units at each level (for example, numbers of seeds in ears, ears in fields, fields from as broad a range of environments within the reference area as possible). In addition, an equal number of seeds should be sampled from each sampling unit (Cleveland *et al.*, 2005). These are concepts originating from population genetics and rarely applied to investigate transgene presence in centers of origin. Drawing on and making better use of state-of-the-art knowledge in population genetics and dynamics could certainly help design the appropriate sampling strategy for specific transgene monitoring objectives and indicates the potential value of incorporating a range of different disciplines in the development of study method and design.

3.2.2 DNA isolation in heterogeneous samples

DNA-based detection methods such as PCR are undoubtedly the most commonly applied approaches to transgene detection. Nonetheless, they depend on efficient DNA extraction, as well as strict cross-contamination controls to be effective. These are not prerequisites specific to landraces and wild relatives, but applying the standard quality control protocols might pose extra challenges for heterogeneous DNA samples, such as those from landraces (Fig. 3.2).

Piñeyro-Nelson *et al.* (2009b) provided rich evidence of how heterogeneous genetic backgrounds might play a role in DNA extraction efficiency and, consequently, PCR quality. These authors suggested that the results obtained by Ortiz-García *et al.* (2005a) could have involved false negatives due to the expression of a variety of secondary metabolites in landrace samples inhibiting PCR amplification. This argument was not accepted by the commercial laboratory Genetic ID in their reply to Piñeyro-Nelson *et al.* (2009a), who confirmed that they had tested for PCR inhibitors through inhibition tests and did not find any such molecules (Schoel and Fagan, 2009). Uncertainty does remain regarding how much interference these characteristics might have on PCR efficiency in detecting transgenes and internal control genes and how to adapt current methodologies to overcome these.

Routine DNA inhibition tests should be able to estimate any contaminants in extracted DNA (Žel *et al.*, 2012). However, DNA extraction methods and DNA quality standards are frequently only tested in highly homogeneous samples, such as modern maize varieties. One example is the International Organization for Standardization (ISO) method of nucleic acid extraction for the detection of genetically modified organisms and derived products (ISO 21571:2005). In applying such protocols, the DNA extraction from a landrace sample will most likely vary in both amount and quality due to its heterogeneity. The guidance document from the European Network of GMO Laboratories (ENGL) (Hougs *et al.*, 2017) described guidelines for how to introduce and verify a validated method in the laboratory. The document further acknowledged methodological bottlenecks, such as the impact of different instrumentation, PCR reaction mixes, primer concentrations, etc., but did not make any reference to heterogeneous samples, or samples from landraces and wild relatives. For instance, the acceptance criteria for DNA concentration and quality are based on three conditions: (i) the slope of the regression line must be between −3.6 and −3.1; (ii) the coefficient of determination (R^2) is equal to or above 0.98; and (iii) the difference between the measured threshold cycle (Ct) and the extrapolated Ct value (ΔCt) is below 0.5. In heterogeneous samples, these criteria may be very difficult to achieve. This is because slight differences in DNA target and inhibitor concentrations might affect Ct values and therefore perhaps other criteria should be applied to measure extracted DNA quality from such sources.

In a more recent effort, the National Center of Metrology in Mexico started a project related to the development of a certified reference material (CRM) for GMO specific events containing promoter 35S and terminator

Fig. 3.2. A farmer shucking landrace maize cobs harvested from his maize field in Oaxaca region in Mexico. (Photograph by Sarah Z. Agapito-Tenfen, July 2015.)

NOS sequences (Castro Galvan *et al.*, 2019). Although the authors claimed that such CRM would provide standards for the evaluation of GMO dispersion in Mexican maize varieties, it is unclear what genetic background has been used in such materials. The authors further described the material as donated from a commercial company, which suggests biological material derived from elite and not landrace varieties. It is likely that these new CRMs will not fulfill the validation or quality control gap for landrace analysis that has been left by the other already developed CRMs in Europe.

3.2.3 The potential for false negatives and false positives in PCR

Although the ability to detect transgenic constructs is a prerequisite for effective risk analysis, regulation and monitoring of GMOs (Lezaun, 2006), this does not mean that the methods used necessarily operate in a flawless manner, nor that the knowledge available to carry out the methods is always complete. This is especially the case when the detection method is being applied to landraces and wild relatives rather than the GM crop itself or its conventional alternatives.

The ultimate transgenic DNA detection method of PCR is at the same time the most controversial. This is because PCR is an indirect method of DNA detection and relies on primer binding efficiency to target DNA sequences in the genome. Many inhibitory molecules can interfere with primer binding and DNA amplification, consequently leading to false negative results. On the other hand, primers are also prone to unspecific binding to highly homologous, but not identical, sequences, which might then lead to false positive results. False positive results might also arise from cross-contamination of samples during PCR reaction preparation. Therefore, many of the previous studies have performed other forms of confirmatory analysis in addition to PCR, such as Southern blots.

Macarthur *et al.* (2007) applied a model to examine the detection of unauthorized events in oilseed rape (*Brassica napus*), and in particular to explore how heterogeneity in the sampled lot affects the limit of detection (LOD), as well as how LOD values can be modified by choice of sampling plan, analytical replication scheme, and reliable indication of false positive rate. LOD refers to the lowest amount of transgenic material that the method is able to detect. The model of Macarthur *et al.* (2007) reveals three things: (i) that the LOD can vary by a factor of 100, depending on the degree of lot heterogeneity; (ii) that it can vary by an order of magnitude depending on the control plan used to detect and monitor GMOs, such as pooling of sub-samples and/or number of replicates; and (iii) that it can be underestimated by a factor of 20 if it is estimated using validation of the analytical method alone. The authors concluded that these three factors show the importance of an integrated assessment of the whole detection system and consideration of potential lot heterogeneity, which is frequently overlooked in practice (Macarthur *et al.*, 2007). This is particularly the case for commercially prepared PCR kits for GMO detection, which still do not give any particular consideration to heterogeneous samples.

The LOD is always a critical matter. At near-LOD concentrations, there is always a significant risk of false negative results for individual tests (Holst-Jensen *et al.*, 2012). When transgene flow or introgression has taken place in a landrace or wild relative, the copy numbers of PCR targets (i.e. transgene elements) will most likely differ, depending on both the number of transgene events that have occurred and the number of crossings. If two screening targets are present in a GMO but with different insert copy numbers, e.g. one and four, the relative LOD for these will differ fourfold for a DNA solution obtained solely from that GMO (Holst-Jensen *et al.*, 2012). However, this information is always unknown. According to Holst-Jensen *et al.* (2012), the pragmatic approach is therefore to consider not only the observed presence/absence pattern but also the approximate absolute concentration of detected targets. While the latter can be extrapolated from standard curves, it should be remembered that there is a possibility that the target present in the GMO may exhibit slightly divergent PCR performance from the target present in the standards.

The practical evaluation of PCR parameters and acceptance criteria were established in 2008 by ENGL to validate qualitative and quantitative PCR methods, as well as in earlier efforts by others, such as the Food and Agriculture Organization (FAO) from 1998 and the ISO from 2005 (Broeders *et al.*, 2014). In our previous review, none of the published papers on GM detection in maize in Mexico made reference to the application of such international guidelines in their testing. Partly this is simply because these guidelines were not fully developed at the time of the first publications; and for later studies, the authors may not have found them fit for purpose. In fact, as noted in our review, PCR testing was frequently criticized and the call for non-PCR validation was observed in all critiques. International guidelines such as those from ENGL provide a set of parameters and acceptance criteria but do not mention the need for non-PCR methods to validate a PCR method, as was frequently called for in the Mexican case.

3.3 Research Needs and Areas Requiring Attention for Improving Effective Transgene Monitoring in Landraces and Wild Relatives

Monitoring the presence of transgenes depends on analytical methods and their measurement

uncertainties and there is currently no agreed and defined framework or harmonized methods specific for the detection of transgenes in landrace varieties and wild relatives in which low-level presence may be the norm (CBD, 2014). There also seems to be a vacuum in the current international regulatory arena that deals with the potential risks of GMOs to the environment and animal and human health – the Cartagena Protocol on Biosafety under the umbrella of the Convention on Biological Diversity – since it also gives no specific attention to the unique challenges facing GM detection and monitoring in landraces and wild relatives.

Despite the lack of clear international guidelines specific for sampling and testing heterogeneous and dispersed landraces and wild relatives, and the pitfalls and potential misapplications highlighted above, transgene detection methods have developed significantly over the past 15 years. There is now an immensely valuable set of tools, approaches, and international guidelines that continue to develop and provide a way to evaluate transgene presence. To support this ongoing development and to steer it in directions that are particularly useful for addressing the challenges associated with detection and monitoring in landraces and wild relatives, here we offer lessons from our previous review of the published literature and scientific debate on the Mexican maize case for future work in this area.

The determination of a positive or negative result from a transgene testing analysis has two major stages. The first stage is related to the analysis and control of steps prior to the endpoint measurement (sampling strategy, DNA extraction, PCR reaction, cross-contamination, etc.) that might affect the value obtained in the endpoint measurement. In other words, these are the factors affecting the Ct value obtained in real-time experiments, or the presence, absence or even the intensity of a band in a gel, and the confidence in the obtained value. The second stage is the interpretation of the endpoint measurement. We discuss issues connected to both of these stages in turn below.

The start of any study is sampling design. For the challenges associated with environmental sampling, it is important to have established *a priori* what scope of inference the research is going to take and then use this to help identify the appropriate sampling strategy. For environmental sampling and GMO monitoring in landraces or wild relatives, particular attention should also be given to identifying the environmental protection goals that are of interest, including what is important for different stakeholders such as regulators, farmers, researchers, etc. (Wickson *et al.*, 2013; Wickson, 2014). This can help to guide the selection of boundaries for the environmental sampling. A classification of ecoregions for environmental and GMO monitoring networks was proposed by Graef *et al.* (2005). These authors suggested that a variety of ecological data should be combined with socio-economic data (such as land use), to help integrate the spatial and temporal complexity and heterogeneity of ecosystems being surveyed and thus provide more reliable, accurate and reproducible data. Rather than random sampling, sampling in such an approach is a stratified procedure integrating different layers of information. The authors also highlighted that GMO monitoring networks must be integrated with agricultural data and might also require adaptation and improvement with regards to implementation over long periods of time and large areas, such as in many agriculture-based regions or countries (Graef *et al.*, 2005).

An important element of environmental monitoring is the temporal factor. The previous studies on transgene detection in Mexican maize landrace, particularly the Piñeyro-Nelson *et al.* (2009a) study, showed temporal dependence of transgene frequency. Especially for land use data, in which the agricultural landscape might change more frequently even within a single year (depending on the crop season), integration of a temporal dimension to environmental sampling strategies and design can be learned from landscape structure analysis (Walz *et al.*, 2016). Indeed, a recent monitoring effort in Mexico that analyzed samples from 2008 until 2018 revealed a highly dynamic spatio-temporal replacement over time, mainly due to introduction of foreign material, which must be taken into account when designing monitoring strategies (Rendón-Aguilar *et al.*, 2019).

Environmental sampling strategies need to try to reflect the geographic heterogeneity of occurrence and agricultural use of the species in question in order to estimate percentage and frequency of transgene presence. This is,

of course, highly challenging in the context of crops and seed systems, and indeed handling this heterogeneity is arguably the primary challenge facing a determination of transgene frequency using environmental sampling. Transgene frequency must be determined based on the population size that has been sampled. Sampled material (e.g. seeds, cobs or leaves) directly affects statistical calculation of transgene frequency and, therefore, inference of transgene presence in a region or sample. Cleveland *et al.* (2005) described the maternal effect (when seeds are taken from a few cobs, versus seeds taken from a seed lot or a mixture of many harvested cobs) on transgene frequency estimation. This is because seeds from the same cob are overrepresenting the maternal genetic contribution but also because, in maize, each seed contains layers of tissues (e.g. endosperm, seedcoat, embryo) with different genomic material. Therefore, since maize genetic structure directly influences DNA-based (e.g. PCR) detection and quantitation, DNA extraction from single plants can provide different results from DNA extractions from a pool of plants (or seeds). As a consequence, plant tissue or organ composition can influence the establishment of a link between the way GM concentrations are determined in seed, feed, food, and in the environment (Holst-Jensen *et al.*, 2006). Several guidelines were provided in Trifa and Zhang (2004), Cleveland *et al.* (2005), and also in Holst-Jensen *et al.* (2006) for calculating effective population size and transgene frequency depending on sample material and size. It is important to realize that each of these proposed guidelines will need to be considered and tailored to the specific material and sample size in use, and that it may not be possible to harmonize a single ideal across them.

Spatio-temporal criteria have also recently been discussed with regard to the risk assessment of GM plants that can persist and propagate in the environment, such as maize. The environmental risk assessment of such plants cannot be reduced to the specific traits and characteristics known at the time of application, since the assessment also needs to consider effects that can emerge after a number of generations, in other genetic backgrounds, and under specific environmental conditions (Bauer-Panskus *et al.*, 2020). The consideration of gene flow between elite and landrace varieties and wild relatives should include comprehensive detection protocols that can anticipate challenges in monitoring. This is arguably the responsibility of the developers of the technology and owners of intellectual property rights.

For the challenges associated with DNA extraction and PCR inhibition, the key issue is to ensure high-quality DNA. Throughout the years, different methods for GMO detection have been validated using certified reference materials that are in the form of powdered grain material. However, everyday routine laboratories must perform GM detection on a wide variety of sample matrixes. In such cases, molecules of plant origin or from other sources that affect PCR amplification will influence the reliability of transgene detection and monitoring. The extraction method is therefore key to ensuring high yield and quality of the DNA obtained and must be carefully selected or adapted.

Cankar *et al.* (2006) analyzed the effects of DNA extraction methods and sample matrices on quantification of GMOs. These authors tested four maize and four soybean samples and found crucial influences on the results of GMO quantification from the extraction technique and sample matrix properties. Although they suggested the development of appropriate extraction techniques for each matrix, the authors also highlighted that, for samples with certain compositional specificities, it will be impossible to define strict quality controls (such as acceptance criteria to compare the efficiency of the sample with that of the standard curve) to be introduced to monitor PCR (Cankar *et al.*, 2006). Adaptation of DNA extraction protocols for heterogeneous samples will then have to be done by different groups and is most likely unable to be harmonized for all GMO detection work. In addition, new quality control measurements might have to be developed, since current amplification of an endogenous positive control gene, as an indicator for the absence of PCR inhibitors, is not always valid (Holden *et al.*, 2003).

Much can be learned from other sciences that deal with genetic analysis of non-model species as well as from food science, which frequently works with highly processed and altered materials. Schrader *et al.* (2012) reviewed the literature providing general guidelines for the removal of PCR inhibitors that could also be applied to heterogeneous samples being tested for

GMO content. In fact, while it arguably always goes back to the assessment of PCR inhibitors by PCR control reactions, for most of the inhibitory substances practicable tools for their analysis are yet to be developed and this could be an important area for future research in this field (Schrader *et al.*, 2012).

Turning to the second stage of GM detection work, the interpretation of endpoint measurements, it is relevant to note that an endpoint measurement can be interpreted as a positive or negative result, depending on the threshold level established for the analysis. Generally, most GMO detection and identification methods have been developed to meet the purpose of fitting into a labeling requirement or law. This means that the threshold levels are set to the labeling requirement (for example, 0.9% presence of a GMO in a food sample) and the tested sample is determined positive if equal to or above that limit and negative if below that limit. Holst-Jensen *et al.* (2006) discussed coherence between legal requirements and analytical approaches for detection of GMOs and recommended key points where coherence should be developed. These include: (i) the definition of units of measurements; (ii) expression of GM material quantities; (iii) terminology; and (iv) inconsistent legal status of products derived from related but slightly different transformation routes. Although some improvement has been achieved in this area (such as the decision by the European Commission to recommend the use of DNA ratios to express GMO quantity (Commission Regulation (EU) No 619/2011)), the debate still continues to focus on GM labeling laws and does not necessarily include GM detection for environmental monitoring. Thus, such analytical coherence is not necessarily fit for the purpose of analysing GM content in landraces and wild relatives.

Given that Europe has established strict requirements for GM regulation, detection and labeling, significant development in setting guidelines for testing and endpoint measurements has been achieved by the ENGL, hosted by the Joint Research Centre (JRC) of the European Commission. However, this work has generally been developed for application within the European context and therefore does not include guidelines for GM detection in landraces and wild relatives, nor does it take account of the socio-political realities of other regions or countries in which this issue may be of lower cultural or economic priority. Whereas low-level presence (for example, Ct values close to final PCR cycles) might not generate labeling concerns, it might well be relevant for environmental risk assessment. The ENGL guidance document 'Overview on the detection, interpretation and reporting on the presence of unauthorized genetically modified materials' (Arne *et al.*, 2011) recognized the absence of a solid reference framework for the interpretation and reporting of results on unauthorized GM presence and recommended primarily focusing on reliability. The document further provided a brief general consideration on convergence between datasets and ruling out or minimizing the possible occurrence of false positives or negatives. Therefore, we propose that future work would benefit significantly from the development of guidelines beyond a European context on how to proceed to verify and confirm positive results for the types of low-level presence that are likely to be typical in the case of contamination of land races and wild relatives in centres of origin and diversity.

Finally, given the debate and learning that has occurred in the Mexican maize case, we propose that future research on transgene detection and monitoring in landraces and wild relatives should also work to include the collection of socio-cultural data to help develop a comprehensive view of gene flow and its influence on both maize and human populations (Fig. 3.3). This collection of socio-cultural data in addition to samples for genetic testing could include information such as how seeds are acquired and selected. Questions of relevance here would be whether seed is saved or purchased, the extent to which grain sold for consumption may also be planted, whether seeds are exchanged within the community and with other communities, whether migrating community members ever return with new seeds, what characteristics are favoured in processes of seed selection, whether planting time varies across plots or varieties, etc. This type of information can usefully help to build a complementary understanding of local seed systems and management practices in the areas being sampled. It can also help build an understanding of how the work and choices of farmers can affect landrace population structure and dynamics, as well as how such practices may work to contain or spread

Fig. 3.3. A Mexican farmer walking with maize tortillas, the main food source in the community located in the Oaxaca region in Mexico. (Photograph by Sarah Z. Agapito-Tenfen, July, 2015.)

transgene contamination (Agapito-Tenfen *et al.*, 2017).

It is clear that significant work remains to be done to optimize and potentially standardize GM detection and monitoring methods for working with landraces and wild relatives in which low levels of presence can be the norm. It is also clear that, although the field has evolved significantly and useful new techniques have been developed to detect even smaller traces of transgenes in a sample, the cost of applying such methods across a broad enough sample to be representative is still a major inhibiting factor. It is certainly questionable to what extent the farmers growing landraces of maize are able to afford to carry out such analyses, and to what extent the governments of centers of crop origin and diversity are willing and able to pay for testing of either landraces or wild relatives. The situation becomes even more challenging and urgent as we realize that these problems exist and persist for the GMOs that are commercially available today while new GMO events continue to be approved all the time.

Although the number of GM events has increased from less than ten in 1998 to over 500 today (Convention on Biological Diversity's Biosafety Clearing House database, available at https://bch.cbd.int/database/lmo-registry/, accessed 12 July 2021), only basic information on each event is available and this does not include the level of sequence information that is needed to develop detection methods (due to claims to confidential business information) (Nielsen, 2013). It is also relevant to note that only parties to the Cartagena Protocol have a legal duty to release information about new GMO approvals and that the largest GMO adopter – the USA – is not a party to the Protocol. To advance the ability to monitor for GM spread effectively, regulatory agencies need to make information publicly available on all GM events that they approve; and developers need to make information on transgene sequences and reference materials available.

Furthermore, it is important to note that while the first wave of GMOs was relatively similar (for example, all using the same CaMV 35S promoter, which allowed for a broad-based screening in a non-event-specific manner), the field is evolving quickly and new GMOs share little similarity, making it harder to detect and monitor a wide range of events by using a screening matrix. The field of biotechnology also now uses techniques beyond recombinant DNA technology, and not all GMOs are made to contain a transgene (e.g. gene-edited crops) or to produce a transgenic protein (e.g. dsRNA-based GMOs) (Heinemann *et al*., 2013). Crops resulting from these techniques can be difficult to distinguish from conventionally bred crops or from crops produced by natural genetic variation, and their identification therefore poses a whole new level of complexity for the field in the future. With these new developments of GMOs, sampling, detection and identification become only more challenging and complex. It is therefore urgent that work to develop robust and reliable methods for the detection of transgenes in landraces and wild relatives advances at a similar pace as the development of GM technologies.

3.4 Conclusion

The importance of conserving genetic biodiversity in important agricultural crop plants and the rapid expansion of biotechnological techniques and GM organisms makes establishing good practices for transgene detection and monitoring in landraces and wild relatives only more urgent and pressing now than it has been in the past. In this chapter, we have reviewed key methodological challenges for effective transgene detection and monitoring, including appropriate methods for environmental sampling and the scale of inference possible on the basis of this, the ability to isolate high-quality DNA from heterogeneous samples using established methods, and the potential for PCR methods to lead to both false negative and false positive results.

The scientific debate over what constitutes a reliable and effective method for DNA detection in landraces and wild relatives where low-level presence may be the norm, as well as how to interpret variation in results generated using different methods, only serves to perpetuate public and policy uncertainty and debate in an area that is already highly controversial. As biotechnologies continue to evolve and expand, it is becoming increasingly urgent that methods for detecting transgenes in landraces and wild relatives, and particularly low levels of presence, are further discussed, developed and supported so as to better address the unique challenges involved. It is crucial that the limitations of existing approaches are transparently acknowledged and recognized, both within the scientific community and also within national and international policy contexts.

Learning from more than 15 years of scientific research and debate on the iconic case of transgene flow into landraces of maize in Mexico has revealed that future work in this field will benefit from: (i) more dedicated approaches to environmental sampling (which are followed up with ongoing monitoring over time); (ii) better guidance on establishing limits of detection and enhancing the ability to detect low-level traces of transgenes; (iii) improved validation of the results (for example, using both inter-laboratory validation and multiple methods); (iv) the setting of a minimum level of information for publication of transgene detection and monitoring analysis in wild relatives and landraces (with guidelines targeting the reliability of results to help ensure the integrity of the scientific literature); (v) promoting the importance of developing consistency between

studies, an increase in transparency and reporting on experimental methods and reasons why choices for one or another were made; and (vi) increased support and resources dedicated to the ongoing development and application of detection strategies. Without such improvements in this field, the future of GM detection and monitoring in landraces and wild relatives is almost certainly going to be one of continued controversy, uncertainty and debate.

Acknowledgements

We would like to acknowledge the support of the indigenous communities of Oaxaca that participated in our project for sharing their knowledge, seeds and insights in the context of maize farming in Mexico. We would also like to thank Dr Gretta Abou-Sleymane, Dr Miluse Trtikova and PhD student Narmeen Mallah for insightful discussions about transgene detection methodologies during a project workshop. This work was supported by the Research Council of Norway under the LATINAMERIKA Program [grant number 236840].

References

Agapito-Tenfen, S., Lopez, F.R., Mallah, N., Abou-Slemayne, G., Trtikova, M. *et al*. (2017) Transgene flow in Mexican maize revisited: Socio-biological analysis across two contrasting farmer communities and seed management systems. *Ecology and Evolution* 7(22), 9461–9472. DOI: 10.1002/ece3.3415.

Agapito-Tenfen, S.Z. and Wickson, F. (2018) Challenges for transgene detection in landraces and wild relatives: learning from 15 years of debate over GM maize in Mexico. *Biodiversity and Conservation* 27(3), 539–566. DOI: 10.1007/s10531-017-1471-0.

Arne, H., Yves, B., Theo, A., Hermann, B. and Marc, D.L. (2011) *Overview on the detection, interpretation and reporting on the presence of unauthorised genetically modified materials: Guidance document from the European Network of GMO Laboratories (ENGL)*. Publication Office of the European Union, Luxembourg. Available at: http://gmo-crl.jrc.ec.europa.eu/ (accessed 12 July 2021)

Baker, L.E. (2013) *Corn Meets Maize: Food Movements and Markets in Mexico*. Rowman & Littlefield Publishers, Lanham, Maryland.

Bauer-Panskus, A., Miyazaki, J., Kawall, K. and Then, C. (2020) Risk assessment of genetically engineered plants that can persist and propagate in the environment. *Environmental Sciences Europe* 32(1), 32. DOI: 10.1186/s12302-020-00301-0.

Bell, S. (2001) *A Beginner's Guide to Uncertainty of Measurement*. National Physical Laboratory (NPL), Teddington, UK.

Bello, W. (2009) *Food Wars*. Verso, New York.

Broeders, S., Huber, I., Grohmann, L., Berben, G., Taverniers, I. *et al*. (2014) Guidelines for validation of qualitative real-time PCR methods. *Trends in Food Science & Technology* 37(2), 115–126. DOI: 10.1016/j.tifs.2014.03.008.

Cankar, K., Stebih, D., Dreo, T., Zel, J. and Gruden, K. (2006) Critical points of DNA quantification by real-time PCR--effects of DNA extraction method and sample matrix on quantification of genetically modified organisms. *BMC Biotechnology* 6, 37. DOI: 10.1186/1472-6750-6-37.

Carro-Ripalda, S. and Astier, M. (2014) Silenced voices, vital arguments: smallholder farmers in the Mexican GM maize controversy. *Agriculture and Human Values* 31(4), 655–663. DOI: 10.1007/s10460-014-9533-3.

Carro-Ripalda, S., Astier, M. and Artia, P. (2015) An analysis of the GM crop debate in Mexico. In: Macnaghten, P. and Carro-Ripalda, S. (eds) *Governing Agricultural Sustainability: Global Lessons from GM Crops*. Routledge, London, pp. 33–73.

Castro Galvan, E., Maldonado Torres, M. and Pérez Urquiza, M. (2019) *Zea maize* reference materials for genetically modified organism detection in Mexico. *Ecology and Evolution* 9(21), 12353–12356. DOI: 10.1002/ece3.5667.

CBD (2014) *Summary of the Activities under the Electronic Network of Laboratories for the Detection and Identification of Living Modified Organisms (2012–2014)*. UN Environment Programme, Convention for Biodiversity. Available at: http://www.cbd.int/doc/meetings/bs/mop-07/information/mop-07-inf-09-en.pdf (accessed 6 October 2020)

Christophe Bonneuil., Foyer, J. and Wynne, B. (2014) Genetic fallout in biocultural landscapes: molecular imperialism and the cultural politics of (not) seeing transgenes in Mexico. *Social Studies of Science* 44(6), 901–929. DOI: 10.1177/0306312714548258.

Christou, P. (2002) No credible scientific evidence is presented to support claims that transgenic DNA was introgressed into traditional maize landraces in Oaxaca, Mexico. *Transgenic Research* 11(1), iii–v. DOI: 10.1023/A:1013903300469.

Cleveland, D.A., Soleri, D., Cuevas, F.A., Crossa, J. and Gepts, P. (2005) Detecting (trans)gene flow to landraces in centers of crop origin: lessons from the case of maize in Mexico. *Environmental Biosafety Research* 4(4), 197–208. DOI: 10.1051/ebr:2006006.

Dalton, R. (2009) Mexico's transgenic maize under fire. *Nature* 462(7272), 404. DOI: 10.1038/462404a.

Delborne, J.A. (2008) Transgenes and transgressions: scientific dissent as heterogeneous practice. *Social Studies of Science* 38(4), 509–541. DOI: 10.1177/0306312708089716.

Dyer, G.A., Serratos-Hernández, J.A., Perales, H.R., Gepts, P., Piñeyro-Nelson, A. *et al*. (2009) Dispersal of transgenes through maize seed systems in Mexico. *PLoS ONE* 4(5), e5734. DOI: 10.1371/journal.pone.0005734.

Garcia, D.A. (2017) Monsanto sees prolonged delay on GMO corn permits in Mexico. Thomson Reuters Environment blog, 30 January. Available at: www.reuters.com/article/us-mexico-monsanto-idUSKBN15E1DJ (accessed 6 October 2020).

Gilbert, N. (2013) Case studies: a hard look at GM crops. *Nature* 497(7447), 24–26. DOI: 10.1038/497024a.

Graef, F., Schmidt, G., Schröder, W. and Stachow, U. (2005) Determining ecoregions for environmental and GMO monitoring networks. *Environmental Monitoring and Assessment* 108(1–3), 189–203. DOI: 10.1007/s10661-005-3966-5.

Heinemann, J.A., Agapito-Tenfen, S.Z. and Carman, J.A. (2013) A comparative evaluation of the regulation of GM crops or products containing dsRNA and suggested improvements to risk assessments. *Environment International* 55, 43–55. DOI: 10.1016/j.envint.2013.02.010.

Holden, M.J., Blasic, J.R. Jr., Bussjaeger, L., Kao, C., Shokere, L.A. *et al*. (2003) Evaluation of extraction methodologies for corn kernel (*Zea mays*) DNA for detection of trace amounts of biotechnology-derived DNA. *Journal of Agricultural and Food Chemistry* 51(9), 2468–2474. DOI: 10.1021/jf0211130.

Holst-Jensen, A., De Loose, M. and Van den Eede, G. (2006) Coherence between legal requirements and approaches for detection of genetically modified organisms (GMOs) and their derived products. *Journal of Agricultural and Food Chemistry* 54(8), 2799–2809. DOI: 10.1021/jf052849a.

Holst-Jensen, A., Bertheau, Y., de Loose, M., Grohmann, L., Hamels, S. *et al*. (2012) Detecting unauthorized genetically modified organisms (GMOs) and derived materials. *Biotechnology Advances* 30(6), 1318–1335. DOI: 10.1016/j.biotechadv.2012.01.024.

Hougs, L., Gatto, F., Goerlich, O., Grohmann, L. and Lieske, K. (2017) *Verification of Analytical Methods for GMO Testing When Implementing Interlaboratory Validated Methods*. Publication Office of the European Union, Luxembourg.

Kaplinsky, N., Braun, D., Lisch, D., Hay, A., Hake, S. *et al*. (2002) Biodiversity (Communications arising): Maize transgene results in Mexico are artefacts. *Nature* 416(6881), 601–602. DOI: 10.1038/nature739.

Lezaun, J. (2006) Creating a new object of government: making genetically modified organisms traceable. *Social Studies of Science* 36(4), 499–531.

Macarthur, R., Murray, A.W., Allnutt, T.R., Deppe, C., Hird, H.J. *et al*. (2007) Model for tuning GMO detection in seed and grain. *Nature Biotechnology* 25(2), 169–170. DOI: 10.1038/nbt0207-169.

McAfee, K. (2008) Beyond techno-science: transgenic maize in the fight over Mexico's future. *Geoforum* 39(1), 148–160. DOI: 10.1016/j.geoforum.2007.06.002.

Mercer, K.L. and Wainwright, J.D. (2008) Gene flow from transgenic maize to landraces in Mexico: an analysis. *Agriculture, Ecosystems & Environment* 123(1-3), 109–115. DOI: 10.1016/j.agee.2007.05.007.

Mercer, K.L., Perales, H. and Wainwright, J. (2012) Climate change and the transgenic adaptation strategy: smallholder livelihoods, climate justice, and maize landraces in Mexico. *Global Environmental Change* 22(2), 495–504. DOI: 10.1016/j.gloenvcha.2012.01.003.

Metz, M. and Fütterer, J. (2002) Biodiversity (Communications arising): Suspect evidence of transgenic contamination. *Nature* 416(6881), 600–602. DOI: 10.1038/nature738.

Nielsen, K.M. (2013) Biosafety data as confidential business information. *PLoS Biology* 11(3), e1001499. DOI: 10.1371/journal.pbio.1001499.

Ortiz-García, S., Ezcurra, E., Schoel, B., Acevedo, F., Soberón, J. *et al*. (2005a) Absence of detectable transgenes in local landraces of maize in Oaxaca, Mexico (2003–2004). *Proceedings of the National Academy of Sciences of the United States of America* 102(35), 12338–12343. DOI: 10.1073/pnas.0503356102.

Ortiz-García, S., Ezcurra, E., Schoel, B., Acevedo, F., Soberón, J. *et al*. (2005b) Reply to Cleveland et al.'s 'Detecting (trans)gene flow to landraces in centers of crop origin: lessons from the case of maize in Mexico'. *Environmental Biosafety Research* 4, 209–215. DOI: 10.1051/ebr:2006007.

Piñeyro-Nelson, A., Van Heerwaarden, J., Perales, H.R., Serratos-Hernández, J.A., Rangel, A. *et al*. (2009a) Transgenes in Mexican maize: molecular evidence and methodological considerations for GMO detection in landrace populations. *Molecular Ecology* 18(4), 750–761. DOI: 10.1111/j.1365-294X.2008.03993.x.

Piñeyro-Nelson, A., Van Heerwaarden, J., Perales, H.R., Serratos-Hernández, J.A., Rangel, A. *et al*. (2009b) Resolution of the Mexican transgene detection controversy: error sources and scientific practice in commercial and ecological contexts. *Molecular Ecology* 18(20), 4145–4150. DOI: 10.1111/j.1365-294X.2009.04369.x.

Quist, D. (2007) Vertical (trans)gene flow: implications for crop diversity and wild relatives. In: Traavik, T. and Lim, L.C. (eds) *Biosafety First*. Tapir Academic Press, Trondheim, Norway, pp. 1–14.

Quist, D. and Catacora-Vargas, G. (2011) *Transgenes in Mexican Maize, Ten Years On: Still not addressing the right questions*. GenØk Biosafety Report 2011/04. GenØk Centre for Biosafety, Tromsø, Norway.

Quist, D. and Chapela, I.H. (2001) Transgenic DNA introgressed into traditional maize landraces in Oaxaca, Mexico. *Nature* 414(6863), 541–543. DOI: 10.1038/35107068.

Quist, D. and Chapela, I. (2002) Suspect evidence of transgenic contamination/Maize transgene results in Mexico are artefacts (see editorial footnote). *Nature* 416(6881), 602. DOI: 10.1038/nature740.

Rendón-Aguilar, B., Bernal-Ramírez, L.A., Bravo-Avilez, D. and Rocha-Munive, M.G. (2019) Temporal dynamics of detected transgenes in maize landraces in their center of origin. *Revista Mexicana de Biodiversidad* 90, e902653. DOI: 10.22201/ib.20078706e.2019.90.2653.

Rivera López, F., Wickson, F. and Helen Hausner, V. (2020) Bridging different perspectives for biocultural conservation: art-based participatory research on native maize conservation in two Indigenous farming communities in Oaxaca, Mexico. *Environment, Development and Sustainability* 22(8), 7427–7451. DOI: 10.1007/s10668-019-00530-1.

Ruiz, M.T.G., Knapp, A.N. and Garcia-Ruiz, H. (2018) Profile of genetically modified plants authorized in Mexico. *GM Crops & Food* 9(3), 152–168. DOI: 10.1080/21645698.2018.1507601.

Sarewitz, D. (2004) How science makes environmental controversies worse. *Environmental Science & Policy* 7(5), 385–403. DOI: 10.1016/j.envsci.2004.06.001.

Schoel, B. and Fagan, J. (2009) Insufficient evidence for the discovery of transgenes in Mexican landraces. *Molecular Ecology* 18(20), 4143–4144. DOI: 10.1111/j.1365-294X.2009.04368.x.

Schrader, C., Schielke, A., Ellerbroek, L. and Johne, R. (2012) PCR inhibitors – occurrence, properties and removal. *Journal of Applied Microbiology* 113(5), 1014–1026. DOI: 10.1111/j.1365-2672.2012.05384.x.

Serratos-Hernández, J.-A., Islas-Gutiérrez, F., Buendía-Rodríguez, E. and Berthaud, J. (2004) Gene flow scenarios with transgenic maize in Mexico. *Environmental Biosafety Research* 3(3), 149–157. DOI: 10.1051/ebr:2004013.

Serratos-Hernández, J.-A., Gómez-Olivares, J.-L., Salinas-Arreortua, N., Buendía-Rodríguez, E., Islas-Gutiérrez, F. *et al*. (2007) Transgenic proteins in maize in the soil conservation area of federal district, Mexico. *Frontiers in Ecology and the Environment* 5, 247–252. DOI: 10.1890/1540-9295(2007)5[247:TPIMIT]2.0.CO;2.

Stewart, C.N. Jr., Halfhill, M.D. and Warwick, S.I. (2003) Transgene introgression from genetically modified crops to their wild relatives. *Nature Reviews Genetics* 4(10), 806–817. DOI: 10.1038/nrg1179.

Trifa, Y. and Zhang, D. (2004) DNA content in embryo and endosperm of maize kernel (*Zea mays* L.): impact on GMO quantification. *Journal of Agricultural and Food Chemistry* 52(5), 1044–1048. DOI: 10.1021/jf034574+.

van Heerwaarden, J., Ortega Del Vecchyo, D., Alvarez-Buylla, E.R. and Bellon, M.R. (2012) New genes in traditional seed systems: diffusion, detectability and persistence of transgenes in a maize metapopulation. *PLoS ONE* 7(10), e46123. DOI: 10.1371/journal.pone.0046123.

Vargas-Parada, L. (2014) Gm maize splits Mexico. *Nature* 511(7507), 16–17. DOI: 10.1038/511016a.

Wainwright, J. and Mercer, K.L. (2011) Transnational transgenes: the political ecology of maize in Mexico. In: Peet, R., Robbins, P. and Watts, M. (eds) *Global Political Ecologies*. Routledge, New York, pp. 412–429.

Walz, U., Hoechstetter, S., Drăguţ, L. and Blaschke, T. (2016) Integrating time and the third spatial dimension in landscape structure analysis. *Landscape Research* 41(3), 279–293. DOI: 10.1080/01426397.2015.1078455.

Wickson, F. (2014) Environmental protection goals, policy and publics in the European regulation of GMOs. *Ecological Economics* 108(23), 269–273. DOI: 10.1016/j.ecolecon.2014.09.025.

Wickson, F., Bøhn, T., Wynne, B., Hilbeck, A. and Funtowicz, S. (2013) Science-based risk assessment requires careful evaluation of all studies. *Nature Biotechnology* 31(12), 1077–1078. DOI: 10.1038/nbt.2761.

Žel, J., Milavec, M., Morisset, D., Plan, D. and Van den Eede, G. (2012) *How to Reliably Test for GMOs*. Springer, Boston, Massachusetts.

4 Transgenic Poplar Gene Flow Monitoring in China

Lei Zhang and Jian-Jun Hu*
State Key Laboratory of Tree Genetics and Breeding, Research Institute of Forestry, Chinese Academy of Forestry, Beijing, China

Abstract

Poplar is cultivated widely for pulpwood, firewood, and timber. Transgenic poplar may be part of a solution for wood demand in China. Because transgene escape is an important part of ecological security evaluation of transgenic plants, in this chapter we discuss a real transgenic poplar case study. In this case study, mature transgenic male *Populus nigra* plants harbored a *Bacillus thuringiensis* toxin gene (i.e. *Bt* poplar). A plantation of these plants served as a testbed for a relevant example for gene flow monitoring in China. Furthermore, we discuss environmental risk assessment (ERA) of these transgenic plants. While transgenes can drift to related species through natural and controlled pollination, the probability of transgene drift appears to be very low in the field. The resultant *Bt* poplar seeds occurred at a frequency from about 0.15% at 0 m to about 0.02% at 500 m away from the *Bt* poplar. The *Bt* poplar progeny seeds had decreased germination within 3 weeks in the field (from 68% to 0%), compared with the 48% germination rate after 3 weeks at 4°C. The survival rate of seedlings in the field was 0% without any treatments, but increased to 1.7% under four combined treatments (clean and trim, watering, weeding, and cover with plastic to retain moisture) after being seeded in the field for 8 weeks. Hybrid offspring appeared to possess segregated traits following artificially controlled pollination. While hybrids of transgenic poplar and non-transgenic poplar can be excellent germplasm, gene flow should be monitored. Transgene expression in grafted scion and rootstock of transgenic poplar is reviewed. The transgenic poplar studied appears to be safe; no ecological or environmental harm has been observed in China.

Keywords: biosafety; environment; gene flow; monitoring; transgenic poplar

4.1 Transgenic Poplar in China

Poplar is an important forest tree crop worldwide. In China, for instance, about 7 million hectares of poplar plantation have been established for shelter forest and timber production (Fang, 2008). Insects cause severe damage to poplar plantations, estimated at millions of US dollars annually. The infected trees are usually sprayed with pesticide or cut down in order to control the infestation, thus resulting in serious economic and ecological consequences (Chen *et al.*, 2009). Genetic engineering provides a promising tool to breed superior poplar clones with improved insect resistance, salt tolerance, and wood properties. Salt tolerance is especially

*Corresponding author: hujj@caf.ac.cn

 Gene Flow: Monitoring, Modeling and Mitigation (eds Wei Wei and C. Neal Stewart, Jr)
DOI: 10.1079/9781789247480.0004

of interest since there is an area of 3.67×10^7 ha of salinized land in China that could otherwise produce trees.

As of 2012, there were 34 different transgenic poplar types growing in confined field trials (Häggman *et al.*, 2013). For example, insect-tolerance transgenic *P. nigra* (lines n12, n153 and n172) and *P. alba* × (*P. davidiana* + *P. simonii*) × *P. tomentosa* with *Bacillus thuringiensis* toxin (*Bt*) gene were planted for field tests, and proved to be effective in avoidance of damage by leaf insects (Hu *et al.*, 2001). *P. xiaozhannica* 'Balizhuangyang' with 1-p-mannitol dehydrogenase (*mtl-D*) could grow normally in mid-saline soil (salt concentration about 0.3–0.4%) (Yin *et al.*, 2004). The salt tolerance trial of 3-year-old transgenic *P. alba* × *P. berolinensis* with jasmonic ethylene responsive factor (*JERF*) gene were 14.5% and 33.6% taller than the control plants at two field sites with 0.3% salt concentration (Li *et al.*, 2009). Antisense of 4-coumarate: CoA ligase, a key gene in lignin biosynthesis, significantly reduced the lignin content in *P. tomentosa* (Jia *et al.*, 2004). In recent years many wood formation regulators such as *PtoMYB74*, *PtoMYB152*, *PtoMYB170* and *PtoMYB216* have been discovered that regulate secondary cell wall lignin biosynthesis (Tian *et al.*, 2013; Li *et al.*, 2014, 2018; Xu *et al.*, 2017).

The data of the International Service for the Acquisition of Agri-biotech Applications (ISAAA) showed that 29 countries planted 190.4 million hectares of biotech crops – a slight decline of 0.7% from 191.7 million hectares in 2018 (ISAAA, 2019). China has approved 60 biotech crop events for food and feed use and cultivation, including maize (17 events), Argentine canola (12), cotton (10), soybean (10), tomato (3), poplar (2), rice (2), papaya (1), petunia (1), sugar beet (1) and sweet pepper (1). From 2013 to 2016, a total of 543 ha of genetically modified poplar were planted in China (ISAAA, 2016).With the development of transgenic poplars, the issue of ecological safety caused by transgenic trees has attracted increasing attention (Hu *et al.*, 2014). Transgenic poplar is subject to government approval for small-scale field testing, environmental release, and pilot-scale production and commercialization in China (Wang *et al.*, 2018). The safety evaluation of transgenic organisms mainly includes the evaluation of: (i) the survival competitiveness of transgenic plants, mainly to evaluate whether the transgenic plants have the ability to become weeds; (ii) vertical gene transfer (VGT), mainly to evaluate the possibility of foreign genes transferred into wild relatives through gene flow; (iii) horizontal gene transfer (HGT), mainly to evaluate the transfer of foreign genes to other animals, plants, insects and microorganisms; (iv) the impact of transgenic plants on the target and non-target insects; and (v) the impact of transgenic plants on the soil ecosystem, mainly evaluating the possible impact of exogenous gene products on the species and quantity of soil microorganisms.

In China, transgenic poplar has been planted mainly in the northern areas, where the climate is dry or semi-dry with annual precipitation ranging from 100 mm to 600 mm, concentrated in July and August. Poplar seed dispersal occurs during May to June and thus misses the rainy season, which is favorable for seed germination. The successful establishment of poplar plantations in these areas is dependent on the quality of seedlings planted and the availability of irrigation. No natural regenerated poplar stands have been found and no natural regeneration of poplar plants has been observed in and around plantations. Taken together, it is considered that transgenic poplar plantations will have little chance to spread transgenes out of plantations; the National Forestry Administration has thus granted permission for the commercialization of transgenic poplar clones in six provinces in China that are located far from natural poplar stands (Fig. 4.1).

4.2 Gene Flow in Transgenic Poplar

At present, a large number of studies have shown that it is difficult to avoid gene drift in transgenic plants. Insect-resistant cotton has been shown to spread *Bt* transgenes to wild cotton (Wegier *et al.*, 2011). Transgenic canola genes have been found to be spread to wild turnip, which has resulted in multi-herbicide resistance in the field (Hall *et al.*, 2009; Pandolfo *et al.*, 2018). Thus, a major concern regarding the commercialization of transgenic forest trees is transgene flow, through which the transgenes may spread from

Fig. 4.1. Plantations of (A) *Bt* black poplar and (B) *Bt* white poplar (Renqiu, Hebei province, China). (Photographs by Jianjun Hu, May 2015.)

transgenic trees in plantations to trees in natural forests. Although gene flow is a very common natural phenomenon and an important process of evolution, the movement of transgenes from transgenic plants to wild plants may be considered undesirable with regard to environmental consequences. Fears include the unintended production of more aggressive weeds, loss of germplasm, and altered populations of non-target organisms, with biodiversity losses.

Trees are long-lived, and their exposure to the environment occurs over a much longer time than annual crops (Ahuja, 2009, 2010). Poplar, a fast-growing tree with 8–15 years in rotation, has been used for genetic transformation both for research and for breeding (Strauss *et al.*, 2004; Hu *et al.*, 2010). Cross-pollination of plantation poplars with their wild relatives is of major concern, due to the broad compatibility within and among species, even those belonging to different sections of *Populus* (Eckenwalder, 1996). Several modeling and simulation methods have been proposed to measure the gene flow between poplar stands (Slavov *et al.*, 2009; Niggemann *et al.*, 2012) and co-dominant markers have been developed for monitoring gene flow among related species (Khasa *et al.*, 2005). However, no empirical data on gene flow in transgenic poplar plantations are available at this time.

4.2.1 Gene flow of transgenic *P. nigra* with *Bt*

4.2.1.1 Analysis of T-DNA insertion loci of transgenic poplar

T-DNA with *BtCry1Ac* toxin gene sequence (1.8, 2.1 or 3.6 kb from the 5′- end) was transferred into *P. nigra* (Fig. 4.2A) (Wang *et al.*, 1996). Ten-year-old transgenic lines showed faster growth and higher lignin and lower cellulose than control check (CK) and native cultivars in Xinjiang Uygur Autonomous Region.

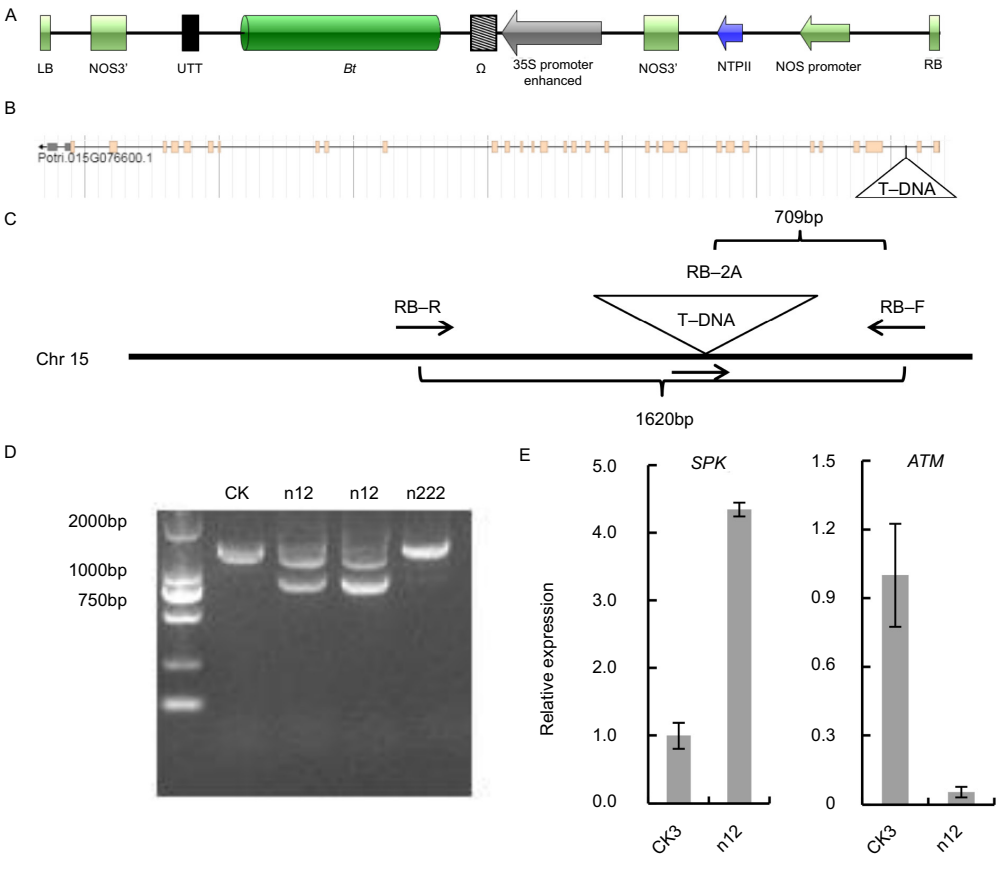


Fig. 4.2. T-DNA insertion loci analysis and specific detection method of transgenic insect-resistant poplar line n12. (A) T-DNA diagram and (B) insertion site of n12. (C) Primers location map for specific detection. (D) Specific PCR assay for insertion site of n12. (E) Expression level of genes located at the T-DNA insertion sites.

By whole-genome resequencing, two transgenic lines (n12 and n153) and untransformed controls were compared with whole-genome insertions/deletions (indels). Gene ontology analysis and protein–protein interaction networks suggest that genes affected by different indels may be associated with programmed cell death and defense responses (Zhang *et al.*, 2018b). We found T-DNA of transgenic poplar line n12 was inserted into the second intron of Potri.015G076600 at 10162773 bp on Chr15 and the nucleotide composition was 65% for AT content by high-efficiency thermal asymmetric interlaced PCR (Fig. 4.2B) (Zhang and Hu, 2020).

Further analysis of the specificity of insertion events is of great significance. According to the determined integration sites of transgenic poplar T-DNA, forward primers and reverse primers RB-F/RB-R were designed according to the genome on both sides of the line n12 integration site (Fig. 4.2C). Using non-transgenic poplar and line n222 as control, the transgenic line n12 was detected by PCR with three primers: RB-F, RB-2A, and RB-R. Electrophoresis showed that 709 bp (transgenic) and 1159 bp (non-transgenic) fragments were amplified in line n12, while only 1159 bp (non-transgenic) fragments were amplified in control and line n222 (Fig. 4.2D).

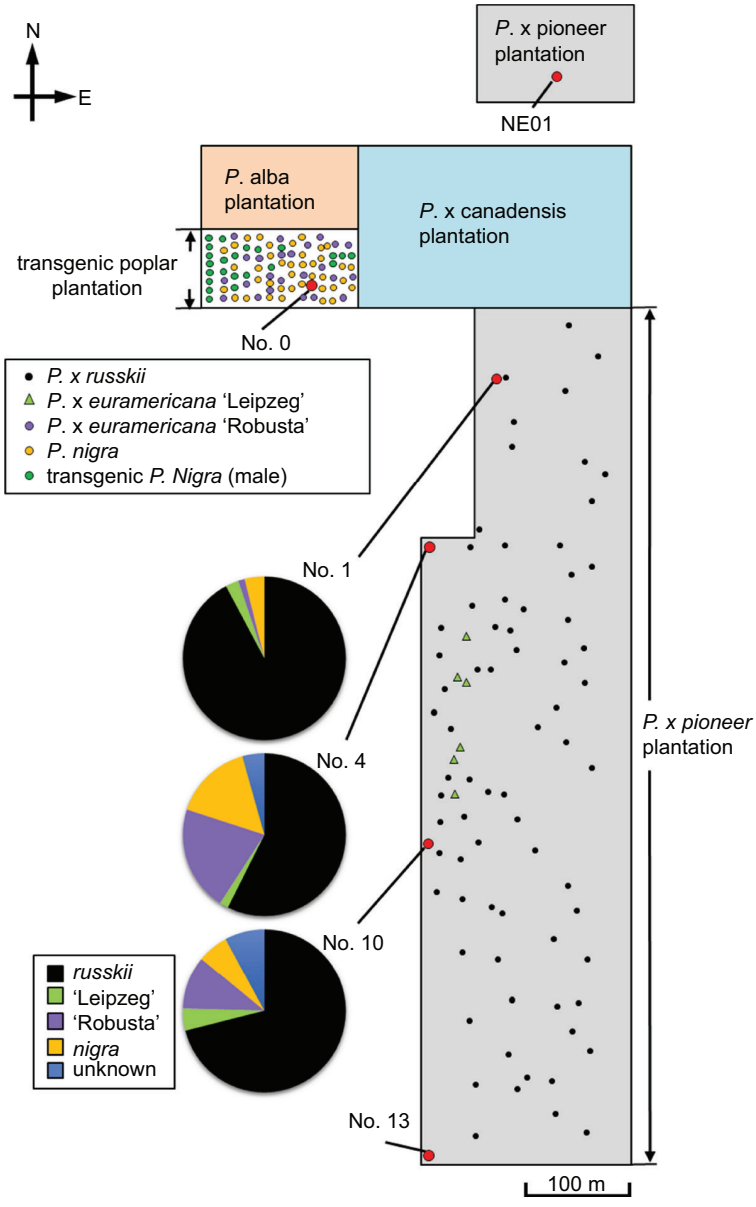


Fig. 4.3. Map of the sites of transgenic poplar plantation (TPP) and the surrounding non-transgenic plantations (Hu *et al*., 2017). No. 0, No. 4, No. 10, and No. 13: sites for seeds collection. The pie chart shows the proportion of candidate paternity of collected seeds at three sites.

Previous studies have shown that the insertion of T-DNA may affect the expression of endogenous genes around the insertion site. Potri.015G076600 encodes serine protein kinase (SPK), and 2.7 kilobase (kb) downstream of the insertion site is Potri.015G076700 ataxia telangiectasia mutated family protein (ATM). The expression of *SPK* in line n12 leaves was 4.3 times higher than that in CK, while *ATM* expression was 20 times lower than that in CK (Fig. 4.2E).

Pb29 is a transgenic high-resistance poplar line 741 with *BtCry1Ac* that has been

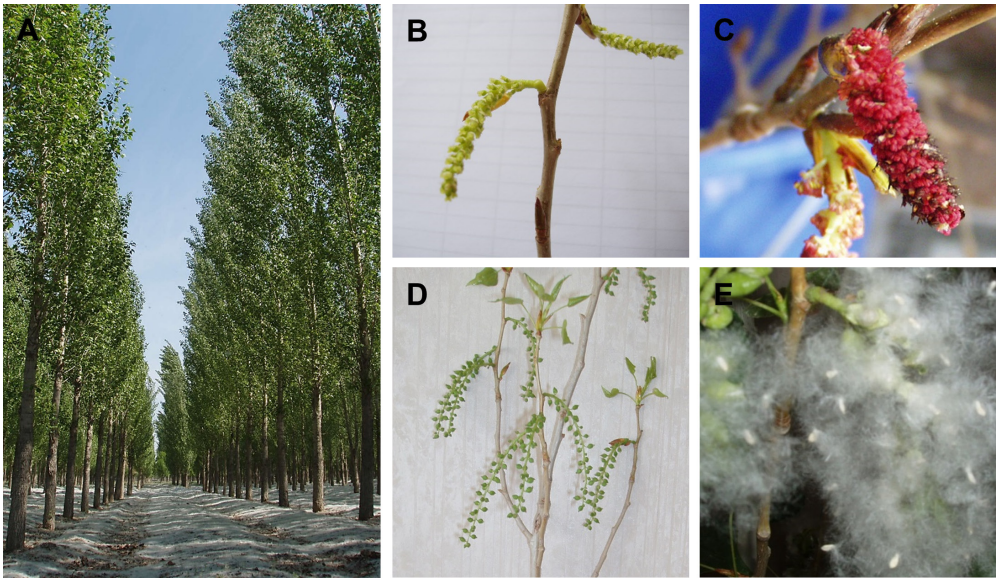


Fig. 4.4. Female, male flowers, hybrid fruits and seeds of *P. nigra* cv. 'Pioneer' and *Bt P. nigra* (Hu *et al*., 2017). (A) 'Pioneer' plantation. (B) Female flowers of 'Pioneer' tree. (C) Male flowers of transgenic *P. nigra* (n222). (D) Fruits of 'Pioneer' pollinated by male *Bt* poplar. (E) Seeds of 'Pioneer'. (Photographs by Jianjun Hu.)

commercialized. Chen *et al*. (2021) identified two T-DNA insertion sites, located at 9,283,905–9,283,937 bp on Chr03 and 10,868,777–10,868,803 bp on Chr10, using next-generation and nanopore sequencing. The T-DNA insertion did not change the gene expression near the insertion site in the leaves of Pb29.

4.2.1.2 Identification of Bt transgene by PCR amplification of Bt fragments from transgenic leaves, pollens, and hybrid seeds

We used the mature *Bt* poplar plantation in Manas County Plain Forest Station (N44°15′56″, E88°19′60″), Xinjiang Uygur Autonomous Region, which was planted with 2-year cuttings on about 1 ha in 1994, to assess the possible gene flow from transgenic poplar plantation through pollen and seed (Fig. 4.3) (Hu *et al*., 2017). The adjacent poplar plantations included *P. alba* plantation in the north, *P.* × *canadensis* (female) in the east and *P. nigra* 'Pioneer' (*P. nigra* 'Italica' × *P. nigra*, female) in the southeast and northeast within 150–800 m of the transgenic poplar. During these studies the following questions were addressed:

1. How far can the *Bt* poplar pollen travel and successfully fertilize other local poplar trees?
2. What is the probability that *Bt* poplar seeds will germinate and establish under different conditions?

The answers to these questions would provide the basis for accessing the safety issue of the application of transgenic trees.

Seeds collected from non-transgenic CK3 (*P. nigra*) in the transgenic poplar plantation (TPP) (Fig. 4.3) and *P. nigra* 'Pioneer' trees (Fig. 4.3 and 4.4A) at four sites (No. 4, No. 10, No. 13 and NE01, Fig. 4.3) were used for DNA extraction and amplification of *Bt* fragment with gene-specific primers. The high cross-compatibility between 'Pioneer' and *Bt* poplars was revealed by pollinating female flowers (Fig. 4.4B) with pollen from transgenic male flowers (Fig. 4.4C) to produce fruits (Fig. 4.4D) and seeds (Fig. 4.4E). The PCR products of *Bt* fragments amplified from leaves of female 'Pioneer' trees, the leaves and pollen of transgenic male tree and the leaves of their progenies are shown in Fig. 4.5A. The successful amplification of the *Bt* fragments from transgenic clones and their progenies indicated

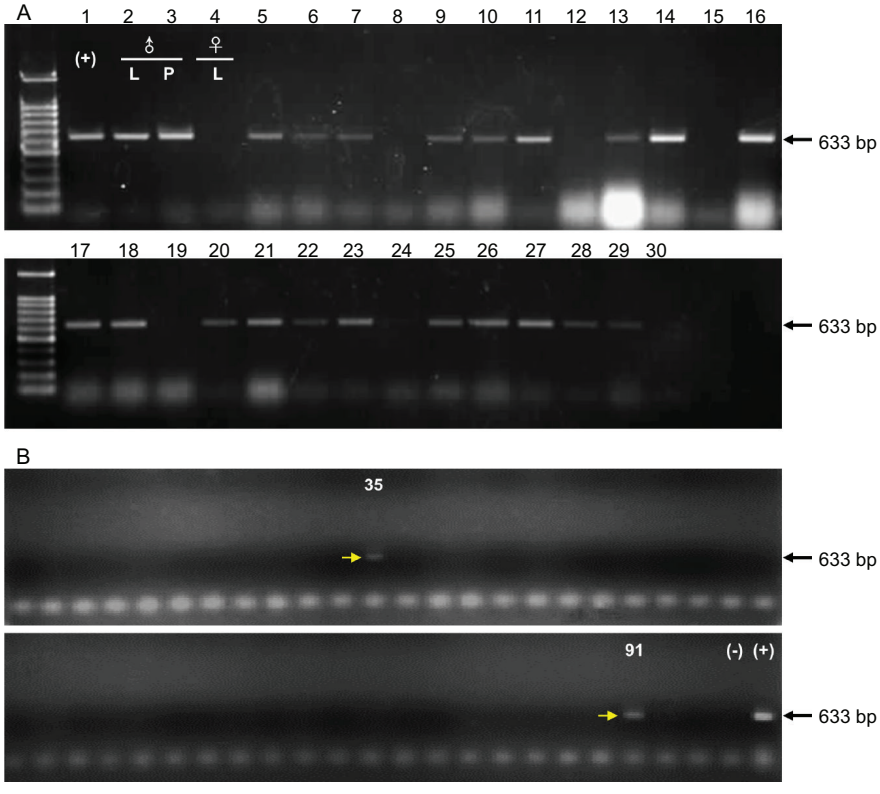


Fig. 4.5. Agarose gel electrophoresis of products amplified from non-transgenic, transgenic poplars (A) and partial seeds collected at No. 4 site (B) (Hu *et al*., 2017). Arrows point to the amplified Bt fragments; (+): positive control; (–): negative control; (L): leaves of 'Pioneer' and n222; (P): pollen of n222.

that the *Bt* gene still existed in the genome of transgenic poplars and could be transferred to its offspring by crossing. The amplified products of seeds collected from the 'Pioneer' trees at the four sites were also performed to measure the probability of the transgene flow and some of them at site No. 4 are shown in Fig. 4.5B (Hu *et al*., 2017).

4.2.1.3 Bt seeds produced with pollen from Bt poplar plantation

The number of *Bt*-containing seeds detected was related not only to the distance from but also to the direction to the TPP. Hu *et al*. (2017) collected 3045 to 6190 seeds of each female tree from four sites (No. 4, No. 10, No. 13 and NE01, Fig. 4.3) for two consecutive years to amplify *Bt* fragments (Fig. 4.5B). *Bt* seeds occurred at the rate of 0.16% and 0.15% (in the two years, respectively) collected at the site No. 0 in the transgenic poplar plantation, but only 0.05% and 0.07% at site No. 4, and 0.03% and 0.02% at site No. 10 (Table 4.1). Although more collected seeds were tested, no *Bt* gene was detected in the seeds at sites No. 13 and NE01. No. 13 was located at 794 m southeast of TPP, while NE01 was located at 368 m northeast.

4.2.1.4 Germination ability and seedling survival of Bt poplar seeds under different conditions

The germination rate of transgenic poplar seeds under natural conditions is an important part of gene drift. The germination rate of *Bt* seeds decreased to 7% and 0%, respectively, after storage at room temperature and in the field for 3 weeks,

Table 4.1. Summary of the proportion of seeds with *Bt* from sampled sites. (Adapted from Hu *et al*. (2017))

Sites (distance to transgenic poplar plantation)	Samples 1^{st} Y /2^{nd} Y	Seedlings /sample	Number of seeds tested 1^{st} Y/2^{nd} Y	*Bt* seed (%) 1^{st} Y /2^{nd} Y
No. 0 (0 m)	609/1238	2 or 5/5	3045/6190	0.16/0.15
No. 4 (210 m)	1236/871	5/5	6180/4355	0.05/0.07
No. 10 (500 m)	1211/863	5/5	6055/4315	0.03/0.02
No. 13 (794 m)	1212/875	5/5	6060/4375	0/0
NE01 (368 m)	1208/947	5/5	6040/4735	0/0

while it was still 48% after 4 weeks of cold storage. When the *Bt* seeds were sown in the field, no seedlings were found. However, the seedling rates of the plots with soil preparation, watering and plastic film mulching were 0.3% and 11.3%, respectively, after 1 week but decreased to 0.2–3.5% after 2 months (Hu *et al.*, 2017). It was determined that poplar seeds remain viable for only 2 weeks in nature (Imbert and Lefèvre, 2003).

4.2.2 Adaptation of hybrid offspring of transgenic and non-transgenic poplar

Hybrid breeding can promote gene exchange and obtain a hybrid with the excellent characteristics of the parents. By controlling pollination, exogenous genes can be transferred to near-origin species such as cultivated and wild species.

The inheritance and expression of the exogenous Bt gene/protein was studied in transgenic hybrid poplar progeny lines. The fast-growing improved variety *P. deltoides* 'DanHong' and transgenic *P. nigra* with *Btcry1Ac* gene were used as hybrid parents to produce hybrid progeny by artificially controlled pollination. Seventeen hybrid progenies verified by PCR testing were performed with insect (*Lymantria dispar*) bioassay and field growth measurement for 4 years. The result of the insect bioassay indicated that the insect resistance of the hybrid progeny was obviously improved compared with *P. deltoides* 'Danhong' and non-transformed *P. nigra*. The height and ground diameter of hybrid progenies B3-44, B3-102, B3-132 and B3-153 were relatively high (Jia *et al.*, 2017). Ren *et al.* (2017) found that transgenic poplar line 741 produced very few seeds. Hybridization was implemented using non-transgenic poplar line 84K (*P. alba* × *P. glandulosa*) as the male parent trees, and transgenic poplar 741 lines, which had different insect-resistance (high insect resistance (pB11, pB29), moderate insect resistance (pB1, pB17), and no insect resistance (pB6)), were used as the female parent trees. The hybrid progeny of pB6 × 84K had a segregation ratio of 3:1. The insect resistance of the hybrid progeny plants was nearly identical to that of the parent plants (Ren *et al.*, 2017). The hybridization of transgenic and non-transgenic poplar is a new way to produce new germplasm, but at the same time we should monitor the process to help avoid transgene escape.

Crossing success between *P. tomentosa* and other poplar trees is very low, except for some *P. tomentosa* × *P. bolleana* and *P. alba* × *P. glandulosa* lines (Kang, 2001; Guo *et al.*, 2018). Triphenyl tetrazolium chloride (TTC) analysis of pollen showed that no pollen of 'T-46' (transgenic *P. tomentosa* with *AhDREB1*) or '401' (non-transgenic triploid hybrid poplar (*P. tomentosa* × *P. bolleana*) × *P. tomentosa*)) was stained and the *in vitro* pollen germination test showed that no pollen of 'T-46' or '401' germinated. Likewise, pollen of hybrid combinations of 'T-46' × 'TB04' and '401' × 'TB04' failed to germinate after 6 h, and the seed setting rate was 0%: no seed was obtained (Guo *et al.*, 2018). Cross-incompatibility between *P. tomentosa* and its close relatives will reduce the risk of gene escape to the non-transgenic population.

4.3 Horizontal Gene Transfer of Transgenic Poplar

Horizontal gene transfer (HGT) refers to the direct transfer of genetic material between different species. Exogenous genes are transferred to soil ecosystem through roots, branches and leaves and transferred between scion and rootstock in graft plants.

4.3.1 Soil ecosystem

The soil ecosystem and surrounding weeds are important to study for horizontal transfer risks. The DNA released into the soil by decaying leaves and roots from transgenic trees may become available for incorporation by soil microbes. Fragments of transgenic DNA were found not to be detectable in the field for more than 4 months (Hay *et al.*, 2002). In soil with degrading leaves under natural conditions, the exogenous gene fragment of *AhDREB1* could be amplified in the first 2 months by PCR, but could not be amplified after the third month (Guo *et al.*, 2018). There were no significant differences in the temporal and seasonal dynamics of species and individuals in arthropod communities between the non-transgenic poplar–cotton and transgenic poplar-741–cotton composite systems (Jiang *et al.*, 2018). The transgenic hybrid poplar *P. euramericana* 'Guariento' with five genes containing *SacB*, *JERF36*, *vgb*, *BtCry3A* and *OC–I* could respond to insect stressors, salinity, and drought both in the greenhouse and in the field (Su *et al.*, 2011). The genomic DNA of transgenic poplar rhizosphere soil microorganisms, insects and surrounding weeds were used as templates, and PCR products of the target genes were not detected in all the samples, which indicated that there was no exogenous gene transfer to other biological levels in the transgenic hybrid poplar (Hou, 2008). The metabolomic soil analysis of two 8-year-old transgenic lines (D520 and D521) and one non-GM line (D50, *P. euramericana* 'Guariento') showed few statistical differences in the bacterial diversity and community structure between transgenic and non-transgenic poplars (Zhu *et al.*, 2016; Ning *et al.*, 2018). The root and soil samples of *JERF36*-OE and non-transgenics in Daqing with and without salinity suggested that transgenic events did not affect the endophytic bacterial and fungal diversity of poplar trees, while the pH and the soil organic matter content influenced the bacterial and fungal community structure (Ning *et al.*, 2018). 'Nanlin 895' with *AtSnRK2C* improved salt and drought resistance in the field. The exogenous gene was stable for expression *in vivo*, while the soil around transgenic lines did not yield PCR products of marker genes and the *AtSnRK2C* target fragment and had no significant effect on lettuce seed growth, suggesting that the exogenous genes were not HGT (Ma *et al.*, 2019). All the results indicated that the planting of transgenic poplar did not result in HGT to the soil ecosystem.

4.3.2 Graft of transgenic and non-transgenic poplar

Grafting has been widely used as an important means of asexual reproduction to maintain varieties' features, and to improve yield and quality. Grafting can cause HGT between scion and rootstock, RNA long-distance transportation, and epigenetic regulation (Zhang *et al.*, 2018a). It is interesting to study whether grafting can cause gene exchange between transgenic and non-transgenic plants. Analysis of the graft sites of two transgenic tobaccos (*Nicotiana tabacum*) which carried different reporter genes in different cellular compartments, the nucleus and the plastic, revealed the frequent occurrence of cells harboring both antibiotic resistances and both fluorescent reporters. The results demonstrated that plant grafting can result in the exchange of genetic information via either large DNA pieces or entire plastid genomes (Stegemann and Bock, 2009). The graft of two transgenic tobacco plants carrying different selectable marker genes caused HGT and produced new species with 60, 76, 84, and 96 chromosomes (Fuentes *et al.*, 2014). Pb29 (transgenic poplar 741 with *BtCry1Ac* genes), CC71 (transgenic poplar 741 with *BtCry3A* genes) and non-transgenic poplar 741 were either as scion or as rootstock (Wang and Yang, 2010; Wang *et al.*, 2012; Chen *et al.*, 2016). No mRNA of the *Bt* gene was detected in the branch and leaf of 741, no matter whether its material was used as scion

or stock, which suggested mRNA of *Bt* was not transported between the stock and scion. Bt protein existed in phloem, xylem, pith and leaves of the grafted poplar and was mainly transported between rootstock and scion through phloem (Wang and Yang, 2010; Chen *et al.*, 2016). This shows that the grafting of transgenic and non-transgenic poplar is safe, which can promote the insect resistance of the non-transgenic grafting part, but will not cause gene escape in transgenic poplar (Wang *et al.*, 2008, 2012).

4.4 Discussion and Outlook

4.4.1 Effects of pollen and seed on gene flow

Gene flow is an important process of biological evolution and a common natural process that affects biodiversity, as DNA is exchanged among organisms. The concern about transgene escape via pollen and seeds of transgenic trees is the possible negative effect on natural forest ecology, which may play a role in hampering widespread use of transgenic trees (Häggman *et al.*, 2013). It is thus important to assess pollen dispersal, seed production, and seedling survival of transgenic trees and their progeny. Previous studies have used established stands and simulation models to explore the consequences of introducing new genes into the environment. In this chapter, we have taken advantage of the commercialized insect-resistant transgenic poplar (Ewald *et al.*, 2006) and for the first time we have provided empirical data to evaluate the possibility of gene flow through pollen and seeds.

Without commercialized transgenic tree plantations, it is not practical to assess transgene escape to conventional populations. The alternative would be to develop and exploit simulation models such as STEVE (Simulation of Transgene Effects in a Variable Environment) and AMELIE (A Modeling framEwork for popuLatIon-gEnotype dynamics) to gain insight into the gene flow in forest trees. Several models of pollen and seed dispersal have been proposed for forest trees (DiFazio, 2002; DiFazio *et al.*, 2004; Kuparinen and Schurr, 2007; Nathan *et al.*, 2008; DiFazio *et al.*, 2012). Pollen propagation distance is affected by environmental factors such as temperature and wind. For example, Kuparinen and Schurr (2007) investigated the rate of transgenic escape in cases where the modified organism carries mitigation genes, and DiFazio (2002) used his model to predict transgenic escape assuming long-distance dispersal as a common phenomenon. Many results indicate that long-distance dispersal is extensive and pollen dispersal curves are similar in populations with very different ecological and demographic characteristics (Slavov *et al.*, 2009). DiFazio *et al.* (2004) reported the mean pollination distances for *P. trichocarpa* as between 140 m and 1100 m, with a strong dependency on the area sampled. Pospíšková and Šálková (2006) found that the effective pollination distances were 10–230 m within a *P. nigra* population along the Morava River. Rathmacher *et al.* (2010) showed that only a minor part of gene flow took place at distances beyond 1 km, and poplar seeds generally had shorter dispersal distances with a maximum distance of 500 m. In central Germany, Bialozyt (2012) reported that most effective pollinations (75%) occurred within a distance of < 1000 m between native black poplar trees (*P. nigra*) and its commercial hybrid (*P.* × *canadensis*), and only a very limited proportion of effective pollinations occurred at distances > 2 km. Therefore, there was a wide range in the distance of pollen dispersal accessed in poplar plantations based on population studies and models. In our study, the results indicated that the pollen dispersal of *Bt* poplars occurred within a distance of 500 m, which is close to the predicted minimum limit based on models or population analysis. Dispersal of *Bt* pollen is also affected by the wind direction during the flowering period in spring. Due to the prevailing wind direction from northwest to southeast in the study site, it is reasonable that no *Bt* pollen-pollinated seeds were detected in the northeast site (NE01), though the separation of about 200 m of poplar plantation in width between TPP and NE01 sites hampered *Bt* pollen dispersal.

The formation of a hybrid plant does not mean that a wild population will be established (Chandler and Dunwell, 2008). Hypocotyls develop within 6–8 h after moisture has reached the seed and the pappus has degraded (Zsuffa, 1974), and the seedling dies if conditions are not favorable for further development. Germination occurs exclusively on bare soil (Barsoum and

Hughes, 1998). The results presented in this study reveal that the transgenic poplar seeds lost germination ability under field conditions after 3 weeks, but retained the ability almost unaffected at 4°C in long storage. Therefore, there is a crucial period of 3 weeks for germination of transgenic seeds in the field. When sown at test sites without watering, no transgenic poplar seeds were germinated in the bare field, but a seedling survival rate of 3.5% was obtained with watering, weeding, and tillage measures. According to Guilloy-Froget *et al.* (2002), successful germination of *P. nigra* seeds depends on a change to hydrated conditions. Therefore, soil, water, and weeds could all affect the germination and survival of seeds, and water seems to be the most important limiting factor. The significance of seed dispersal could also be highly dependent on plantation size (DiFazio *et al.*, 2012). In the latter study, the number of male transgenic poplar trees (23) was small, which might have affected the pollen fertilization of these trees under competition with non-transgenic pollen from 141 males. As a result, only 0.16% and 0.15%, respectively, of *Bt* seeds from the control (CK3) were detected in the transgenic poplar plantation in the two years. The paternity analysis also supports this supposition that the non-transgenic male trees dominating the pollen pool would produce more non-*Bt* seeds, thus few transgenic seeds would be readily produced in such a situation.

In summary, the study provides evidence that the pollen from transgenic trees travels only limited distances, and that, in the presence of a large number of non-transgenic males, the probability for the *Bt* pollen to successfully produce *Bt* progeny is quite rare. In addition, the transgenic seeds could not germinate easily and the seedlings could not survive easily in dry areas in northern China. Therefore, based on this empirical data, it could be concluded that the transgene may not flow easily to the local poplar forestry.

4.4.2 Evaluation of hybrid fitness

The analysis and evaluation of the fitness effect brought by transgenic to wild sibling species is very important for the evaluation of gene escape and ERA (Lu, 2015). The extensive interspecific gene flow of 80 individuals in six sections of the genus *Populus* showed that frequent hybridization occurred among poplar species within the same section or clade (Wang *et al.*, 2020). We should strengthen our work in the following aspects.

1. Hybrid combinations are designed to investigate the cross-compatibility between different species. Transgenic and non-transgenic hybrids were used for backcrossing experiments to evaluate the suitability of backcrossing offspring.
2. Surveys of transgenic poplars planted in a variety of fields are necessary. The analysis of different transgenic events should follow case-by-case and region-by-region.
3. Improved methods for monitoring wild relatives and non-target hosts near transgenic forests should be examined.

4.4.3 A new way to avoid gene escape

It is believed that transgenic escape through pollen is the most common mode of gene flow. The assumption is that transgenic crops hybridize with their sexually compatible wild relatives through pollen transmission, so escapes into the natural environment will occur. To reduce pollen transmission, male and female sterile mutants may be utilized, as well as regulating flowering and pollen in transgenic trees. For example, PrMC2pro-II-barnaseH102E, which was the modified coding sequences barnaseH102E driven by male cone-specific promoter of *Pinus radiata PrMC2* transgenic pine and *Eucalyptus*, grew similarly to control trees in all observed attributes except the pollen-less phenotype. Meanwhile, field experiments showed that the gene could be used for pollination-mediated gene flow caused by large-scale deployment of transgenic trees (Zhang *et al.*, 2012).

CRISPR-mediated gene editing may be used to selectively knock out key genes in plants (Li *et al.*, 2013). For some crops, there is no transgene integration in the plant, and therefore this may be useful as an unregulated plant breeding approach. For poplars, gene editing may be performed to alter traits. For example,

mutating *NST/SND*, *PtrMYB156*, *PtoMYB170* and *PtrMYB57* resulted in knockouts to study xylem development and anthocyanin biosynthesis (Wan *et al*., 2017; Xu *et al*., 2017; Yang *et al*., 2017; Takata *et al*., 2019).

4.4.4 Government regulations

Safety control and management measures should be strengthened to set laws and regulations based on research experience of gene drift of transgenic crops and the progress of research on gene drift between transgenic poplar and non-transgenic trees. The EU strongly advocates a prevention-oriented approach and a product and process management model (Sprink *et al*., 2016). There are many disputes about risk management of transgenic crops in Brazil, but Brazil has signed a new biosafety law to regulate risks (Silva, 2018). In 2017, China promulgated new measures for genetically modified safety management. Although research of transgenic trees relative to crops is still in its infancy, the Chinese government continuously strengthened the supervision of transgenic trees and implemented new measures for the approval and management of genetic engineering activities in 2018.

Acknowledgments

This work was supported by the National Key Program on Transgenic Research (2018ZX08020002) and National High-Tech R&D Program of China (2011AA10020101).

References

Ahuja, M. (2009) Transgene stability and dispersal in forest trees. *Trees* 23, 1125–1135.

Ahuja, M. (2010) Fate of transgenes in the forest tree genome. *Tree Genetics & Genomes* 7, 221–230.

Barsoum, N. and Hughes, F.M.R. (1998) Regeneration response of black poplar to changing river levels. In: Wheather, H. and Kirby, C. (eds) *Hydrology in a Changing Environment*. John Wiley and Sons, Hoboken, New Jersey, pp. 347–412.

Bialozyt, R. (2012) Gene flow in poplar – Experiments, analysis and modeling to prevent transgene outcrossing. *iForest – Biogeosciences and Forestry* 5, 147–152.

Chandler, S. and Dunwell, J. (2008) Gene flow, risk assessment and the environmental release of transgenic plants. *Critical Reviews in Plant Sciences* 27, 25–49.

Chen, L.-J., Diao, Z.-Y., Specht, C. and Sung, Z.R. (2009) Molecular evolution of VEF-Domain-Containing PcG genes in plants. *Molecular Plant* 2(4), 738–754. DOI: 10.1093/mp/ssp032.

Chen, P.-F., Ren, Y.-C., Zhang, J., Wang, J.-M. and Yang, M. (2016) Expression and transportation of Bt toxic protein in 8-year-old grafted transgenic poplar. *Scientia Silvae Sinicae* 52, 46–52.

Chen, X., Dong, Y., Huang, Y., Fan, J., Yang, M. *et al*. (2021) Whole-Genome resequencing using next-generation and nanopore sequencing for molecular characterization of T-DNA integration in transgenic poplar 741. *BMC genomics* 22(1), 329. DOI: 10.1186/s12864-021-07625-y.

DiFazio, S. (2002) Measuring and modeling gene flow from hybrid poplar plantations: Implications for transgenic risk assessment. PhD, Oregon State University, Oregon.

DiFazio, S., Slavov, G., Burczyk, J., Leonardi, S. and Strauss, S. (2004) Gene flow from tree plantations and implications for transgenic risk assessment. In: Walter, C. and Carson, M.J. (eds) *Plantation Forest Biotechnology for the 21st Century*. Research Signpost, Kerala, India, pp. 405–422.

DiFazio, S.P., Leonardi, S., Slavov, G.T., Garman, S.L., Adams, W.T. *et al*. (2012) Gene flow and simulation of transgene dispersal from hybrid poplar plantations. *New Phytologist* 193(4), 903–915. DOI: 10.1111/j.1469-8137.2011.04012.x.

Eckenwalder, J.E. (1996) Systematics and evolution of *Populus*. In: Stettler, R.F. (ed.) *Biology of Populus and Its Implications for Management and Conservation*. NRC Research Press, Ottawa, Canada, pp. 7–32.

Ewald, D., Hu, J. and Yang, M. (2006) Transgenic forest trees in China. In: Fladung, M. and Ewald, D. (eds) *Tree Transgenesis: Recent Developments*. Springer, Berlin, pp. 25–45.

Fang, S. (2008) Silviculture of poplar plantation in China: a review. *Journal of Applied Ecology* 19(10), 2308–2316.

Fuentes, I., Stegemann, S., Golczyk, H., Karcher, D. and Bock, R. (2014) Horizontal genome transfer as an asexual path to the formation of new species. *Nature* 511(7508), 232–235. DOI: 10.1038/nature13291.

Guilloy-Froget, H., Muller, E., Barsoum, N. and Hughes, F. (2002) Dispersal, germination, and survival of *Populus nigra* L. (Salicaceae) in changing hydrologic conditions. *Wetlands* 22, 478–488.

Guo, Q., Lu, N., Luo, Z., Sun, Y. and Jin, S. (2018) An assessment of the environmental impacts of transgenic triploid *Populus tomentosa* in field condition. *Forests* 9, 482–494.

Häggman, H., Raybould, A., Borém, A., Fox, T., Handley, L. *et al*. (2013) Genetically engineered trees for plantation forests: key considerations for environmental risk assessment. *Plant Biotechnology Journal* 11(7), 785–798. DOI: 10.1111/pbi.12100.

Hall, L., Topinka, K., Huffman, J., Davis, L. and Good, A. (2009) Pollen flow between herbicide-resistant *Brassica napus* is the cause of multiple-resistant *B. napus* volunteers. *Weed Science* 48, 688–694.

Hay, I., Morency, M.-J. and Seguin, A. (2002) Assessing the persistence of DNA in decomposing leaves of genetically modified poplar trees. *Canadian Journal of Forest Research* 32, 977–982.

Hou, Y. (2008) Preliminary study on the ecological safety assessment of transgenic poplar. PhD thesis. Chinese Academy of Forestry, Beijing.

Hu, J.-J., Tian, Y.-C., Han, Y.-F., Li, L. and Zhang, B.-E. (2001) Field evaluation of insect-resistant transgenic *Populus nigra* trees. *Euphytica* 121, 123–127.

Hu, J.J., Yang, M.-S. and Lu, M.Z. (2010) Advances in biosafety studies on transgenic insect-resistant poplars in China. *Biodiversity Science* 18, 336–345. [In Chinese with English abstract.]

Hu, J.-J., Wang, L.J., Yan, D.-H. and Lu, M.-Z. (2014) Research and application of transgenic poplar in China. In: Fenning, T. (ed.) *Challenges and Opportunities for the World's Forests in the 21st Century*. Springer, Dordrecht, Netherlands, pp. 567–584.

Hu, J.-J., Zhang, J., Chen, X.-L., Lv, J.-H., Jia, H.-X. *et al*. (2017) An empirical assessment of transgene flow from a Bt transgenic poplar plantation. *PLoS ONE* 12(1), e0170201. DOI: 10.1371/journal.pone.0170201.

Imbert, E. and Lefèvre, F. (2003) Dispersal and gene flow of *Populus nigra* (Salicaceae) along a dynamic river system. *Journal of Ecology* 91, 447–456.

ISAAA (2016) *Global Status of Commercialised Biotech/GM Crops: 2016*. ISAAA Brief No. 52. International Service for the Acquisition of Agri-biotech Applications, Ithaca, New York.

ISAAA (2019) *Global Status of Commercialized Biotech/GM Crops: 2019*. ISAAA Brief No. 55. International Service for the Acquisition of Agri-biotech Applications, Ithaca, New York.

Jia, C., Zhao, H., Wang, H., Xing, Z. and Du, K. (2004) Obtaining the transgenic poplars with low lignin content through down-regulation of 4CL. *Chinese Science Bulletin* 49, 905–909.

Jia, H., Sun, P., Li, J., Li, X. and Su, X. (2017) Insect-resistance and growth measurement of hybrid progeny from *Populus deltoids* CL 'Danhong' and transgenic *P. nigra* with BtCry1Ac gene. *Molecular Plant Breeding* 15, 4101–4109.

Jiang, W., Zhang, D., Liu, J., Li, C. and Lu, Z. (2018) Comparative analysis of arthropod communities in transgenic Bt and non-transgenic poplar-cotton composite systems. *Scientia Silvae Sinicae* 54, 73–79.

Kang, X. (2001) Study on mechanism of pollen abortion in Chinese white poplar (*Populus tomentosa* Carr.). *Scientia Silvae Sinicae* 37, 35–39.

Khasa, D., Pollefeys, P., Navarro Quezada, A., Perinet, P. and Bousquet, J. (2005) Species-specific microsatellite markers to monitor gene flow between exotic poplars and their natural relatives in eastern North America. *Molecular Ecology Notes* 5, 920–923.

Kuparinen, A. and Schurr, F. (2007) A flexible modelling framework linking the spatio-temporal dynamics of plant genotypes and populations: application to gene flow from transgenic forests. *Ecological Modelling* 202, 476–486.

Li, C., Wang, X., Lu, W., Liu, R. and Tian, Q. (2014) A poplar R2R3-MYB transcription factor, PtrMYB152, is involved in regulation of lignin biosynthesis during secondary cell wall formation. *Plant Cell, Tissue and Organ Culture* 119, 553–563.

Li, C., Ma, X., Yu, H., Fu, Y. and Luo, K. (2018) Ectopic expression of PtoMYB74 in poplar and Arabidopsis promotes Secondary Cell wall formation. *Frontiers in Plant Science* 9, 1262. DOI: 10.3389/fpls.2018.01262.

Li, J., Zhang, Y., Chen, K., Shan, Q. and Wang, Y. (2013) CRISPR/Cas: a novel way of RNA-guided genome editing. *Hereditas* 35(11), 1265–1273. DOI: 10.3724/sp.j.1005.2013.01265.

Li, Y., Su, X., Zhang, B., Huang, Q., Zhang, X. *et al*. (2009) Expression of jasmonic ethylene responsive factor gene in transgenic poplar tree leads to increased salt tolerance. *Tree Physiology* 29(2), 273–279. DOI: 10.1093/treephys/tpn025.

Lu, B. (2015) Analysis of fitness effect and its application in assessing environmental risk caused by transgene flow. *Biotechnology Bulletin* 31, 7–16.

Ma, X., Sun, W., Wei, H., Liu, L. and Yu, X. (2019) Field trials and biosafety analysis of transgenic poplar with AtSnRK2C gene. *Molecular Plant Breeding* 17, 3570–3578.

Nathan, R., Schurr, F.M., Spiegel, O., Steinitz, O., Trakhtenbrot, A. *et al*. (2008) Mechanisms of long-distance seed dispersal. *Trends in Ecology & Evolution* 23(11), 638–647. DOI: 10.1016/j.tree.2008.08.003.

Niggemann, M., Wiegand, T., Robledo-Arnuncio, J. and Bialozyt, R. (2012) Marked point pattern analysis on genetic paternity data for uncertainty assessment of pollen dispersal kernels. *Journal of Ecology* 100, 264–276.

Ning, K., Ding, C., Zhu, W., Zhang, W., Dong, Y. *et al*. (2018) Comparative metabolomic analysis of the cambium tissue of non-transgenic and multi-gene transgenic poplar (*Populus* × *euramericana* 'Guariento'). *Frontiers in plant science* 9, 1201. DOI: 10.3389/fpls.2018.01201.

Pandolfo, C.E., Presotto, A., Carbonell, F.T., Ureta, S., Poverene, M. *et al*. (2018) Transgene escape and persistence in an agroecosystem: the case of glyphosate-resistant *Brassica rapa* L. in central Argentina. *Environmental Science and Pollution Research International* 25(7), 6251–6264. DOI: 10.1007/s11356-017-0726-3.

Pospíšková, M. and Šálková, I. (2006) Population structure and parentage analysis of black poplar along the Morava River. *Canadian Journal of Forest Research* 36, 1067–1076.

Rathmacher, G., Niggemann, M., Köhnen, M., Ziegenhagen, B. and Bialozyt, R. (2010) Short-distance gene flow in *Populus nigra* L. accounts for small-scale spatial genetic structures: implications for in situ conservation measures. *Conservation Genetics* 11, 1327–1338.

Ren, Y., Zhang, J., Liang, H., Wang, J. and Yang, M. (2017) Inheritance and expression stability of exogenous genes in insect-resistant transgenic poplar. *Plant Cell, Tissue and Organ Culture* 130, 567–576.

Silva, J.F. (2018) *Brazil Agricultural Biotechnology Annual*. GAIN (Global Agricultural Information Network) Report, USDA Foreign Agricultural Service, Washington, DC.

Slavov, G.T., Leonardi, S., Burczyk, J., Adams, W.T., Strauss, S.H. *et al*. (2009) Extensive pollen flow in two ecologically contrasting populations of *Populus trichocarpa*. *Molecular Ecology* 18(2), 357–373. DOI: 10.1111/j.1365-294X.2008.04016.x.

Sprink, T., Eriksson, D., Schiemann, J. and Hartung, F. (2016) Regulatory hurdles for genome editing: process- vs. product-based approaches in different regulatory contexts. *Plant Cell Reports* 35(7), 1493–1506. DOI: 10.1007/s00299-016-1990-2.

Stegemann, S. and Bock, R. (2009) Exchange of genetic material between cells in plant tissue grafts. *Science* 324(5927), 649–651. DOI: 10.1126/science.1170397.

Strauss, S.H., Brunner, A.M., Busov, V.B., Ma, C. and Meilan, R. (2004) Ten lessons from 15 years of transgenic Populus research. *Forestry* 77, 455–465.

Su, X., Chu, Y., Li, H., Hou, Y., Zhang, B. *et al*. (2011) Expression of multiple resistance genes enhances tolerance to environmental stressors in transgenic poplar (*Populus* × *euramericana* 'Guariento'). *PLoS ONE* 6(9), e24614. DOI: 10.1371/journal.pone.0024614.

Takata, N., Awano, T., Nakata, M.T., Sano, Y., Sakamoto, S. *et al*. (2019) Populus NST/SND orthologs are key regulators of secondary cell wall formation in wood fibers, phloem fibers and xylem ray parenchyma cells. *Tree Physiology* 39(4), 514–525. DOI: 10.1093/treephys/tpz004.

Tian, Q., Wang, X., Li, C., Lu, W., Yang, L. *et al*. (2013) Functional characterization of the poplar R2R3-MYB transcription factor PtoMYB216 involved in the regulation of lignin biosynthesis during wood formation. *PLoS ONE* 8(10), e76369. DOI: 10.1371/journal.pone.0076369.

Wan, S., Li, C., Ma, X. and Luo, K. (2017) PtrMYB57 contributes to the negative regulation of anthocyanin and proanthocyanidin biosynthesis in poplar. *Plant Cell Reports* 36(8), 1263–1276. DOI: 10.1007/s00299-017-2151-y.

Wang, L. and Yang, M. (2010) Transportation of the mRNA and protein of the foreign Bt gene in transgenic poplar. *Scientia Silvae Sinicae* 46, 49–54.

Wang, G., Castiglione, S., Chen, Y., Li, L. and Han, Y. (1996) Poplar (*Populus nigra* L.) plants transformed with a *Bacillus thuringiensis* toxin gene: insecticidal activity and genomic analysis. *Transgenic Research* 5, 289–301.

Wang, Y.P., Li, J., Yang, M.S. and Liang, H.Y. (2008) Insect resistance selectivity of transgenic hybrid poplar 741. *Scientia Silvae Sinicae* 44, 67–71.

Wang, L., Yang, M., Akinnagbe, A., Liang, H., Wang, J. *et al*. (2012) *Bacillus thuringiensis* protein transfer between rootstock and scion of grafted poplar. *Plant Biology* 14(5), 745–750. DOI: 10.1111/j.1438-8677.2011.00555.x.

Wang, G., Dong, Y., Liu, X., Yao, G., Yu, X. *et al*. (2018) The current status and development of insect-resistant genetically engineered poplar in China. *Frontiers in Plant Science* 9, 1408–1422. DOI: 10.3389/fpls.2018.01408.

Wang, M., Zhang, L., Zhang, Z., Li, M., Wang, D. *et al*. (2020) Phylogenomics of the genus *Populus* reveals extensive interspecific gene flow and balancing selection. *New Phytologist* 225(3), 1370–1382. DOI: 10.1111/nph.16215.

Wegier, A., Piñeyro-Nelson, A., Alarcón, J., Gálvez-Mariscal, A., Alvarez-Buylla, E. *et al*. (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. *Molecular Ecology* 20(19), 4182–4194. DOI: 10.1111/j.1365-294X.2011.05258.x.

Xu, C., Fu, X., Liu, R., Guo, L., Ran, L. *et al*. (2017) PtoMYB170 positively regulates lignin deposition during wood formation in poplar and confers drought tolerance in transgenic Arabidopsis. *Tree Physiology* 37(12), 1713–1726. DOI: 10.1093/treephys/tpx093.

Yang, L., Zhao, X., Ran, L., Li, C., Fan, D. *et al*. (2017) PtoMYB156 is involved in negative regulation of phenylpropanoid metabolism and secondary cell wall biosynthesis during wood formation in poplar. *Scientific Reports* 7, 41209. DOI: 10.1038/srep41209.

Yin, J., Sun, Z., Wang, Y., Li, D. and Liu, G. (2004) Field test of saline resistant transgenic *Populus xiaozhannica* cv. 'Balizhuangyang'. *Journal of Northeast Forest University*, 23–25.

Zhang, C., Norris-Caneda, K.H., Rottmann, W.H., Gulledge, J.E., Chang, S. *et al*. (2012) Control of pollen-mediated gene flow in transgenic trees. *Plant Physiology* 159(4), 1319–1334. DOI: 10.1104/pp.112.197228.

Zhang, G., Qi, L., Nie, X. and Guo, H. (2018a) Genetic information exchange and variation mechanism of grafted plants. *Journal of Yunnan Agricultural University (Natural Science)* 33, 947–957. [In Chinese with English abstract.]

Zhang, J., Zhan, T.-I., Jia, H., Wu, L. and Lu, M. (2018b) Whole-genome re-sequencing reveals molecular mechanisms of biomass changes in 11-year-old Bt transgenic poplar. *Trees* 32, 1609–1620.

Zhang, L. and Hu, J. (2020) T-DNA insertion loci analysis and specific detection method of transgenic *Populus nigra* lines with *BtCry1Ac*. *Scientia Silvae Sinicae* 56, 45–52.

Zhu, W., Chu, Y., Ding, C., Huang, Q. and Zhang, B. (2016) Assessing bacterial communities in the rhizosphere of 8-year-old genetically modified poplar (*Populus* spp.). *Journal of Forestry Research* 27, 939–947.

Zsuffa, L. (1974) The genetics of *Populus nigra* L. *Annales Forestales* 6, 29–53.

5 Monitoring Gene Flow from Genetically Modified Soybean to Cultivated Soybean and Wild Soybean in China

Lai-Pan Liu[1], Kun Xue[2] and Biao Liu[1]*

[1]Nanjing Institute of Environmental Sciences, Ministry of Ecology and Environment, Nanjing, China; [2]College of Life and Environmental Sciences, Minzu University of China, Beijing, China

Abstract

With the large-scale commercial planting of genetically modified (GM) crops in the world, the gene flow from GM crops to their wild relatives and its environmental risks have become a hot topic in the field of biosafety of GM organisms (GMOs). Wild soybean is one of the important plant genetic resources in China. China has not only imported a large amount of GM soybeans every year, but also started to carry out field experiments of GM soybeans with intellectual property rights; therefore, the gene flow of GM soybean to wild relatives and its influence on natural resources should be assessed before the commercial planting of GM soybean in China. In this chapter, the research progress of gene flow from GM soybean to cultivated soybean and wild soybean and the fitness of hybrid offspring are reviewed. This chapter reviews the current studies on gene flow from GM soybean and its consequences and also proposes further research topics.

Keywords: cultivated soybean; fitness; gene flow; GM soybeans; wild soybeans

5.1 Introduction

In 2019, the global planting area of genetically modified (GM) crops was 190.4 million hectares with a slight decline of 0.7% from 191.7 million hectares in 2018, in which the planting area of GM soybean was still the largest among crop species: 91.9 million hectares with a reduction of 4% from 95.9 million hectares in 2018 (ISAAA, 2019). Transgenes in GM soybeans were mainly those endowing herbicide resistances via the glyphosate resistant *epsps* (5-enol pyruvylshikimate-3-phosphate synthase) gene, insect resistance with *Bt* genes, and stacked insect and herbicide resistance with *epsps*+*Bt* genes (ISAAA, 2019). The commercial cultivation of GM crops has not only produced huge economic benefits, but also protected the environment and human health by reducing the use of chemical pesticides and herbicides (Pellegrino *et al.*, 2018). However, GM crops may also cause environmental and human health risks, among which foreign gene flow is one of the main environmental risks, especially if foreign genes escape from GM crops to wild

*Corresponding author: liubiao@nies.org

 Gene Flow: Monitoring, Modeling and Mitigation (eds Wei Wei and C. Neal Stewart, Jr)
DOI: 10.1079/9781789247480.0005

relatives of crops and their potential environmental risks through pollen-mediated gene flow (Dale, 1992; Prakash, 2001; Snow and Allison, 2002; Hancock, 2003; Lu and Yang, 2009).

Gene flow was originally a concept in population genetics, which refers to the phenomenon or process of the transfer of one or more genes from one Mendelian genetic population to another (Lu *et al.*, 2016). According to this definition, gene flow can occur between different groups of a species and between species with different genetic relationships. In fact, gene flow and genetic material exchanges between different varieties of the same crop, between crops and their wild relatives and weed types is a natural phenomenon that has existed in the evolution of plants (such as crops) for thousands of years (Ellstrand *et al.*, 1999; Lu *et al.*, 2009). Among the world's major crops, 12 species can cross with their wild relatives, and have an important impact on the evolution of wild relatives: gene flow from seven crops to their wild relatives leads to more aggressive weeds of their wild relatives, or the extinction of some rare wild relatives under certain circumstances (Ellstrand *et al.*, 1999; Ellstrand, 2001; Wang and Peng, 2003; Haygood *et al.*, 2003; Campbell *et al.*, 2006). The transgene transfer to wild related species through gene flow theoretically may not only affect the genetic integrity of wild related species, but also make wild related species rapidly expand their distribution space by changing their survival and reproductive ability. The direst purported outcome is the formation of uncontrollable weeds, resulting in unpredictable environmental risks (Dale, 1992; Crawley *et al.*, 1993; Stewart *et al.*, 2003; Lu *et al.*, 2009; Lu, 2015). With the gradual expansion of commercial cultivation and trade scale of GM crops, the phenomenon of foreign gene flow in a variety of GM crops has been confirmed. After a variety of GM herbicide-resistant oilseed rape (*Brassica napus* L.) was commercially planted in Canada, the gene flow between GM oilseed rape resistant to different herbicides led to the emergence of autogenous seedlings of oilseed rape containing glyphosate-resistant and imidazoline-resistant herbicides in Canadian GM oilseed rapeseed fields and their surrounding environment (Hall *et al.*, 2000; Beckie *et al.*, 2003). Moreover, there was gene flow between GM oilseed rape and wild turnip (*Brassica rapa*), which resulted in the herbicide resistance gene persisting in the natural population of wild turnip for several years (Warwick *et al.*, 2003, 2010). In Argentina, there are reports of gene flow from *epsps* GM oilseed rape (*B. napus* L.) to turnip (*B. rapa* L.), resulting in the existence and spread of glyphosate-resistant turnips expressing the *epsps* gene in farmland (Pandolfo *et al.*, 2018). Mexico is the world's center of the biodiversity origin of maize. With the large-scale cultivation of GM maize in the USA (especially near the border area of Mexico) and its export to Mexico, foreign genetic components of GM corn have been found in local maize varieties in Mexico (Quist and Chapela, 2001; Dalton, 2008; Piñeyro-Nelson *et al.*, 2010). *Bt* transgenic cotton (*Gossypium hirsutum*) was approved for commercial cultivation in Mexico in 1996, and a paper published in 2011 confirmed the detection of the *Bt* gene in wild upland cotton populations in Mexico (Wegier *et al.*, 2011). In China, foreign genes such as *Bt* and *epsps* can be transferred from GM rice to cultivated rice (*Oryza sativa* L.), weedy rice (*O. sativa* f. *spontanea*) and common wild rice (*O. rufipogon* Griff.) by gene flow and are normally expressed in hybrid progenies (Song *et al.*, 2003; Lu *et al.*, 2016).

In short, a large number of studies and investigations have shown that if hybridization is possible and environmental conditions are favorable in time and space, the transgenes in GM crops could escape to their wild relatives through gene flow and be expressed normally in hybrid progenies (Cao *et al.*, 2014; Lu, 2014; Lu *et al.*, 2016; Ellstrand, 2018).

5.2 Gene Flow between Cultivated Soybean and Wild Soybean

5.2.1 Distribution of wild soybean

Wild soybean (*Glycine soja* Seib. et Zucc.) is a relative of cultivated soybean (*Glycine max* (L.) Merr.), an annual, self-pollinating herb, and is one of the national second-class protected wild plants in China. Wild soybean is distributed in China, Russia, Japan, North Korea, and South Korea, and China is the center of biodiversity origin of wild soybean, which is widely distributed in all provinces except Xinjiang, Qinghai

and Hainan. It mainly grows in farmland and natural ecosystems near fields, gardens, ditches, riverbanks, and lakes at altitudes of 150–2650 m above sea level. Although cultivated soybean and wild soybean belong to two different species, there is no obvious reproductive isolation, and they have the same genome (both chromosomes are $2n = 40$), meiotic chromosome pairing pattern and similar genome structure and composition. They are not only easy to cross and bear good fruit, but also the genetic mode of hybrid progeny is similar to that of hybrid progenies of cultivated soybean varieties; the exchange of genetic material is relatively easy between the two species (Fan *et al.*, 2012; Dong and Yang, 2015). The natural distribution area of wild soybean in China highly overlaps that of the cultivated soybean planting area, and the growth period is basically the same (Wang and Li, 2013a). Interspecific hybrids are fertile. If GM soybeans are planted commercially on a large scale and for a long time in China, gene flow may potentially affect wild soybean populations. In addition, wild soybean has a large number of seeds per plant, strong shattering, small seeds and strong dormancy. Once the transgene flows into the genome of wild soybean, it is necessary to manage and control its biosafety issues. Cultivated soybeans can transmit their genes to wild soybeans through pollen, causing the latter to contain some genes of cultivated soybeans (wild soybeans have infiltrated 0.73% of cultivated soybean genes) (Wang *et al.*, 2012). Some rare traits of wild soybeans (such as white flowers, gray hairs and yellow, green and brown seed coats) may come from cultivated soybeans (Kuroda *et al.*, 2006; Wang *et al.*, 2010, 2012; Wang and Li, 2012; Liu *et al.*, 2020). Therefore, in order to protect the biological resources of wild soybean in China, the scientific problem of gene flow to wild soybean must be studied systematically and deeply before the commercial planting of GM soybean.

5.2.2 Gene flow between cultivated soybean and wild soybean

The frequency of hybridization between different varieties of cultivated soybeans is usually less than 3% (Ahrent and Caviness, 1994; Ray *et al.*, 2003). In the case of interrow planting, the natural outcrossing rate is 0.65–6.32%, the average natural outcrossing rate is 1.8%, and the natural outcrossing rate at a distance of more than 10 m is less than 0.01%. The intraspecies crossing between cultivated soybeans and cultivated soybeans can be carried out by pollination by insects (such as bees), but wind pollination is almost negligible (Caviness, 1966; Erickson, 1975; Ahrent and Caviness, 1994; Ray *et al.*, 2003; Yoshimura, 2011).

The outcrossing rate between cultivated soybean and wild soybean is generally lower than that between cultivated soybean plants. The average natural outcrossing rate of Japanese cultivated soybean and wild soybean was 0.73% under conditions of 50 cm spacing and 30-day flowering overlap (Nakayama and Yamaguchi, 2010. Kuroda *et al.* (2006) found that 6.8% of Japanese wild soybean plants contained genes from cultivated soybean. The hybrid offspring of cultivated soybean and wild soybean can survive for a long time under natural conditions (Oka, 1983; Wang and Li, 2011, 2013a). Cultivated soybean can transmit its genes to wild soybean through pollen, that is, the genes of cultivated soybean have infiltrated into the related wild soybean. The study by Liu (2008) on germplasm material of 96 wild soybean from 26 provinces showed that some wild soybean individuals contained genes or genetic components from cultivated soybean. *Megachile tsurugensis* and other flower-visiting insects are the main vectors of gene flow from cultivated soybean to wild soybean (Nakayama and Yamaguchi, 2010; Wang and Li, 2011, 2013a, b; Kuroda *et al.*, 2010).

There is also gene flow among different individuals and populations of wild soybean. The results of Japanese scholars show that, because pollinating insects such as European honey bee (*Apis mellifera*) and carpenter bee (*Xylocopa appendiculata*) regularly visit the flowers of wild soybean, the outcrossing rate of multiple loci among four wild soybean populations in Japan can reach 9.3–19%, and the gene flow frequency within wild soybean populations is higher than that between cultivated soybean and wild soybean (Fujita *et al.*, 1997; Kuroda *et al.*, 2008). There are also reports on gene flow within wild soybean populations and

among adjacent different populations in China. The outcrossing rate among wild soybean natural populations is from 0% to 3.5%; and, within the range of 50 km, the closer the distance between different wild soybean materials, the higher is the frequency of gene flow (He *et al.*, 2012; Wang and Li, 2012; Wang *et al.*, 2014, 2015).

It can be seen that low-frequency gene flow occurs between cultivated soybean and wild soybean and between different populations of wild soybean, and that insects are the main vectors of gene flow. If GM soybeans are grown commercially in China, transgenes can flow into wild soybeans and may spread further among different wild soybean populations.

5.2.3 Gene flow from GM soybeans to non-GM soybeans and wild soybeans

In order to study and evaluate the risk of gene flow of GM soybeans, scientists have carried out field experiments on gene flow from GM soybeans to non-GM cultivars as well as from GM soybeans to wild soybeans (Table 5.1).

In a field experiment of gene flow between GM and non-GM soybeans in Brazil, when non-GM soybeans were 1 m away from *epsps* GM soybean pollen donors, the average gene flow frequency was 0.52%. The average gene flow frequency at 2 m was 0.12%, and no gene flow was detected when the distance increased to 10 m (Abud *et al.*, 2007). The results of a 4-year experiment conducted by Japanese scholars showed that the gene flow frequency between *epsps* GM (AG3701RR, Event 40-3-2) and non-GM soybeans was the highest (0.19%) at a distance of 0.7 m, and that gene flow could not be detected when the distance increased to 10.5 m (Yoshimura *et al.*, 2006). In the interrow planting and concentric circle planting experiments in China, the natural outcrossing rate of *epsps* GM herbicide-resistant soybean AG5601 to 36 conventional non-GM soybean varieties ranged from 0.039% to 0.934% at the distance of 0.5–15 m, and the outcrossing rate between AG5601 and cv. 'Zhonghuang13' at the farthest flow distance of 15 m was 0.012% (Huang *et al.*, 2014). The results of Liu *et al.* (2012) showed that the gene flow frequency was 0.03% at a distance of 5 m between *epsps* GM soybeans and various varieties of non-GM soybeans, but decreased to 0.001% at 29 m. Experiments in China and Japan have found that gene flow between GM soybeans and conventional soybeans occurs mainly through pollination insects, and the probability of gene flow caused by wind pollination is very low (Yoshimura *et al.*, 2006; Huang *et al.*, 2014).

The published studies on gene flow between GM soybeans and wild soybeans are mainly by Chinese and Japanese scholars. Chen *et al.* (2003) found that the frequency of gene flow between *epsps* GM soybean ARG04 and wild soybean was about 0.09%, and insects were the main pollen vectors. The results of another experiment showed that the farthest flow distance between *epsps* GM soybean ARG04 and wild soybean was 10 m, and the outcrossing frequency between GM soybean ARG04 and wild soybean was less than 1/10,000 (Chen *et al.*, 2004). In two other field experiments, no gene flow from *epsps* GM soybean to wild soybean was detected (Lü *et al.*, 2009; Liu *et al.*, 2012). In a gene flow test in Japan, it was found that the flowering coincidence period between GM soybean AG3701RR and wild soybean was 18–24 days (Mizuguti *et al.*, 2009). Although only one of the 32,502 F_1 soybean seedlings in this test was GM positive, this hybrid GM soybean produced 947 seeds, of which 824 seeds developed into normal plants. Another study by Mizuguti *et al.* (2010) showed that the hybridization frequency between GM soybeans and wild soybeans was from 0% to 0.097%, and that the higher the flowering coincidence degree between GM soybeans and wild soybeans, the higher was the hybridization frequency between them.

In 2017, we used *epsps+pat* GM soybean S4003.14 developed by China Dabeinong Biotechnology Co. Ltd as material to study its gene flow to five non-GM cultivated soybeans and five wild soybeans under field conditions in Yitong County, Jilin Province (Fig. 5.1). The results showed that the flowering overlap time of S4003.14 with the five non-GM cultivated soybeans and the five wild soybeans was 17–27 days and 19–23 days, respectively. With interrow planting, the average outcrossing rates of gene flow of S4003.14 to five non-GM cultivated soybeans and to

Table 5.1. Outcrossing rate of gene flow among different soybean materials.

Direction	Distance (m)	Outcrossing rate (%)	References
Cultivated soybean to wild soybean	0.5	0.73	Nakayama and Yamaguchi, 2010
	2	0.88	Wang and Li, 2011
	Field sampling (no distance data available)	2.2	Kuroda *et al*., 2006
	Field sampling (no distance data available)	2.4–3.0	Kiang *et al*., 1992
	Field sampling (no distance data available)	9.3–19	Fujita *et al*., 1997
GM soybean to non-GM cultivars	0.5	0.263	Huang *et al*., 2014
	0.7	0.19	Yoshimura *et al*., 2006
	0.7	0.05	Liu *et al*., 2012
	1	0.52	Abud *et al*., 2007
	1	0.104	Huang *et al*., 2014
	1.4	0.04	Yoshimura *et al*., 2006
	1.5	0.045	Huang *et al*., 2014
	2	0.12	Abud *et al*., 2007
	2	0.025	Huang *et al*., 2014
	2.1	0.034	Yoshimura *et al*., 2006
	2.5	0.023	Huang *et al*., 2014
	2.8	0.017	Yoshimura *et al*., 2006
	3.5	0.038	Yoshimura *et al*., 2006
	5	0.022	Huang *et al*., 2014
	5	0.03	Liu *et al*., 2012
	7.0	0.040	Yoshimura *et al*., 2006
	10	0	Abud *et al*., 2007
	10	0.023	Huang *et al*., 2014
	10.5	0	Yoshimura *et al*., 2006
	15	0.012	Huang *et al*., 2014
	20	0	Huang *et al*., 2014
	29	0.001	Liu *et al*., 2012

Continued

Table 5.1. Continued

Direction	Distance (m)	Outcrossing rate (%)	References
GM soybean to wild soybean	0	0.296	Yook *et al*., 2021
	0.05	0.003	Mizuguti *et al*., 2009
	0.05	0.0388–0.0971	Mizuguti *et al*., 2010
	0.25	0.259	Yook *et al*., 2021
	0.5	0.192, 0.202	Yook *et al*., 2021
	1	0.206, 0.177	Yook *et al*., 2021
	2	0–0.0133	Mizuguti *et al*., 2010
	2	0.085, 0.204	Yook *et al*., 2021
	4	0–0.0134	Mizuguti *et al*., 2010
	4	0.057, 0.096	Yook *et al*., 2021
	6	0–0.0133	Mizuguti *et al*., 2010
	6	0.059, 0.045	Yook *et al*., 2021
	8	0	Mizuguti *et al*., 2010
	8	0.025	Yook *et al*., 2021
	10	0	Mizuguti *et al*., 2010
	0–50	0	Kuroda *et al*., 2008
	0–50[a]	0.007 9	Chen *et al*., 2004
	5–29	0	Liu *et al*., 2012

[a]Only one plant survived at the distance of 10 m.

five wild soybeans were 0.16% and 0.06%, respectively, and the fertility of hybrid progenies was normal. Figure 5.2 shows one of the surviving hybrid progenies under the selection pressure of herbicide application. When the distance between S4003.14 soybean and non-GM soybean was more than 1 m, no gene flow was detected. The results of this case show that although the outcrossing rate of gene flow from GM soybeans to non-GM cultivated soybeans and wild soybeans is very low, due to the large planting area of cultivated soybeans in China and the high overlapping between the distribution of wild soybeans and cultivated soybeans (Liu *et al*., personal observation; Wang and Li, 2013a), transgenes in GM soybeans could potentially flow into non-GM cultivars and wild soybeans through pollen with large-scale release of GM soybeans in China. If the gene flow occurs to the non-GM cultivated soybean and the seeds of the hybrid offspring are mainly used as processing raw materials after harvest, the environmental risk may be smaller and easier to control; if the gene flow occurs to the wild soybean in nature, it means that the transgenes will enter the wild populations that mainly grow in the natural environment, which may cause unpredictable risk to the wild soybean resources.

Although gene flow frequency is low, the hybrids formed between cultivated soybean and wild soybean are fertile and mainly self-pollinated to produce further generations in the field (Guan *et al*., 2015; Kan *et al*., 2015; Liu *et al*., 2021); thus the risk of transgene introgression is not negligible. Besides sexual compatibility, other main factors affecting gene flow frequency from cultivated soybean to wild soybean are the physical distance and flowering-overlap time. If the transgenes in GM soybeans are introduced into wild soybeans, the fitness effect caused by GM soybeans should be analyzed and evaluated (Yook *et al*., 2021; Liu *et al*., 2021), which is very important for the risk assessment of gene flow and for the protection of wild plant genetic resources (Lu, 2015; Lu *et al*., 2016).

Fig. 5.1. Planting and experimental scheme of gene flow from GM *epsps+pat* soybean to non-GM soybeans. The central part of the concentric arcs is the area planted with GM soybean, and the red concentric arcs represent a series of sampling distances.

5.3 Ecological Consequences of Gene Flow from GM Soybean to Wild Soybean

Fitness analysis of hybrid progenies of GM crops and wild related species is an important step to evaluate the ecological consequences caused by gene flow (Song *et al.*, 2004; Snow *et al.*, 2010; Lu, 2015). Kuroda *et al.* (2013) thought that some artificial domestication genes from cultivated soybean may not be conducive to the survival of hybrid offspring in the natural environment, so the fitness of hybrid offspring may be lower than that of wild soybean. There are only a few published reports regarding the consequences of gene flow so far.

Guan *et al.* (2015) obtained hybrid F_1 seeds of GM soybean with *epsps* gene and wild soybean by artificial pollination, and then studied the fitness of F_1 and F_2 generations under greenhouse and field conditions. It was found that although the pod-setting rate of the F_1 generation was very low, the F_1 soybean with *epsps* gene had longer vegetative growth periods, higher biomass and lower 100-seed weight than the F_1 soybean without *epsps* gene, and that the 100-seed weight of F_1 and F_2 was significantly higher than that of wild soybean. The results showed that the *epsps* gene had no adverse effect on the growth of hybrid progenies containing the *epsps* gene, and it was speculated that the escaped *epsps* gene could exist in wild soybean population without glyphosate herbicide selection pressure. The results of the study carried out by Kan *et al.* (2015) in the greenhouse showed that four wild soybean materials and *epsps*

Fig. 5.2. One of the soybean plants of positive hybrid progenies survived after spraying with herbicides.

GM soybean could cross and produce hybrid F_1 and F_2 progeny with glyphosate resistance. In the absence of glyphosate selection pressure, the relative fitness of hybrid progenies in some characters was higher than that of female wild soybean lines (especially reproduction indices such as pod setting rate, 100-seed weight and seed quality per plant). There was no significant difference in other fitness indices (Kan *et al.*, 2015). Yook *et al.* (2021) assessed the potential weed risk of hybrids produced by gene flow from glyphosate resistant (GR) soybean to wild soybean in South Korea through a 2-year field study. Hybrid F_1 and F_2 progeny showed the intermediate characters of parent soybean during vegetative growth and reproductive growth, and the numbers of flowers, pods, and seeds per plant were similar to those of wild soybean, but significantly higher than those of transgenic soybean. Liu *et al.* (2021) measured the fitness of F_1 hybrids obtained from ten wild soybean populations collected from China and for comparison with transgenic glyphosate-resistant soybean. The results showed that the emergence rate of all hybrid seeds was in the range of 13.33–63.33% in the absence of weed competition. Compared with wild progenitors, the composite fitness of nine F_1 hybrids was significantly lower. One of the special F_1 was IMBT F_1, whose composite fitness is 1.28, which is similar to its wild ancestor, and all the F_1 hybrids could produce offspring.

The results of those reported studies (Guan *et al.*, 2015; Kan *et al.*, 2015; Yook *et al.*, 2021; Liu *et al.*, 2021) showed that in the absence of glyphosate, some fitness parameters of the progenies of GM soybean–wild soybean hybrids were significantly higher than those of wild soybean. That finding was not consistent with the prediction of Kuroda *et al.* (2013), which indicates that additional research is needed.

5.4 Current Problems in the Study of Gene Flow from GM Soybeans to Wild Soybeans

To sum up, although there have been some research reports on gene flow from GM soybeans to wild soybeans, there are still unanswered questions.

First of all, the huge genetic differences among wild soybean populations in different regions of China were not taken into account. Except for Qinghai, Xinjiang and Hainan provinces, wild soybean is distributed in most

provinces in China. Due to the huge differences in ecological, geographic and climatic conditions, wild soybeans in different regions of China have formed various ecotypes and populations adapted to the local environment. There are great differences in phenotype and genetic diversity among different ecotypes of wild soybean in leaf type, flower color, fluff color, seed size, seed coat color, stem morphology, and growth period. In particular, the differences in growth period (especially flowering), flower color and seeds of different ecotypes of wild soybean will directly affect the frequency of gene flow between wild soybean and transgenic soybean, as well as the survival and fitness of hybrid offspring under different environmental conditions (Zhao *et al.*, 2008; Wang *et al.*, 2012). The published studies on gene flow from GM soybeans to wild soybeans only involve a few geographic populations. Without taking into account the huge genetic differences among wild soybeans in different regions of China, the risk levels of gene flow from GM soybeans to wild soybeans may not be concluded comprehensively. The growth and fitness of hybrid offspring may perform differently under different soil, temperature and other local environmental conditions in different regions, which may affect the results of environmental risk assessments.

Secondly, the transgenes involved in GM soybeans have focused on the *epsps* gene, and only a few of them involve other transgenes (such as the *Bt* gene) and multiple transgenes of stacked traits. In addition to the *epsps* gene, the transgenes expressed in a large number of GM soybeans imported into China and self-developed GM soybeans include: the aroxy chain alkylate dioxygenase-12 gene (*aad-12*), which can degrade 2,4-D herbicides; phosphonectin acetyltransferase gene (*pat*), which inactivates glyphosate herbicides; *Bt* genes, which are toxic to Lepidoptera pests; and multiple foreign genes in GM soybean of stacked traits (for example, GM soybean DAS-44406-6 developed by Dow Yinong Company contains three transgenes: 2*mepsps*, *aad*-12 and *pat*) (Ministry of Agriculture and Rural Affairs and PRC, 2019). If different foreign genes and combinations of foreign genes enter the genome of wild soybean through gene flow, the fitness effect and environmental risks could become complicated, which should be evaluated and studied according to the case-by-case principle.

Thirdly, there is a lack of research on the fitness of hybrid backcross progenies produced by gene flow from GM soybean to wild soybean. The potential ecological risks caused by gene flow from GM soybean to wild soybean depend on the expression of foreign genes in hybrid offspring and the fitness of hybrid progenies (Hails and Morley, 2005; Lu, 2015). However, only a few research cases have been publicly reported, and the contents of the studies do not involve the changes of transgene expression, growth period, dormancy, and shattering of hybrid offspring under different environmental conditions. Moreover, they do not take into account the risk that transgenes may spread among different wild soybean populations through backcrossing with wild soybean. In addition, the technical standard for environmental safety assessment of GM soybeans in China (Ministry of Agriculture notice No. 2031-3-2013: 'Environmental Safety testing of GM plants and their products – part 3: foreign gene flow') only stipulates the need to evaluate the distance and frequency of gene flow from GM soybeans to cultivated and wild soybeans. There is no fitness evaluation of hybrid progenies produced by gene flow from GM soybeans to wild soybeans. It can be seen that there is a lack of requirements for evaluating the fitness of hybrid offspring caused by gene flow from GM soybeans to wild soybeans in the current technical standards for environmental safety assessment of GM soybeans issued and implemented by the Chinese government, and there are also deficiencies in the breadth and depth of related scientific research.

5.5 Perspectives on Future Study of Gene Flow from GM Soybean to Wild Soybean

In order to meet the huge demand for soybeans in China, in addition to continuing to import GM soybeans on a large scale, the voice of China's soybean industry for the commercial cultivation of GM soybeans is getting louder. The protection of wild related species of crops has attracted the attention of international organizations

and governments, and wild soybean is a valuable biological genetic resource in China and in the world (Kofsky *et al.*, 2018; FAO, 2019). Large-scale commercial cultivation of GM soybeans will lead to the risk of gene flow to wild soybeans. In order to carry out the commercialization of GM soybeans in the future while protecting wild soybeans, it is necessary to conduct in-depth and systematic evaluation and research on the risks caused by gene flow from GM soybeans to wild soybeans from the following aspects.

5.5.1 The risk of gene flow between GM soybean and wild soybean in different geographic regions of China

In regions that are not only important producing areas of cultivated soybean but are also known for the large-scale distribution of wild soybean in China, the geographic distribution, flowering-overlap time, species, and number of pollinating insects of cultivated soybean and wild soybean should be investigated. The geographic overlap, hybridization compatibility, distance, and frequency of gene flow between GM soybean and wild soybean in these regions should be evaluated. It is important to determine the risk level of gene flow between GM soybean and wild soybean in different regions, and to provide scientific data for determining the possible key areas of GM soybean release and carrying out safety management.

5.5.2 Fitness of backcross progenies formed between GM soybean and wild soybean

First of all, it is necessary to study the fitness of hybrid progenies between wild soybean and GM soybean expressing other transgenes in addition to *epsps* gene, or GM soybean expressing multiple transgenes. Secondly, if transgenes flow from GM soybean into wild soybean, it may not only be passed on through self-crossing of hybrid offspring, but may also spread within and between wild soybean populations through backcrossing between hybrid offspring and wild soybean. Therefore, it is also necessary to study and evaluate the fitness of backcross progenies produced by wild soybean and hybrid progeny between GM soybean and wild soybean. Thirdly, we should consider the diversity of the growth environment of wild soybean, and should study and evaluate the fitness of GM soybean–wild soybean hybrid progeny under various typical natural growth conditions of wild soybean as far as possible. Fourthly, the measured indicators should include not only vegetative growth and reproductive growth, but also foreign gene expression, dormancy, and shattering.

5.5.3 Effect of *epsps* gene on fitness of wild soybean

From a worldwide point of view, there are many kinds of herbicide tolerance genes that have been put into commercial application, among which the *epsps* gene is the most widely used (Yang *et al.*, 2014). Most of the GM soybean varieties imported into China contain the *epsps* gene. At present, the foreign genes in self-developed GM soybeans with industrialization value are mainly *epsps* gene. Under the condition of applying the target herbicide glyphosate, hybrid backcross progenies of GM crops–wild related species expressing the *epsps* gene may have significant selection advantage over hybrids without *epsps* gene expression. Some recent research results suggest that wild relatives of crops that are transferred and express *epsps* genes through gene flow may significantly increase fitness in natural ecosystems without glyphosate, resulting in unpredictable risks (Guan *et al.*, 2015; Kan *et al.*, 2015; Yook *et al.*, 2021; Liu *et al.*, 2021). The possible reason for this phenomenon is that the aromatic amino acids produced by the shikimic acid pathway play an important role in maintaining plant metabolism, growth and reproduction, and the EPSPS protein is the key enzyme dependent on shikimic acid pathway in almost all plants and microorganisms (Beres *et al.*, 2018; Fang *et al.*, 2018; Yang *et al.*, 2017). Although the structures of EPSPS proteins from different sources are different, their biological functions are similar (Herrmann, 1995; Gong *et al.*, 2015). Therefore, unlike the results of fitness cost of *Bt* GM crops in the absence of target pests (Lu *et al.*,

2016), the risk of the glyphosate-resistant *epsps* transgene flowing to wild soybean needs to be further evaluated and studied.

5.5.4 The risk of gene flow from imported GM soybean to wild soybean

Over the past decade, China has been the world's largest importer of GM soybean, with imports reaching 95.53 million tons in 2017 (Yao, 2020). In the process of domestic transport, storage, and processing of a large amount of imported GM soybean, a small number of seeds may be accidentally lost to farmland or natural ecosystems, causing gene flow to cultivated and wild soybean. In Japan and South Korea, there have been cases in which imported GM soybean and maize have been spilled and grow near transport routes and storage sites (Lee *et al*., 2009; Goto *et al*., 2017). In view of the fact that China has imported and will continue to import a large number of GM soybeans, investigations on the residues of imported GM soybean in the domestic transportation routes, storage, and processing sites and their surrounding environment should be carried out to provide more data support for the assessment and management of the risk of gene flow caused by imported GM soybean.

5.5.5 Safety control, monitoring and management measures for gene flow from GM soybean to wild soybean

In summary, there are reported cases of gene flow between conventional soybean and wild soybean. Taking into account the extremely wide geographic distribution of wild soybean in China and the scattered planting model of cultivated soybean dominated by smallholder farmers, it is essential to develop safety management measures in advance, before commercial release of GM soybeans that are suitable for China's national conditions. After commercial release, it is especially important to monitor hybrid offspring in the field formed between GM soybean and wild soybean to eliminate voluntary seedlings of hybrid offspring with foreign genes in time.

Acknowledgments

This paper is supported by the 'National Major Special Projects for Genetically Modified Organisms (2016ZX08012-005)'.

References

Abud, S., Souza, P.I.M., Vianna, G.R., Leonardecz, E., Moreira, C.T. *et al*. (2007) Gene flow from transgenic to nontransgenic soybean plants in the Cerrado region of Brazil. *Genetics and Molecular Research* 6(2), 445–452. DOI: 10.1590/S1415-47572007000600029.

Ahrent, D.K. and Caviness, C.E. (1994) Natural cross-pollination of twelve soybean cultivars in Arkansas. *Crop Science* 34, 376. DOI: 10.2135/cropsci1994.0011183X003400020013x.

Beckie, H.J., Warwick, S.I., Nair, H. and Séguin-Swartz, G. (2003) Gene flow in commercial fields of herbicide-resistant Canola (*Brassica napus*). *Ecological Applications* 13, 1276–1294. DOI: 10.1890/02-5231.

Beres, Z.T., Yang, X., Jin, L., Zhao, W. and Mackey, D.M. (2018) Overexpression of a native gene encoding 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) may enhance fecundity in *Arabidopsis thaliana* in the absence of glyphosate. *International Journal of Plant Sciences* 179, 390–401. DOI: 10.1086/696701.

Campbell, L.G., Snow, A.A. and Ridley, C.E. (2006) Weed evolution after crop gene introgression: greater survival and fecundity of hybrids in a new environment. *Ecology Letters* 9(11), 1198–1209. DOI: 10.1111/j.1461-0248.2006.00974.x.

Cao, D., Stewart, C.N. Jr., Zheng, M., Guan, Z.-J., Tang, Z.-X. *et al*. (2014) Stable *Bacillus thuringiensis* transgene introgression from *Brassica napus* to wild mustard *B. juncea*. *Plant Science* 227, 45–50. DOI: 10.1016/j.plantsci.2014.06.018.

Caviness, C.E. (1966) Estimates of natural crosspollination in Jackson soybeans in Arkansas. *Crop Science* 6, 211. DOI: 10.2135/cropsci1966.0011183X000600020034x.

Chen, X., Wang, C.-Y., Zhu, C.-S., Yan, J.-Y. and Gao, B. (2003) Evaluation of safety for transgenic Roundup Ready soybean and its influence on environment. *Jiangsu Agricultural Sciences* 6, 36–40. [In Chinese with English abstract.] DOI: 10.3969/j.issn.1002-1302.2003.06.013.

Chen, X., Yan, J.-Y. and Gao, B. (2004) Preliminary study on Roundup Ready soybean's roundup ready gene move to wild soybean. *Chinese Journal of Oil Crop Sciences* 26, 89–91. [In Chinese with English abstract.] DOI: 10.3321/j.issn:1007-9084.2004.02.022.

Crawley, M.J., Hails, R.S., Rees, M., Kohn, D. and Buxton, J. (1993) Ecology of transgenic oilseed rape in natural habitats. *Nature* 363, 620–623. DOI: 10.1038/363620a0.

Dale, P.J. (1992) Spread of engineered genes to wild relatives. *Plant Physiology* 100(1), 13–15. DOI: 10.1104/pp.100.1.13.

Dalton, R. (2008) Modified genes spread to local maize. *Nature* 456(7219), 149. DOI: 10.1038/456149a.

Dong, Y.S. and Yang, G.Y. (2015) *Research and Utilization of Wild Soybean Resources in China*. Shanghai Science and Technology Education Press, Shanghai, China.

Ellstrand, N.C. (2001) When transgenes wander, should we worry? *Plant Physiology* 125(4), 1543–1545. DOI: 10.1104/pp.125.4.1543.

Ellstrand, N.C. (2018) 'Born to Run'? Not necessarily: species and trait bias in persistent free-living transgenic plants. *Frontiers in Bioengineering and Biotechnology* 6, 88. DOI: 10.3389/fbioe.2018.00088.

Ellstrand, N.C., Prentice, H.C. and Hancock, J.F. (1999) Gene flow and introgression from domesticated plants into their wild relatives. *Annual Review of Ecology and Systematics* 30, 539–563. DOI: 10.1146/annurev.ecolsys.30.1.539.

Erickson, E.H. (1975) Effect of honey bees on yield of three soybean cultivars. *Crop Science* 15, 84. DOI: 10.2135/cropsci1975.0011183x001500010025x.

Fan, H., Zhao, T.J., Ding, Y.L., Xing, G.N. and Gai, J.Y. (2012) Genetic analysis of the characteristics and geographic differentiation of Chinese wild soybean population. *Scientia Agricultura Sinica* 45, 414–425.

Fang, J., Nan, P., Gu, Z.-Y., Ge, X.-C., Feng, Y.-Q. *et al*. (2018) Overexpressing exogenous 5-enolpyruvyl shikimate-3-phosphate synthase (*epsps*) genes increases fecundity and auxin content of transgenic *Arabidopsis* plants. *Frontiers in Plant Science* 9, 233. DOI: 10.3389/fpls.2018.00233.

FAO (2019) Bélanger, J. and Pilling, D. (eds) *The State of the World's Biodiversity for Food and Agriculture*. FAO Commission on Genetic Resources for Food and Agriculture Assessments, Rome.

Fujita, R., Ohara, M., Okazaki, K. and Shimamoto, Y. (1997) The extent of natural cross-pollination in wild soybean (*Glycine soja*). *Journal of Heredity* 88, 124–128. DOI: 10.1093/oxfordjournals.jhered.a023070.

Gong, Y.-Y., Guo, S.-Q., Shu, H.-M., Ni, W.-C. and Shen, X.-L. (2015) Analysis of molecular evolution and gene structure of epsps protein in plant shikimate pathway. *Chinese Bulletin of Botany* 50, 295–309. [In Chinese with English abstract.] DOI: 10.3724/SP.J.1259.2015.00295.

Goto, H., Mcpherson, M.A., Comstock, B.A., Stojsin, D. and Ohsawa, R. (2017) Likelihood assessment for gene flow of transgenes from imported genetically modified soybean [*Glycine max* (L.) Merr.] to wild soybean (Glycine soja Seib. et Zucc.) in Japan as a component of environmental risk assessment. *Breeding Science* 67(4), 348–356. DOI: 10.1270/jsbbs.16134.

Guan, Z.-J., Zhang, P.-F., Wei, W., Mi, X.-C., Kang, D.-M. *et al*. (2015) Performance of hybrid progeny formed between genetically modified herbicide-tolerant soybean and its wild ancestor. *AoB Plants* 7, plv121. DOI: 10.1093/aobpla/plv121.

Hails, R.S. and Morley, K. (2005) Genes invading new populations: a risk assessment perspective. *Trends in Ecology & Evolution* 20(5), 245–252. DOI: 10.1016/j.tree.2005.02.006.

Hall, L., Topinka, K., Huffman, J. and Good, D.A. (2000) Pollen flow between herbicide-resistant *Brassica napus* is the cause of multiple-resistant *B. napus* volunteers. *Weed Science* 48, 688–694. DOI: 10.1614/0043-1745(2000)048[0688:PFBHRB]2.0.CO;2.

Hancock, J.F. (2003) A framework for assessing the risk of transgenic crops. *Bio Science* 53, 512–519. DOI: 10.1641/0006-3568(2003)053[0512:AFFATR]2.0.CO;2.

Haygood, R., Ives, A.R. and Andow, D.A. (2003) Consequences of recurrent gene flow from crops to wild relatives. *Proceedings of the Royal Society of London B: Biological Sciences* 270(1527), 1879–1886. DOI: 10.1098/rspb.2003.2426.

He, S., Wang, Y., Volis, S., Li, D. and Yi, T. (2012) Genetic diversity and population structure: implications for conservation of wild soybean (*Glycine soja* Sieb. et Zucc) based on nuclear and chloroplast microsatellite variation. *International Journal of Molecular Sciences* 13(10), 12608–12628. DOI: 10.3390/ijms131012608.

Herrmann, K.M. (1995) The shikimate pathway: early steps in the biosynthesis of aromatic compounds. *Plant Cell* 7, 907–919. DOI: 10.1105/tpc.7.7.907.

Huang, W.-K., Peng, H., Wang, G.-F., Cui, J.-K. and Zhu, L.-F. (2014) Assessment of gene flow from glyphosate-resistant transgenic soybean to conventional soybean in China. *Acta Physiologiae Plantarum* 36, 1637–1647. DOI: 10.1007/s11738-014-1539-3.

ISAAA (2019) *Global Status of Commercialized Biotech/GM Crops: 2019. ISAAA Brief No. 55*. International Service for the Acquisition of Agri-biotech Applications, Ithaca, New York.

Kan, G.-Z., Tong, Z.-F., Hu, Z.-B., Ma, D.-Y. and Zhang, G.-Z. (2015) Fitness of hybrids between wild soybeans (*Glycine soja*) and the glyphosate-resistant transgenic soybean (*Glycine max*). *Soybean Science* 34, 177–184. [In Chinese with English abstract.] DOI: 10.11861/j.issn.1000-9841.2015.02.0177.

Kiang, Y.-T., Chiang, Y.-C. and Kaizuma, N. (1992) Genetic diversity in natural populations of wild soybean in Iwate prefecture, Japan. *Japanese Journal of Heredity* 83, 325–329. DOI: 10.1007/BF01900718.

Kofsky, J., Zhang, H.-Y. and Song, B.-H. (2018) The untapped genetic reservoir: the past, current, and future applications of the wild soybean (*Glycine soja*). *Frontiers in Plant Science* 9, 949. DOI: 10.3389/fpls.2018.00949.

Kuroda, Y., Kaga, A., Tomooka, N. and Vaughan, D.A. (2006) Population genetic structure of Japanese wild soybean (*Glycine soja*) based on microsatellite variation. *Molecular Ecology* 15(4), 959–974. DOI: 10.1111/j.1365-294X.2006.02854.x.

Kuroda, Y., Kaga, A., Tomooka, N. and Vaughan, D.A. (2008) Gene flow and genetic structure of wild soybean (*Glycine soja*) in Japan. *Crop Science* 48, 1071–1079. DOI: 10.2135/cropsci2007.09.0496.

Kuroda, Y., Kaga, A., Tomooka, N. and Vaughan, D. (2010) The origin and fate of morphological intermediates between wild and cultivated soybeans in their natural habitats in Japan. *Molecular Ecology* 19(11), 2346–2360. DOI: 10.1111/j.1365-294X.2010.04636.x.

Kuroda, Y., Kaga, A., Tomooka, N., Yano, H., Takada, Y. *et al*. (2013) QTL affecting fitness of hybrids between wild and cultivated soybeans in experimental fields. *Ecology and Evolution* 3(7), 2150–2168. DOI: 10.1002/ece3.606.

Lee, B., Kim, C.-G., Park, J.-Y., Park, K.-W. and Kim, H.-J. (2009) Monitoring the occurrence of genetically modified soybean and maize in cultivated fields and along the transportation routes of the Incheon port in South Korea. *Food Control* 20, 250–254. DOI: 10.1016/j.foodcont.2008.05.006.

Liu, B., Xue, K., Liu, L.-P., Zhou, Y. and Han, J. (2020) Progress on the gene flow from genetically modified soybeans to wild soybeans. *Journal of Ecology and Rural Environment* 36, 833–841. [In Chinese with English abstract.] DOI: 10.19741/j.issn.1673-4831.2019.0407.

Liu, J., Zhou, B., Yang, C.-Y., Li, Y.-R. and Jiang, L.-X. (2012) Gene flowing of genetically modified glyphosate-resistant soybean with *epsps*. *Soybean Science* 31, 517–521. [In Chinese with English abstract.] DOI: 10.3969/j.issn.1000-841.2012.04.002.

Liu, J.-Y., Sheng, Z.-W., Hu, Y.-Q., Liu, Q., Qiang, S. *et al*. (2021) Fitness of F_1 hybrids between 10 maternal wild soybean populations and transgenic soybean. *Transgenic Research* 30(1), 105–119. DOI: 10.1007/s11248-020-00230-x.

Liu, Y.-N. (2008) An analysis of genetic diversity in a mini-core collection of wild soybean (G. soja) in China. Master's thesis [in Chinese]. Chinese Academy of Agricultural Sciences, Beijing. DOI: 10.7666/d.Y1421977.

Lu, B.-R. (2014) Genetic and evolutionary effects of hybridization-introgression and their implications for conservation of crop wild relative species. *Chinese Science Bulletin* 59, 479–492. [In Chinese with English abstract.] DOI: 10.1360/972013-660.

Lu, B.-R. (2015) Analysis of fitness effect and its application in assessing environmental risk caused by transgene flow. *Biotechnology Bulletin* 31, 7–16. [In Chinese with English abstract.] DOI: 10.13560/j.cnki.biotech.bull.1985.2015.03.007.

Lu, B.-R. and Yang, C. (2009) Gene flow from genetically modified rice to its wild relatives: assessing potential ecological consequences. *Biotechnology Advances* 27(6), 1083–1091. DOI: 10.1016/j.biotechadv.2009.05.018.

Lu, B.-R., Xia, H., Yang, X., Jin, X. and Liu, P. (2009) Evolutionary theory of hybridization-introgression: its implication in environmental risk assessment and research of transgene escape. *Biodiversity Science* 17, 362–377. [In Chinese with English abstract.] DOI: 10.3724/SP.J.1003.2009.09132.

Lu, B.-R., Wang, L. and Wang, Z. (2016) Can transgene flow lead to environmental biosafety impacts in rice? *Scientia Sinica Vitae* 46, 420–431. [In Chinese with English abstract.] DOI: 10.1360/N052016-00127.

Lü, X.-B., Wang, H.-Y., Liu, Q., Zhao, G. and Li, X.-C. (2009) Biosafety of Roundup Ready soybean planted in black soil ecosystem. [In Chinese with English abstract]. *Soybean Science* 28, 260–265.

Ministry of Agriculture and Rural Affairs, PRC (2019) Approval List of Safety Certificate for Agricultural Genetically Modified Organisms in 2018. Available at: http://www.moa.gov.cn/ztzl/zjyqwgz/spxx/201901/t20190108_6166293.htm (accessed 8 January 2021).

Mizuguti, A., Yoshimura, Y. and Matsuo, K. (2009) Flowering phenologies and natural hybridization of genetically modified and wild soybeans under field conditions. *Weed Biology and Management* 9, 93–96. DOI: 10.1111/j.1445-6664.2008.00324.x.

Mizuguti, A., Ohigashi, K., Yoshimura, Y., Kaga, A., Kuroda, Y. *et al*. (2010) Hybridization between GM soybean (*Glycine max* (l.) merr.) and wild soybean (*Glycine soja* sieb. et zucc.) under field conditions in Japan. *Environmental Biosafety Research* 9(1), 13–23. DOI: 10.1051/ebr/2010004.

Nakayama, Y. and Yamaguchi, H. (2010) Natural hybridization in wild soybean (*Glycine max* ssp. *soja*) by pollen flow from cultivated soybean (*Glycine max* ssp. *max*) in a designed population. *Weed Biology and Management* 2, 25–30. DOI: 10.1046/j.1445-6664.2002.00043.x.

Oka, H.I. (1983) Genetic control of regenerating success in semi-natural conditions observed among lines derived from a cultivated x wild soybean hybrid. *Journal of Applied Ecology* 20, 937–949. DOI: 10.2307/2403138.

Pandolfo, C.E., Presotto, A., Carbonell, F.T., Ureta, S., Poverene, M. *et al*. (2018) Transgene escape and persistence in an agroecosystem: the case of glyphosate-resistant *Brassica rapa* L. in Central Argentina. *Environmental Science and Pollution Research* 25(7), 6251–6264. DOI: 10.1007/s11356-017-0726-3.

Pellegrino, E., Bedini, S., Nuti, M. and Ercoli, L. (2018) Impact of genetically engineered maize on agronomic, environmental and toxicological traits: a meta-analysis of 21 years of field data. *Scientific Reports* 8(1), 6485. DOI: 10.1038/s41598-018-24840-y.

Piñeyro-Nelson, A., Heerwaarden, J.V., Perales, H.R., Serratos-Hernández, J.A. and Alvarez-Buylla, E.R. (2010) Transgenes in Mexican maize: molecular evidence and methodological considerations for GMO detection in landrace populations. *Molecular Ecology* 18, 750–761. DOI: 10.1111/j.1365-294X.2008.03993.x.

Prakash, C.S. (2001) The genetically modified crop debate in the context of agricultural evolution. *Plant Physiology* 126(1), 8–15. DOI: 10.1104/pp.126.1.8.

Quist, D. and Chapela, I.H. (2001) Transgenic DNA introgressed into traditional maize landraces in Oaxaca, Mexico. *Nature* 414(6863), 541–543. DOI: 10.1038/35107068.

Ray, J.D., Kilen, T.C., Abel, C.A. and Paris, R.L. (2003) Soybean natural cross-pollination rates under field conditions. *Environmental Biosafety Research* 2, 133–138. DOI: 10.1051/ebr:2003005.

Snow, A.A. and Allison, A. (2002) Transgenic crops: why gene flow matters. *Nature Biotechnology* 20(6), 542. DOI: 10.1038/nbt0602-542.

Snow, A.A., Andersen, B. and Jørgensen, R.B. (2010) Costs of transgenic herbicide resistance introgressed from *Brassica napus* into weedy *B. rapa*. *Molecular Ecology* 8, 60–615. DOI: 10.1046/j.1365-294x.1999.00596.x.

Song, Z.-P., Lu, B.-R., Zhu, Y.G. and Chen, J.-K. (2003) Gene flow from cultivated rice to the wild species *Oryza rufipogon* under experimental field conditions. *New Phytologist* 157(3), 657–665. DOI: 10.1046/j.1469-8137.2003.00699.x.

Song, Z.-P., Lu, B.-R., Wang, B. and Chen, J.-K. (2004) Fitness estimation through performance comparison of F_1 hybrids with their parental species *Oryza rufipogon* and *O. sativa*. *Annals of Botany* 93(3), 311–316. DOI: 10.1093/aob/mch036.

Stewart, C.N. Jr., Halfhill, M.D. and Warwick, S.I. (2003) Transgene introgression from genetically modified crops to their wild relatives. *Nature Reviews Genetics* 4(10), 806–817. DOI: 10.1038/nrg1179.

Wang, K.-J. and Li, X.-H. (2011) Interspecific gene flow and the origin of semi-wild soybean revealed by capturing the natural occurrence of introgression between wild and cultivated soybean populations. *Plant Breeding* 130, 117–127. DOI: 10.1111/j.1439-0523.2010.01815.x.

Wang, K.-J. and Li, X.-H. (2012) Genetic characterization and gene flow in different geographical-distance neighbouring natural populations of wild soybean (*Glycine soja* Sieb. & Zucc.) and implications for protection from GM soybeans. *Euphytica* 186, 817–830. DOI: 10.1007/s10681-012-0621-4.

Wang, K.-J. and Li, X.-H. (2013a) Genetic diversity and gene flow dynamics revealed in the rare mixed populations of wild soybean (*Glycine soja*) and semi-wild type (*Glycine gracilis*) in China. *Genetic Resources and Crop Evolution* 60, 2303–2318. DOI: 10.1007/s10722-013-9998-8.

Wang, K.-J. and Li, X.-H. (2013b) Pollen dispersal of cultivated soybean into wild soybean under natural conditions. *Crop Science* 53, 2497–2505. DOI: 10.2135/cropsci2012.07.0423.

Wang, K.-J., Li, X.-H., Zhang, J.-J., Chen, H., Zhang, Z.-L. *et al*. (2010) Natural introgression from cultivated soybean (*Glycine max*) into wild soybean (*Glycine soja*) with the implications for origin of populations of semi-wild type and for biosafety of wild species in China. *Genetic Resources and Crop Evolution* 57, 747–761. DOI: 10.1007/s10722-009-9513-4.

Wang, K.-J., Li, X.-H. and Liu, Y. (2012). Fine-scale phylogenetic structure and major events in the history of the current wild soybean (*Glycine soja*) and taxonomic assignment of semi-wild type (*Glycine gracilisSkvortz*.) within the Chinese subgenus *soja*. *Journal of Heredity* 103, 13–27. DOI: 10.1093/jhered/esr102.

Wang, K.-J., Li, X.-H. and Yan, M.-F. (2014) Microsatellite markers reveal genetic diversity of wild soybean in different habitats and implications for conservation strategies (*Glycine soja*) in China. *Conservation Genetics* 15, 605–618. DOI: 10.1007/s10592-014-0564-5.

Wang, Y.-H., Zhang, X.-J. and Fan, S.-J. (2015) Genetic diversity of wild soybean populations in Dongying, China, by simple sequence repeat analysis. *Genetics and Molecular Research* 14(3), 11613–11623. DOI: 10.4238/2015.September.28.13.

Wang, Z.-F. and Peng, S.-L. (2003) Plant hybridization and its harmful genetic consequences. *Biodiversity Science* 11, 333–339. [In Chinese with English abstract.] DOI: 10.3321/j.issn:1005-0094.2003.04.008.

Warwick, S.I., Simard, M.-J., Légère, A., Beckie, H.J., Braun, L. *et al*. (2003) Hybridization between transgenic *Brassica napus* L. and its wild relatives: *Brassica rapa* L., *Raphanus raphanistrum* L., *Sinapis arvensis* L., and *Erucastrum gallicum* (Willd.) O.E. Schulz. *Theoretical and Applied Genetics* 107(3), 528–539. DOI: 10.1007/s00122-003-1278-0.

Warwick, S.I., Légère, A., Simard, M.J. and James, T. (2010) Do escaped transgenes persist in nature? The case of an herbicide resistance transgene in a weedy *Brassica rapa* population. *Molecular Ecology* 17, 1387–1395. DOI: 10.1111/j.1365-294X.2007.03567.x.

Wegier, A., Piñeyro-Nelson, A., Alarcón, J., Gálvez-Mariscal, A., Alvarez-Buylla, E.R. *et al*. (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. *Molecular Ecology* 20(19), 4182–4194. DOI: 10.1111/j.1365-294X.2011.05258.x.

Yang, C.-F., Gao, C.-L., Qiao, F., An, Y.-X. and Kong, J.-J. (2014) Present situation and prospect of commercial research and development of herbicide-tolerant transgenic soybean. *Shaanxi Journal of Agricultural Sciences* 60, 53–55. [In Chinese with English abstract.] DOI: 10.3969/j.issn.0488-5368.2014.05.020.

Yang, X., Beres, Z.T., Jin, L., Parrish, J.T., Zhao, W. *et al*. (2017) Effects of over-expressing a native gene encoding 5-enolpyruvylshikimate-3-phosphate synthase (*epsps*) on glyphosate resistance in *Arabidopsis thaliana*. *PLoS ONE* 12(4), e0175820. DOI: 10.1371/journal.pone.0175820.

Yao, L. (2020) Current situation and development trend of China's soybean industry under Sino-US trade friction. *China Oils and Fats* 45, 10–14. [In Chinese with English abstract.] DOI: 10.12166/j.zgyz.1003-7969/2020.02.003.

Yook, M.-J., Park, H.-R., Zhang, C.-J., Lim, S.-H., Jeong, S.-C. *et al*. (2021) Environmental risk assessment of glufosinate-resistant soybean by pollen-mediated gene flow under field conditions in the region of the genetic origin. *Science of the Total Environment* 762, 143073. DOI: 10.1016/j.scitotenv.2020.143073.

Yoshimura, Y. (2011) Wind tunnel and field assessment of pollen dispersal in soybean [*Glycine max* (L.) Merr.]. *Journal of Plant Research* 124, 109–114. DOI: 10.1007/s10265-010-0357-y.

Yoshimura, Y., Matsuo, K. and Yasuda, K. (2006) Gene flow from GM glyphosate-tolerant to conventional soybeans under field conditions in Japan. *Environmental and Biosafety Research* 5(3), 169–173. DOI: 10.1051/ebr:2007003.

Zhao, T.-J., Ding, Y.-L. and Gai, J.-Y. (2008) Genetic diversity and ecological differentiation of Chinese annual wild soybean (*Glycine soja*). *Biodiversity Science* 16, 133–142. [In Chinese with English abstract.] DOI: 10.3724/SP.J.1003.2008.07286.

6 Monitoring Herbicide Resistance Gene Flow in Weed Populations

Hugh J. Beckie[1]* and Sara L. Martin[2]
[1]Australian Herbicide Resistance Initiative, School of Agriculture and Environment, The University of Western Australia, Perth, Western Australia, Australia; [2]Ottawa Research and Development Centre, Agriculture and Agri-Food Canada, Ottawa, Ontario, Canada

Abstract

Although herbicide-resistant (HR) weeds can be regularly monitored in fields via surveys, areawide monitoring of both cropland and ruderal (non-crop disturbed) areas is required for species with high propagule mobility. With increasing occurrence of HR weed populations in many agro-ecoregions, the relative contribution of independent evolution through herbicide selection and movement of HR alleles via pollen or seed needs to be elucidated to inform management and help preserve the remaining public good and common resource of herbicide susceptibility. Molecular markers available for many weed species can be utilized to assess regional gene flow accurately. In this chapter, we outline recommended principles and protocols for areawide monitoring of herbicide resistance gene flow in weed populations, exemplified by a case study of glyphosate resistance in kochia (*Bassia scoparia* A.J. Scott syn. *Kochia scoparia* (L.) Schrad.) in western Canada. Since being introduced from Eurasia to the Americas over a century ago, both seed- and pollen-mediated gene flow in the species have aided rapid range expansion and the spread of herbicide resistance.

Keywords: herbicide resistance; pollen-mediated gene flow; pollen movement; resistance spread; seed-mediated gene flow; weed seed dispersal

6.1 Introduction

Monitoring can be defined as the detection of changes and effects related to specific causes (Hellawell, 1991), or the systematic measurement of selected variables and processes that may be affected by a given practice (EFSA, 2006). For example, random field surveys to determine the nature and occurrence of herbicide-resistant (HR) weeds have been conducted regularly across western Canada and Western Australia during the past 20–25 years (Owen *et al*., 2007; Beckie *et al*., 2020). Monitoring can help discern temporal trends, quantify year-to-year stochasticity, or assess ecoregion or management influences on the subject of interest if periodically conducted using similar methodology. The magnitude and extent of HR gene flow in weeds may be inferred by how rapidly the incidence of an HR weed biotype increases across a region. However, determining the contribution of

*Corresponding author: hugh.beckie@uwa.edu.au

 Gene Flow: Monitoring, Modeling and Mitigation (eds Wei Wei and C. Neal Stewart, Jr)
DOI: 10.1079/9781789247480.0006

the spread of HR alleles via gene flow versus independent evolution through herbicide selection to resistance occurrence is difficult unless suitable molecular markers and data analyses are employed. The elements of HR weed surveys and areawide weed (including crop volunteers) monitoring programs have been detailed previously (Beckie *et al.*, 2000, 2010).

Independent evolution is often cited as the explanation for the incidence of HR weeds in a region. However, HR allele movement through propagules (pollen, seeds, or vegetative) may provide an underestimated contribution to the distribution of HR alleles across the landscape (Davis and Frisvold, 2017; Beckie *et al.*, 2019a). Given the potentially long waiting time for *de novo* mutations conferring HR (estimated at 1 in 10^5 to 10^6 gametes per locus per generation) (Crow, 1983) and given that the rate of gene flow among populations even in highly autogamous species is greater than the mutation rate, the appearance of resistance alleles can be greatly accelerated by gene flow among populations (Jasieniuk *et al.*, 1996; Barrett and Schluter, 2008). A weed population is often defined from an agronomic perspective as individuals of a weedy species within a field, as this is the largest scale where agronomic management practices and selection pressures are relatively homogeneous (Délye *et al.*, 2010). A strong role for the process of gene flow has been noted in several studies of HR weed populations. For example, Délye *et al.* (2010) and Busi *et al.* (2011) found that the genetic composition of black-grass (*Alopecurus myosuroides* Huds.) or annual ryegrass (*Lolium rigidum* Gaud.) populations in organic fields, including the presence of HR alleles, was the result of high pollen-mediated gene flow (PMGF) from conventional fields where resistance genes were present. Comparably, the similarity in the locus providing glyphosate resistance in Palmer amaranth (*Amaranthus palmeri* S. Wats.) indicated that the resistance alleles likely evolved once and then spread to geographically distant locations within the USA (Molin *et al.*, 2018). Further, long-distance gene flow (up to 9.6 km) has been documented in weedy beet (*Beta vulgaris* L.) in France in a study also indicating that a significant proportion of mating events occurred between individuals located several kilometers apart (Fénart *et al.*, 2007). These examples and others (reviewed by Hawkins *et al.*, 2019) suggest that gene flow contributes to the incidence of resistance at a similar or greater level than *de novo* evolution. Indeed, the introduction of HR alleles into previously herbicide-susceptible (HS) populations via gene flow from HR weed populations can be seen as a form of evolutionary rescue when populations are under strong selection pressure from herbicide application (Jasieniuk *et al.*, 1996; Orr and Unckless, 2014; Kreiner *et al.*, 2018). In these cases, rare long-distance dispersal events of HR genes can allow for the rapid long-distance spread of HR ecotypes. Long-distance dispersal events through pollen and seed have been quantified with varying degrees of precision. Developing our understanding of these factors can clarify targets for reducing the spread of HR alleles.

The extent of PMGF (percentage outcrossing) and potential distance that pollen can travel from an individual plant depends on the biology of the species, the environment, and the interplay between these factors. One key biological parameter is the length of time that pollen remains viable after shedding, which is correlated with PMGF levels seen in weed species. The pollen of waterhemp (*Amaranthus tuberculatus* L.) is non-viable after only 2 h (Liu *et al.*, 2012), resulting in a limited PMGF level compared with kochia (*Bassia scoparia* A.J. Scott syn. *Kochia scoparia* (L.) Schrad.), where pollen retains 10% viability even after 2 days (Friesen *et al.*, 2009). The length of time that pollen remains viable can be affected by environmental conditions, with high temperatures and low humidity reducing this period.

The phenology and abundance of the species can also play a role in the distance that pollen can travel successfully. Unlike crop species, which have been selected to mature in synchrony and are usually sown across a wide area within a short period, weed species often emerge throughout the season and can vary widely in growth stage. The variable phenology could limit the overlap in flowering time between cohorts or populations of determinant plant species (i.e. a set flowering period), while this is likely to be mitigated in indeterminate species that continue to flower until frost. The abundance and density of a weed species in an agro-ecoregion are positively correlated with rates of PMGF, with higher abundance and density leading to

higher rates of PMGF than at lower levels. Two environmental factors that shape PMGF at the landscape level are land-use patterns and the prevailing wind speed and direction. Specifically, the landscape can impose significant barriers to gene flow. For example, the canopy of a crop can provide a barrier to pollen movement (Murray *et al.*, 2002), while areas that are fallow or ruderal can provide routes enabling PMGF (Busi *et al.*, 2008). Similarly, strong winds that are predominantly from a particular direction can increase the frequency and distance of PMGF in the downwind direction (Beckie *et al.*, 2016). Despite the large variation possible in the interplay of these factors, an exponential decline is often observed between pollination (outcrossing) rate and distance from the pollen source. As a result, the majority of PMGF will typically occur within a few meters of the source, though the maximum distance of PMGF is less certain. When measured experimentally, this distance often corresponds to the geographic limits of the study area and is likely underestimated because of a strong dependence on sample size and the limits of detection.

Seed-mediated gene flow (SMGF) has the potential to allow HR gene movement across much larger distances than PMGF (Bagavathiannan and Norsworthy, 2013). While there are some reports of long-distance gene flow by pollen (e.g. Délye *et al.*, 2010), the majority of gene flow in most weedy species at long distances probably results from seed movement; PMGF is more important for short-distance HR gene movement (Jasieniuk *et al.*, 1996; Diggle and Neve, 2001). Similar to PMGF, the distance that seeds travel from their source populations is a result of the interplay between the biology of the species and the environment. Several factors, including the potential for anthropogenic (human-mediated) actions that facilitate spread (Benvenuti, 2007), contribute to increasing the scale over which these propagules can introduce genes. Biological factors contributing to SMGF include high fecundity, early seed shattering, and, perhaps most importantly, specialized adaptations for dispersal. For example, some weed species in the Asteraceae such as prickly lettuce (*Lactuca serriola* L.) are wind dispersed by a specialized feathery pappus that acts as a parachute, allowing seeds, and the HR alleles they carry, to spread rapidly and widely (Thill and Mallory-Smith, 1997; Shields *et al.*, 2006; Okada *et al.*, 2013). In our connected world, anthropogenic actions can spread seed through the movement of manure, feed, or commercial seed stocks, or as a result of seed loss during transport of crops or equipment (e.g. combine harvesters). This movement can occur over much greater distances (hundreds of kilometers) than natural seed dispersal (Stephenson *et al.*, 1990; Eberlein *et al.*, 1992; Anderson *et al.*, 1996; Norsworthy *et al.*, 2009; Aper *et al.*, 2012; Tafoya-Razo *et al.*, 2017). In this context, seeds that are difficult to remove from a crop or from crop machinery also have an advantage for the rapid and efficient dispersal of HR genes, as this characteristic increases anthropogenic dispersal (Darmency, 1996; Llewellyn and Allen, 2006; Benvenuti, 2007). While PMGF declines exponentially with distance from the pollen source, seed dispersal patterns are often less predictable and less restricted to specific regions or fields, making long-distance SMGF more difficult to measure or track through monitoring than PMGF (Dauer *et al.*, 2007).

In the next two sections, we briefly describe some basic principles and protocols for monitoring HR gene flow in weed populations, which are exemplified in several examples from the literature and a case study of glyphosate-resistant (GR) kochia. As detailed in the case study, the magnitude of gene flow in weed populations has implications for the rate of spread of HR alleles, population genetic diversity, and ability of the species to become locally adapted.

6.2 Principles and Protocols for Areawide Gene Flow Monitoring

Monitoring is most valuable during the early detection phase of resistance to a specific herbicide site of action (SOA). As incidence becomes widespread across a region, continued monitoring may be less justified. One of the first issues to address is the spatial scale of monitoring within a weed species' geographic range – possibly ranging from adjacent fields of an initially suspected or confirmed population, to county/municipal district level, to state/provincial level to an entire agro-region. Initial monitoring efforts usually focus on a relatively small area, which later

expands as more information on the incidence of resistance in the weed species is known. Even for weeds with low propagule dispersal capability, anthropogenic movement of seed can be in the order of hundreds of kilometers within a short time period (Beckie *et al.*, 2019a). Ultimately, the geographic area monitored is largely dictated by the perceived potential adverse economic impact of the HR weed biotype, and available human and financial resources for the monitoring program. Monitoring sites can include both agricultural fields and ruderal areas such as roadsides and railway rights of way where herbicides are applied (Bagavathiannan and Norsworthy, 2016). Random site selection is important to obtain an unbiased estimate of the incidence of resistance. However, the number of monitoring sites may be stratified or weighted by ecoregion or landscape, soil type, crop type or cultivated area, and weed species abundance or range limits (Beckie *et al.*, 2010, 2020). Depending upon the type of molecular marker used to assess gene flow, collected samples may comprise seeds or plant tissue. Protocols for HR weed survey sample collection and screening have been described previously (Beckie *et al.*, 2000; Burgos *et al.*, 2013). Experimental methods to study gene flow in weeds or crops were reviewed by Mallory-Smith *et al.* (2015).

Evidence of areawide HR gene movement via seed and pollen is becoming more accessible with the flourishing of molecular weed science. Increasing focus in this area has led to the development of resources for weed genomics and the study of HR evolution (Basu *et al.*, 2004; Ravet *et al.*, 2018). With these advances and reduction in sequencing costs, the techniques used are shifting from the use of a handful of markers such as amplified fragment length polymorphisms (AFLPs), microsatellites, and inter-simple sequence repeats (ISSRs), to those that provide a large number of markers. They often are single nucleotide polymorphisms (SNPs), or are spread across the genome, such as double-digest restriction enzyme-associated sequencing (Yang *et al.*, 2016) or whole genome resequencing (Ravet *et al.*, 2018). Numerous studies have demonstrated that different protocols and molecular or phenotypic markers can be successfully employed to understand how the patterns of gene flow shape the distribution of HR alleles across different environments, in different regions, and in weed species with different biology. These studies often have taken advantage of HR alleles in combination with other molecular markers that allow estimation of population genetic parameters. Herein, we briefly highlight four studies from different regions of the world, but many others are available in the literature (e.g. Fénart *et al.*, 2007; Aper *et al.*, 2012; Okada *et al.*, 2013; Küpper *et al.*, 2018; Molin *et al.*, 2018). In section 6.3, we examine glyphosate resistance in kochia as a case study in monitoring HR gene flow.

6.2.1 Acetyl-CoA carboxylase inhibitor resistance in French blackgrass (*Alopecurus myosuroides* Huds.)

Resistance to acetyl-CoA carboxylase (ACCase) inhibitors can occur following point mutations at several locations in the ACCase gene. This gene and AFLP markers were used to understand levels of gene flow and diversity in populations of blackgrass, an allogamous wind-pollinated species, in organic and conventional fields in Côte d'Or, France, encompassing 8800 km^2 (880,000 ha) (Délye *et al.*, 2010). Seed samples were randomly collected over multiple years from both types of field. The samples were screened for ACCase-inhibitor resistance using a classical pot assay. The frequencies of seven mutant HR ACCase alleles were assessed through genotyping and sequencing. This assessment determined that 74% of plants in organic fields and 80% of plants in conventional fields were HR. To confirm that HR ACCase alleles resulted from gene flow versus independent selection, a highly polymorphic 1257 bp fragment from the gene encoding ACCase enabled discrimination of distinct haplotypes corresponding to distinct occurrences of the allele. The AFLPs generated 116 markers, 102 of which were polymorphic. Analysis of these markers indicated no significant differences for the expected genetic diversity within populations (H_w = 0.2396 and 0.2406 for organic and conventional fields, respectively, p = 0.44), low overall genetic differentiation among populations (F_{ST} = 0.0234 and 0.0201, p = 0.94), and average pairwise genetic differentiation (F_{ST} pair = 0.0233 and 0.0200) (Holsinger and Weir,

2009). The authors thereby concluded that organic and conventional field populations had similar genetic diversity and that the prevalence of HR alleles in the organic fields was likely a result of extensive PMGF rather than persistence, seed dispersal (estimated at 4 m), or anthropogenic seed movement. Further, their models of PMGF for the species suggested that, given periodic demographic collapses of blackgrass populations in organic fields resulting from the cultivation of highly competitive fodder crops, pollen dispersal rate (probability of successful pollen migration) of only 10^{-3} from the larger and more stable populations in the conventional fields would be needed to produce the observed frequency of HR alleles in the organic fields. This study determined that HR gene movement occurred a maximum distance of 9 km over a 4-year period.

6.2.2 Acetyl-CoA carboxylase inhibitor resistance in Mexican wild oat (*Avena fatua* L.)

Herbicide-resistant ACCase alleles were also used, in combination with microsatellites, to investigate the spread of HR wild oat in northwestern Mexico (Tafoya-Razo *et al.*, 2017). Seeds were sampled from six populations, three from each of the two main agricultural zones where ACCase inhibitor resistance had become problematic. As in the study of blackgrass, these wild oat samples were screened for ACCase inhibitor resistance using a pot assay and the HR ACCase alleles were sequenced. The authors also sequenced nine microsatellite markers to assess population genetics parameters such as the number of alleles per locus; they used Migrate-N (Beerli, 2009) for estimating mean diversity and effective population size; and they conducted a spatial analysis of molecular variance (SAMOVA) to establish the relationship between geographic and genetic distances. They determined that all biotypes sampled were HR and that all shared the same mutation at codon 1781 in the ACCase allele. This result suggested that gene flow, not independent origins, was a more likely explanation for the distribution of HR alleles. The number of microsatellite markers per locus averaged 4.7 and ranged between 3.3 and 5.8. The estimated mean diversity of the populations was very similar for five of the six populations investigated, averaging 0.0965, and with estimated effective population sizes of between 4580 and 6270. However, both parameters were substantially lower for one population at 0.0087 and 3470, respectively, possibly due to a genetic bottleneck. The SAMOVA indicated that there were three genetic groups, comprising two populations from the Valle de Mexicali in Baja California, one strongly differentiated population in the Valle de Yaqui in Sonora, and one population from Valle de Mexicali as well as two from the Valle de Hermosillo in Sonora. This finding suggested that biotypes in the Valle de Hermosillo shared a genetic background with the population in Valle de Mexicali; the latter population was the likely origin of the resistance in the two Valle de Hermosillo populations, one of which showed the evidence of a bottleneck caused by founder effects. The authors concluded that the introduction of ACCase inhibitor resistance to the Valle de Hermosillo populations was the result of anthropogenic dispersal via the movement of contaminated wheat seedlots, but that the genetically differentiated population in the Valle de Yaqui could represent an independent origin for the mutation. This study determined that SMGF of the HR allele likely occurred over a distance of 1000 km.

6.2.3 Acetolactate synthase inhibitor resistance in South Australian prickly lettuce (*Lactuca serriola* L.)

Point mutations in the acetolactate synthase (ALS) gene can also provide genetic markers for ALS inhibitor resistance; they were combined with ISSRs to study the relative importance of independent selection events and HR seed dispersal for prickly lettuce, an autogamous Asteraceae species, in South Australia (Lu *et al.*, 2007). Samples were collected across a 350 km^2 area from roadside and cereal fields over 5 years. Prickly lettuce seedlings were screened for ALS inhibitor resistance and a portion of the ALS gene was sequenced. Further, the authors employed four ISSR primers that resulted in 179 bands. These data were used to investigate genetic relationships (dendrograms) using

unweighted pair group method with arithmetic mean (UPGMA) cluster analysis, and to determine genetic distances (e.g. Nei, 1978). The correlation between genetic and geographic distances between samples was examined using the Mantel test (Diniz-Filho *et al.*, 2013). Of the 69 populations studied, 22 were susceptible while the majority of the populations sampled had at least some resistant individuals, with four different mutations found within the ALS gene. The clusters of populations observed in the UPGMA contained individuals from different populations, some as far as 66 km apart. This lack of population structure was also indicated by Mantel tests that showed there was no correlation between the genotype and geographical distance ($r = 0.055$, $p = 0.10$). However, genetic variation increased between samples with increasing geographic distance when more distant populations were considered, suggesting population structure may exist at a larger scale. Correlation of clusters observed in the ISSR patterns with the polymorphisms observed in the ALS sequence supported gene flow as being a major contributor to the occurrence of herbicide resistance in clusters of approximately 5 km in radius. However, the presence of multiple ALS inhibitor resistance alleles suggested that several independent selection events had also occurred within the region. As the species is primarily self-pollinated, PMGF is expected to be relatively rare. The authors suggested that SMGF was likely responsible for most of the HR allele spread, since prickly lettuce seeds are small and adapted to wind dispersal via a large pappus and are difficult to clean from both seed and farm equipment. Based on identical genetic clusters of HR individuals, the authors concluded that the extent of HR spread was as far as 43 km.

6.2.4 Enolpyruvylshikimate-3-phosphate synthase inhibitor resistance in North American waterhemp

More recently, glyphosate resistance has been paired with whole-genome resequencing to investigate the role of gene flow and *de novo* evolution in populations of waterhemp, a dioecious and wind-pollinated species, in southern Ontario, Canada, and the Midwestern USA (Kreiner *et al.*, 2018). Glyphosate resistance in the species can be conferred by increased 5-enolp yruvylshikimate-3-phosphate synthase (EPSPS) gene copy number, which results in increased production of this enzyme normally inhibited in glyphosate-susceptible (GS) plants. Glyphosate resistance can also be conferred by a point mutation at codon 106 of the gene. Seeds were collected from natural populations and agricultural fields where chemical control was poor. Herbicide resistance was determined using a pot assay, the number of EPSPS copies was measured, and represented EPSPS sequences were investigated. This information was combined with data from 10,280,132 SNPs identified by whole-genome resequencing and the software 'freebayes' (Garrison and Marth, 2012). These SNPs were used in conjunction with a chromosome-level genome assembly to investigate population structure, complete demographic modeling, and detect selective sweeps around the amplified EPSPS locus. The populations showed significant geographic structure with clusters corresponding to the areas sampled, including distinct clusters for the natural and agricultural populations sampled in Ontario. Plants from Essex County, Ontario, showed genetic similarity to the westernmost population in Missouri, USA. These populations in Essex County were as genetically distinct from the other populations in Ontario as those located at the furthest geographic range of the study. The authors concluded that GR waterhemp in Essex County was likely introduced through SMGF, via either agricultural machinery or animal-mediated seed dispersal from Missouri. However, the analysis also indicated that glyphosate resistance in the Walpole, Ontario, populations represented an independent origin. These conclusions were supported by the evidence that the selective sweep around EPSPS was hard, resulting in a marked reduction in genetic diversity, increased genetic differentiation, and extended haplotype homozygosity in GR individuals from the Walpole populations, but not in individuals from Essex County or the Midwest USA, where the sweep was softer. Indeed, further investigation indicated a single evolutionary origin of EPSPS amplification at the Walpole location, while multiple independent origins among and within the Midwest USA populations contributed to glyphosate resistance in these populations.

Based on the similarity of waterhemp genotypes in Essex County and in Illinois, USA, this study indicates that HR allele movement via SMGF may have occurred over approximately 640 km. As shown in the above-mentioned cases of wild oat and prickly lettuce, independent origins also play a role in the distribution of HR alleles across the landscape.

6.3 Case Study: Glyphosate-resistant Kochia

Kochia is an annual weed species introduced from Eurasia into the Americas in the mid- to late 19th century as an ornamental for horticulture use (reviewed in Friesen *et al.*, 2009). During the past 50 years, kochia's range has expanded northward in the Northern Great Plains of Canada as well as southward into the High Plains of Texas, USA (Forcella, 1985; Beckie *et al.*, 2002). This C_4 species is drought tolerant and grows in abundance in both ruderal and cropping areas in western North America. Kochia's strong competitiveness in cropland and ruderal areas is the culmination of its drought, heat, and salinity tolerance, an ability to germinate and grow rapidly at cool soil temperatures and exert allelopathic effects, as well as its propensity for evolution of resistance to multiple herbicide SOAs (Friesen *et al.*, 2009). The evolution of resistance to multiple herbicide SOAs has compromised herbicidal control, thereby directly promoting the spread and persistence of the species. For example, most populations became resistant to ALS-inhibiting herbicides by the mid-2000s following the spread of HR alleles, as determined by surveys and random testing (Beckie *et al.*, 2011, 2019b; Varanasi *et al.*, 2015). The areawide frequency of HR biotypes of this species has progressed rapidly since it was first reported in the late 1980s (Heap, 2020).

The occurrence of resistance to glyphosate or synthetic auxin herbicides in kochia populations across the western Great Plains has also increased rapidly over the past decade. GR kochia was first reported in Kansas, USA, in 2007, and is now listed as occurring in ten US states and three Canadian provinces (Heap, 2020). As in waterhemp, glyphosate resistance can be conferred in kochia by increased gene copy number and overexpression of the EPSPS enzyme (Wiersma *et al.*, 2015). The EPSPS copy number correlates with the level of resistance, with four or more copies required for resistance to glyphosate when applied at label rates (Godar *et al.*, 2015; Varanasi *et al.*, 2015). Inheritance of glyphosate resistance follows a single locus Mendelian pattern in kochia, as the EPSPS gene copies are arranged in a tandem array on a single chromosome (Jugulam *et al.*, 2014).

In Canada, GR kochia was first confirmed in three close-proximity chemical fallow fields in one county in Alberta in 2011; later that year, 50 fields were surveyed within a 20 km radius of the three fields (Beckie *et al.*, 2013). Based on additional confirmed cases within that survey radius, a random survey of 309 kochia populations across southern Alberta (80,000 km^2) was conducted the following year. Five per cent of kochia populations in three counties were found to be resistant to glyphosate, but no populations were found to have individuals resistant to dicamba, an auxinic herbicide (Hall *et al.*, 2014). A second random survey of 305 kochia populations, conducted 5 years later in the same region, found that the number of GR populations had increased to 50% (15 counties) while 18% were now resistant to dicamba (nine counties) (Beckie *et al.*, 2019b). Moreover, 10% of the populations in the 2017 survey were resistant to herbicides of three different SOAs: ALS inhibitors, glycine, and synthetic auxins. Similarly, in southern Manitoba, Canada (75,000 km^2), incidence of GR kochia increased from 1% of 283 populations surveyed in 2013 to 59% of 297 populations in 2018 (Beckie *et al.*, 2015; Geddes *et al.*, 2019).

The fitness costs of the tandem duplication of the EPSPS gene in kochia were investigated using six segregating F_2 lines derived from some of the first Canadian populations to show resistance (Martin *et al.*, 2017). These lines were generated using controlled crosses between low and high EPSPS copy number (EPSPSCN) individuals from within the same populations collected in Alberta and Saskatchewan. These individuals were determined to be either high or low EPSPSCN by measuring the relative EPSPSCN compared with ALS copy number using quantitative PCR (qPCR) (Gaines *et al.*, 2010; Wiersma *et al.*, 2015). Segregating

F_2 lines were used so that the low and high EPSPSCN individuals that were compared had similar genetic backgrounds, an essential aspect for studying the fitness effects of a specific HR gene (Vila-Aiub *et al.*, 2009; Giacomini *et al.*, 2014). The majority of the F_2 families included in the study segregated for one, five and ten copies of the EPSPS gene. The 896 F_2 seedlings that germinated showed the expected pattern of EPSPSCN segregation and thus provided no evidence that EPSPSCN altered seed viability. While there were no significant differences in emergence time between high and low EPSPSCN individuals, emergence time from high EPSPSCN maternal lineages was delayed by 2.7 days in five of the six populations. This may be particularly detrimental in this species where competitiveness and persistence are partly attributed to its ability to germinate and emerge early. Plants from high EPSPSCN maternal lines were also significantly shorter, as indicated by growth models, and showed delayed flowering times and reduced seed production. High EPSPSCN individuals produced significantly fewer seeds by weight than those with low EPSPSCN. This finding suggested that high EPSPSCN individuals were passing on significant maternal effects to their offspring, which were fewer in number compared with low EPSPSCN individuals.

Therefore, high EPSPSCN individuals within kochia populations may be at a fitness disadvantage compared with wild-type individuals in the absence of glyphosate in some genetic backgrounds. This conclusion supports the hypothesis that GR alleles spread rapidly via PMGF and SMGF from a *de novo* event that occurred in the southern USA after the introduction of selection pressure from glyphosate, rather than rising to prominence from within the standing variation in multiple locations throughout kochia's range. The latter scenario would be more likely for a variant that was either neutral or beneficial before herbicide application changed the selective landscape (Barrett and Schluter, 2008; Orr and Unckless, 2014; Kreiner *et al.*, 2018; Hawkins *et al.*, 2019).

The extremely rapid range expansion of HR biotypes of kochia, potentially from a single source, strongly suggests high inter-population gene flow of HR alleles. This wind-pollinated species shows moderate levels of PMGF. For example, the rate of PMGF from GR to non-GR kochia was 0.25% at 96 m, the farthest distance tested (Beckie *et al.*, 2016). Kochia has a tumbling mechanism, which is an efficient and effective adaptation for the dispersal of HR weed seeds over short and long distances by wind (Stallings *et al.*, 1995a, b; Beckie *et al.*, 2016). It is estimated that an average kochia plant with tumbleweed architecture contains 100,000 viable, non-dormant seeds. These tumbleweeds provide a mechanism for rapid mass seed dispersal as they can travel 1 km in 5 min and drop 90% of the seeds during the journey (Beckie *et al.*, 2016).

To better understand the consequences of this moderate PMGF and efficient seed dispersal for the population genetics of kochia, a double-digest restriction enzyme-associated sequencing study was conducted using populations sampled from across western Canada (Martin *et al.*, 2020). The study also compared the genetic diversity of populations and of individuals with high EPSPSCN, with the expectation that strong selection pressure and the rapid spread of glyphosate resistance may have resulted in reduced standing genetic variation.

The materials used for this study were derived from bulk-collected seed from kochia populations tested during surveys to determine the extent of GR kochia in Alberta (Beckie *et al.*, 2013; Hall *et al.*, 2014), Saskatchewan, and Manitoba (Beckie *et al.*, 2015). Populations were identified as containing either only GS individuals among the 100 plants tested or at least some GR individuals (Fig. 6.1). Individuals were considered GR if they were not controlled by glyphosate applied at a discriminating dose of 900 g ae/ha (Beckie *et al.*, 2013). In each case, a GS population in as close proximity as possible to the GR populations was included in the study to limit the potential for shifts in SNP frequency due to local adaptation rather than the introduction of high EPSPSCN. Seeds from 12 individuals from each of 13 pairs of GS and GR populations from across the Canadian Prairies, a total of 312 individuals, were grown for DNA extraction. At the time of sampling, no individuals from a population designated as GS had increased EPSPSCN relative to ALS, while populations identified as GR were generally a mixture of individuals with and without increased EPSPSCN (Table 6.1). As a result, 89 high and 206 low EPSPSCN individuals were included in the study. Double-digested

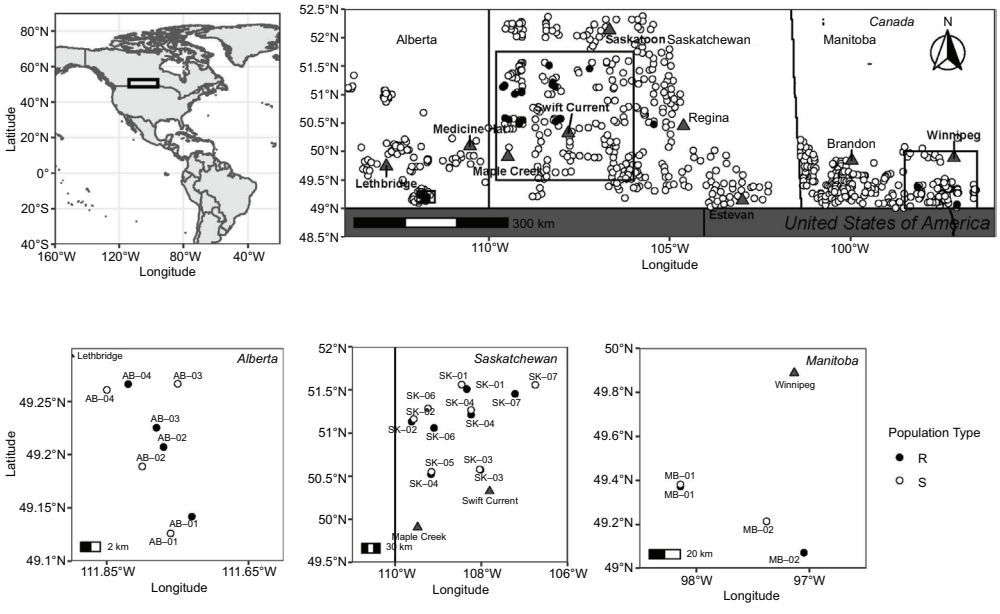


Fig. 6.1. Locations of the kochia populations in Alberta, Saskatchewan, and Manitoba, Canada, surveyed between 2011 and 2013 (Beckie *et al*., 2013, 2015) for glyphosate resistance and found to contain only susceptible (S) individuals (open circle) or at least some resistant individuals (R) (solid circle); greater detail is provided for sites used for experimental work (Martin *et al*., 2017, 2020) as shown along the bottom row.

restriction enzyme-associated marker library preparation and sequencing were completed following the 3DRAD protocol (Peterson *et al*., 2012) by the BadDNA laboratory, University of Georgia.

The final data analysis relied on the availability of kochia's draft genome (Patterson *et al*., 2019) and was primarily processed with 'Stacks' (Catchen *et al*., 2013) as well as population genetics packages such as 'StAMPP' (Pembleton *et al*., 2013), 'PopGenReports' (Adamack and Gruber, 2014) and 'poppr' (Kamvar *et al*., 2018) available for R (R Core Team, 2017). A total of 362 (95%) individuals had sufficient coverage and 3173 variable loci were identified, with 11.2% missing data.

The analysis confirms that pollen- and seed-mediated dispersal likely contribute high connectivity and gene flow among kochia populations (Table 6.2). The study found very low levels of genetic differentiation (overall F_{ST} = 0.01), individuals harbouring the majority of genetic variation (79.8%), an absence of structure by province, population or EPSPSCN indicated in principal components analyses (PCA), and in a lack of clustering in fineRAD-Structure analyses (Malinsky *et al*., 2018). For example, there was no evidence of isolation by distance, with no correlation between pairwise genetic and geographic distances among the 26 populations according to Mantel tests (Diniz-Filho *et al*., 2013) for either Provesti's genetic distance (r^2 = 0.11, p = 0.07) or F_{ST} genetic differentiation among populations (r^2 = –0.02, p = 0.57; Fig. 6.2). Populations of GR plants with similar genetic differentiation (F_{ST} = 0.052) include the obligate outcrosser Palmer amaranth in the USA (Küpper *et al*., 2018). There was less structure than found in previous work that used microsatellites for kochia populations in Minnesota and North Dakota, USA, where a moderate amount of population structure (G_{ST} = 0.09) was found (Mengistu and Messersmith, 2002). Population genetics parameters did not differ by population type or individual EPSPSCN. However, the species as a whole may be genetically depauperate compared with other weed species, regardless of EPSPSCN. Similarly,

Table 6.1. The average and range of enolpyruvylshikimate-3-phosphate synthase copy number (EPSPSCN) individuals found in the population, the total number of individuals with sufficient coverage for analysis (n_T) and the number of individuals with four or more copies of EPSPS relative to ALS (n_R) (adapted from Martin *et al*., 2020 and unpublished data).

	Susceptible populations		Resistant populations		
	EPSPSCN	n_T	EPSPSCN	n_T	n_R
AB-01	1.0 (0.7-1.5)	12	7.7 (0.9-19)	9	7
AB-02	1.2 (1.1-1.4)	12	6.3 (0.8-12.8)	12	8
AB-03	1.0 (0.8-1.2)	12	4.8 (0.9-16.2)	8	4
AB-04	1.3 (1.1-1.9)	11	6.7 (1-13.4)	12	7
MB-01	0.9 (0.8-1.1)	12	7.4 (1-21.9)	12	6
MB-02	1.2 (0.9-1.6)	10	9.0 (0.9-27.3)	9	5
SK-01	1.1 (0.8-1.8)	12	14.75 (5.6-22.3)	12	12
SK-02	1.1 (0.7-1.6)	12	8.0 (4-13.2)	10	10
SK-03	1.0 (0.7-1.8)	12	11.5 (1-30.3)	12	6
SK-04	1.1 (0.7-1.5)	12	5.4 (0.9-10.2)	10	6
SK-05	1.0 (0.9-1.3)	12	8.2 (0.8-18.5)	12	12
SK-06	0.8 (0.5-1.0)	12	21.7 (12.5-48.4)	12	6
SK-07	1.2 (1.0-1.5)	12	*1.0 (0.7-1.3*)	*12*	*0*[a]

[a]The initial survey screen of the population (n =100) indicated 10% resistant individuals (Beckie *et al*., 2015), but no individuals sampled by Martin *et al*. (2020) had increased EPSPSCN.

levels of inbreeding and shared alleles might be higher than expected for random mating in wind-pollinated species. Specifically, overall F_{IS} (inbreeding coefficient) was calculated as 0.23, or 23% higher than expected through random mating, and ranged from 0 to 0.42 within populations. Unfortunately, gene flow via seed versus pollen could not be distinguished, as no variation was found in regions aligning to the chloroplast.

The high gene flow among populations suggests that kochia across the Canadian Prairies could be considered as a single population and raises questions about the basis (genetic versus plasticity) of differences in emergence, flowering time, and leaf morphology seen between populations grown in a common environment (Martin *et al*., 2017). Interestingly, while a key feature of successful weed species invasions is often the ability to evolve locally adapted ecotypes (Barrett, 1982), high gene flow in kochia may strongly limit this ability (Lenormand, 2002). This study provides no indication that the spread of GR alleles will reduce kochia's genetic variability at the population level or its capacity to evolve additional herbicide resistance from standing variation. Future studies, able to take advantage of a chromosome level assembly of kochia, will be better able to examine the strength of the selective sweep of high EPSPSCN through these populations, clarify the number of origins of EPSPS amplification, and shed additional light on the potential consequences of this selection for kochia's adaptive potential (Przeworski, 2002; Nielsen *et al*., 2005; Messer and Petrov, 2013). It is clear, however, that beneficial alleles including those for herbicide resistance in environments with regular herbicide application can be expected to spread rapidly through this species' range as a result of high population connectivity. This prediction is concordant with rapid spread of ALS mutations (Beckie *et al*., 2011), glyphosate resistance (Beckie *et al*., 2015, 2019b; Geddes *et al*., 2019) and, more recently, auxinic resistance (Beckie *et al*., 2019b), as discussed previously. The evidence from the above studies indicates that the tremendous magnitude of recurrent, unrelenting gene flow of HR alleles will result in frequent, widespread multiple-HR trait stacking in kochia individuals and populations. As kochia is adapted to dry, hot conditions, our warming climate will promote range expansion and further increase

Table 6.2. Number of individuals with fewer (n_S) or more (n_R) than four copies of enolpyruvylshikimate-3-phosphate synthase (EPSPS) relative to ALS included in estimates of: allelic richness (A_R), observed heterozygosity (H_O), expected heterozygosity (H_E), the average proportion of polymorphic nucleotide sites within individuals by 10^{-3} (P_n), the average proportion of loci that showed variability within individuals (H_L), bootstrapped estimate of inbreeding coefficient (F_{IS})[a], genetic differentiation among populations (F_{ST}) (population-specific estimates; BayeScan), and geographic distance (km) between population pairs. (Adapted from Martin *et al*., 2020)

	Susceptible populations								Resistant populations									Distance between paired sites (km)
	n_S	A_R	H_O	H_E	P_n	H_L	F_{IS}	F_{ST}	n_S	n_R	A_R	H_O	H_E	P_n	H_L	F_{IS}	F_{ST}	
AB-01	12	1.27	0.18	0.28	2.43	0.11	0.36	0.02	2	7	1.26	0.21	0.27	2.59	0.13	0.28	0.01	2.76
AB-02	12	1.27	0.20	0.28	2.56	0.12	0.32	0.02	4	8	1.26	0.19	0.26	2.57	0.15	0.29	0.03	2.85
AB-03	12	1.26	0.18	0.26	2.45	0.12	0.33	0.02	4	4	1.26	0.17	0.28	2.21	0.14	0.42	0.01	4.95
AB-04	11	1.26	0.23	0.26	2.83	0.14	0.16	0.03	5	7	1.27	0.22	0.27	2.87	0.14	0.23	0.01	2.19
MB-01	12	1.26	0.22	0.26	2.74	0.13	0.16	0.04	6	6	1.27	0.17	0.28	2.28	0.11	0.25	0.02	1.04
MB-02	10	1.27	0.24	0.28	2.95	0.14	0.23	0.02	4	5	1.26	0.23	0.27	2.80	0.13	0.12	0.02	28.68
SK-01	12	1.27	0.23	0.26	2.96	0.14	0.17	0.02	0	12	1.27	0.22	0.27	2.75	0.17	0.18	0.01	10.28
SK-02	12	1.27	0.22	0.27	2.87	0.14	0.28	0.02	0	10	1.26	0.25	0.27	3.20	0.13	0.18	0.02	4.77
SK-03	12	1.27	0.23	0.27	3.01	0.14	0.06	0.01	6	6	1.26	0.23	0.27	2.88	0.14	0.39	0.02	2.03
SK-04	12	1.26	0.20	0.27	2.63	0.12	0.20	0.01	4	6	1.26	0.24	0.28	2.89	0.14	0.28	0.02	6.16
SK-05	12	1.27	0.28	0.27	3.44	0.17	0.17	0.02	0	12	1.26	0.16	0.26	2.33	0.11	0.01[a]	0.06	26.99
SK-06	12	1.26	0.22	0.26	2.90	0.14	0.25	0.03	6	6	1.27	0.21	0.27	2.67	0.13	0.4	0.03	3.3
SK-07	12	1.26	0.23	0.26	2.95	0.14	0.20	0.03	12	0[b]	1.26	0.27	0.26	3.51	0.17	0.19	0.02	34.46

[a]Bootstrapped confidence interval included zero.

[b]The initial survey screen of the population (n =100) indicated 10% resistant individuals, but no individuals sampled here had high EPSPSCN and the population pair was excluded from population level comparisons.

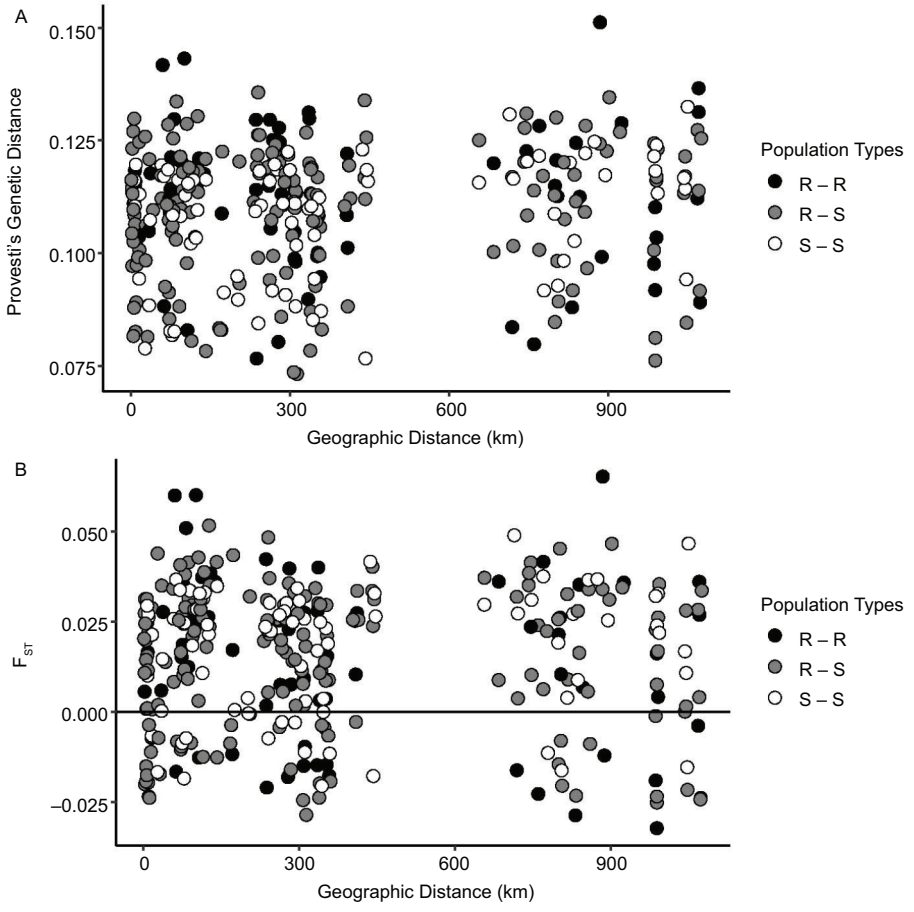


Fig. 6.2. Pairwise genetic and geographic distances between populations containing individuals with increased enolpyruvylshikimate-3-phosphate synthase (EPSPS) copy number based on EPSPS:acetolactate synthase (ALS) ratios (R – R); populations with increased EPSPS copy number and those that did not (R – S); and between population pairs with no evidence of increased EPSPS copy number for both (S – S). (A) Provesti's genetic distance and (B) F_{ST}, genetic differentiation among populations versus geographic distance, indicating little to no isolation by distance in this species. (Adapted from unpublished data and Martin *et al*., 2020)

population abundance and magnitude of gene flow (Beckie *et al*., 2002).

6.4 Monitoring to Inform Management

Over the past decade there have been repeated calls within the weed science community (see Llewellyn and Allen, 2006; Busi *et al*., 2008; Dauer *et al*., 2009; Sosnoskie *et al*., 2012; Okada *et al*., 2013; Délye *et al*., 2016; Ervin and Frisvold, 2016; Sarangi *et al*., 2017; Evans *et al*., 2018; Bagavathiannan *et al*., 2019, among others) for coordinated local and regional responses to reduce HR allele movement in weeds with high dispersal capability (i.e. areawide pest management) to avert a 'tragedy of the commons', i.e. overconsumption and depletion of the common-pool resource of herbicide susceptibility (Hardin, 1968). There are numerous socio-economic challenges to achieving the goal of reducing weed mobility across the landscape. One of the biggest challenges is controlling weed

populations in ruderal areas, because they are not usually the responsibility of the grower or land manager. For example, the collective failure to control HR kochia populations across the western Great Plains is largely because the problem is so daunting: given the vast agricultural and ruderal infestation area, coordination logistics are complex because of the numerous stakeholders involved, and there are no regulations or financial incentives for action.

Nevertheless, a few areawide weed management projects have been initiated. As an example, a 'zero tolerance' program that successfully tackled Palmer amaranth in the southern USA was established initially in two counties in Arkansas as a pilot project (Barber *et al*., 2015). A project, launched in 2019 in three regions of eastern Australia, focuses on flaxleaf fleabane (*Conyza bonariensis* L.) and other mobile weeds (Thyer, 2019). A prerequisite and foundation for mitigation or management is accurately assessing the scope of the problem, achieved through efficient and effective areawide monitoring of HR weed incidence and of HR gene flow among weed populations via seed or pollen dispersal.

References

Adamack, A.T. and Gruber, B. (2014) PopGenReport: simplifying basic population genetic analyses in R. *Methods in Ecology and Evolution* 5, 384–387.

Anderson, D.D., Roeth, F.W. and Martin, A.R. (1996) Occurrence and control of triazine-resistant common waterhemp (*Amaranthus rudis*) in field corn (*Zea mays*). *Weed Technology* 10(3), 570–575. DOI: 10.1017/S0890037X00040458.

Aper, J., Mechant, E., de Riek, J., Van Laere, K., Bulcke, R. *et al*. (2012) Analysis of local spread of metamitron-resistant *Chenopodium album* patches in Belgium. *Weed Research* 52(5), 421–429. DOI: 10.1111/j.1365-3180.2012.00928.x.

Bagavathiannan, M.V. and Norsworthy, J.K. (2013) Postdispersal loss of important arable weed seeds in the midsouthern United States. *Weed Science* 61(4), 570–579. DOI: 10.1614/WS-D-13-00065.1.

Bagavathiannan, M.V. and Norsworthy, J.K. (2016) Multiple-herbicide resistance is widespread in roadside Palmer amaranth populations. *Plos One* 11(4), e0148748. DOI: 10.1371/journal.pone.0148748.

Bagavathiannan, M.V., Graham, S., Ma, Z., Barney, J.N., Coutts, S.R. *et al*. (2019) Considering weed management as a social dilemma bridges individual and collective interests. *Nature Plants* 5(4), 343–351. DOI: 10.1038/s41477-019-0395-y.

Barber, L.T., Smith, K.L., Scott, R.C., Norsworthy, J.K. and Vangilder, A.M. (2015) *Zero Tolerance: A Community-Based Program for Glyphosate-Resistant Palmer Amaranth Management*. University of Arkansas Cooperative Extension Service Bulletin FSA2177, Fayetteville, Arkansas.

Barrett, S.C.H. (1982) Genetic variation in weeds. In: Charudattan, R.C. and Walker, H.L. (eds) *Biological Control of Weeds with Plant Pathogens*. John Wiley, New York, pp. 73–98.

Barrett, R.D.H. and Schluter, D. (2008) Adaptation from standing genetic variation. *Trends in Ecology & Evolution* 23(1), 38–44. DOI: 10.1016/j.tree.2007.09.008.

Basu, C., Halfhill, M.D., Mueller, T.C. and Stewart Jr, C.N. (2004) Weed genomics: new tools to understand weed biology. *Trends in Plant Science* 9(8), 391–398. DOI: 10.1016/j.tplants.2004.06.003.

Beckie, H.J., Heap, I.M., Smeda, R.J. and Hall, L.M. (2000) Screening for herbicide resistance in weeds. *Weed Technology* 14(2), 428–445. DOI: 10.1614/0890-037X(2000)014[0428:SFHRIW]2.0.CO;2.

Beckie, H.J., Weiss, R.M., Leeson, J.Y. and Olfert, O.O. (2002) Range expansion of kochia (*Kochia scoparia*) in North America under a changing climate. In: Ivany, J.A. and Blackshaw, R.E. (eds) *Topics in Canadian Weed Science, Vol. 8. Climate Change and the Canadian Agricultural Environment*. Canadian Weed Science Society, Pinawa, Manitoba, Canada, pp. 33–46.

Beckie, H.J., Hall, L.M., Simard, M.-J., Leeson, J.Y. and Willenborg, C.J. (2010) A framework for postrelease environmental monitoring of second-generation crops with novel traits. *Crop Science* 50(5), 1587–1604. DOI: 10.2135/cropsci2009.12.0743.

Beckie, H.J., Warwick, S.I., Sauder, C.A., Lozinski, C. and Shirriff, S. (2011) Occurrence and molecular characterization of acetolactate synthase (ALS) inhibitor–resistant kochia (*Kochia scoparia*) in Western Canada. *Weed Technology* 25(1), 170–175. DOI: 10.1614/WT-D-10-00067.1.

Beckie, H.J., Blackshaw, R.E., Low, R., Hall, L.M., Sauder, C.A. *et al*. (2013) Glyphosate- and acetolactate synthase inhibitor-resistant kochia (*Kochia scoparia*) in western Canada. *Weed Science* 61(2), 310–318. DOI: 10.1614/WS-D-12-00140.1.

Beckie, H.J., Gulden, R.H., Shaikh, N., Johnson, E.N., Willenborg, C.J. *et al*. (2015) Glyphosate-resistant kochia (*Kochia scoparia* L. Schrad.) in Saskatchewan and Manitoba. *Canadian Journal of Plant Science* 95(2), 345–349. DOI: 10.4141/cjps-2014-328.

Beckie, H.J., Blackshaw, R.E., Hall, L.M. and Johnson, E.N. (2016) Pollen- and seed-mediated gene flow in kochia (*Kochia scoparia*). *Weed Science* 64(4), 624–633. DOI: 10.1614/WS-D-16-00038.1.

Beckie, H.J., Busi, R., Bagavathiannan, M.V. and Martin, S.L. (2019a) Herbicide resistance gene flow in weeds: under-estimated and under-appreciated. *Agriculture, Ecosystems & Environment* 283(Sp Iss), e106566. DOI: 10.1016/j.agee.2019.06.005.

Beckie, H.J., Hall, L.M., Shirriff, S.W., Martin, E. and Leeson, J.Y. (2019b) Triple-resistant kochia [*Kochia scoparia* (L.) Schrad.] in Alberta. *Canadian Journal of Plant Science* 99(2), 281–285. DOI: 10.1139/cjps-2018-0256.

Beckie, H.J., Shirriff, S.W., Leeson, J.Y., Hall, L.M., Harker, K.N. *et al*. (2020) Herbicide-resistant weeds in the Canadian prairies: 2012 to 2017. *Weed Technology* 34(3), 461–474. DOI: 10.1017/wet.2019.128.

Beerli, P. (2009) How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use? In: Bertorelle, G., Bruford, M.W., Hauffe, H.C., Rizzoli, A. and Vernesi, C. (eds) *Conservation Biology, Vol. 17: Population Genetics for Animal Conservation*. Cambridge University Press, Cambridge, UK, pp. 42–79.

Benvenuti, S. (2007) Weed seed movement and dispersal strategies in the agricultural environment. *Weed Biology and Management* 7(3), 141–157. DOI: 10.1111/j.1445-6664.2007.00249.x.

Burgos, N.R., Tranel, P.J., Streibig, J.C., Davis, V.M., Shaner, D. *et al*. (2013) Confirmation of resistance to herbicides and evaluation of resistance levels. *Weed Science* 61(1), 4–20. DOI: 10.1614/WS-D-12-00032.1.

Busi, R.Y., Yu, Q., Barrett-Lennard, R. and Powles, S. (2008) Long distance pollen-mediated flow of herbicide resistance genes in *Lolium rigidum*. *Theoretical and Applied Genetics* 117(8), 1281–1290. DOI: 10.1007/s00122-008-0862-8.

Busi, R., Michel, S., Powles, S.B. and Délye, C. (2011) Gene flow increases the initial frequency of herbicide resistance alleles in unselected *Lolium rigidum* populations. *Agriculture, Ecosystems & Environment* 142(3-4), 403–409. DOI: 10.1016/j.agee.2011.06.012.

Catchen, J.M., Hohenlohe, P.A., Bassham, S., Amores, A. and Cresko, W.A. (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology* 22(11), 3124–3140. DOI: 10.1111/mec.12354.

Crow, J.F. (1983) *Genetics Notes: An Introduction to Genetics*, 8th edn. Burgess Press, Minneapolis, Minnesota.

Darmency, H. (1996) Movement of resistance genes among plants. In: Brown, T.M. (ed.) *Molecular Genetics and Evolution of Pesticide Resistance in Plants: Biology and Biochemistry*. Lewis Publishers, Boca Raton, Florida, pp. 209–220.

Dauer, J.T., Mortensen, D.A. and VanGessel, M.J. (2007) Temporal and spatial dynamics of long-distance *Conyza canadensis* seed dispersal. *Journal of Applied Ecology* 44(1), 105–114. DOI: 10.1111/j.1365-2664.2006.01256.x.

Dauer, J.T., Luschei, E.C. and Mortensen, D.A. (2009) Effects of landscape composition on spread of an herbicide-resistant weed. *Landscape Ecology* 24(6), 735–747. DOI: 10.1007/s10980-009-9345-9.

Davis, A.S. and Frisvold, G.B. (2017) Are herbicides a once in a century method of weed control? *Pest Management Science* 73(11), 2209–2220. DOI: 10.1002/ps.4643.

Délye, C., Clément, J.A.J., Pernin, F., Chauvel, B. and Le Corre, V. (2010) High gene flow promotes the genetic homogeneity of arable weed populations at the landscape level. *Basic and Applied Ecology* 11(6), 504–512. DOI: 10.1016/j.baae.2010.06.008.

Délye, C., Causse, R. and Michel, S. (2016) Genetic basis, evolutionary origin and spread of resistance to herbicides inhibiting acetolactate synthase in common groundsel (*Senecio vulgaris*). *Pest Management Science* 72(1), 89–102. DOI: 10.1002/ps.4058.

Diggle, A.J. and Neve, P. (2001) The population dynamics and genetics of herbicide resistance – a modeling approach. In: Powles, S.B. and Shaner, D.L. (eds) *Herbicide Resistance and World Grains*. CRC Press, New York, pp. 61–99.

Diniz-Filho, J.A.F., Soares, T.N., Lima, J.S., Dobrovolski, R., Landeiro, V.L. *et al*. (2013) Mantel test in population genetics. *Genetics and Molecular Biology* 36(4), 475–485. DOI: 10.1590/S1415-47572013000400002.

Eberlein, C.V., Al-Khatib, K., Guttieri, M.J. and Fuerst, E.P. (1992) Distribution and characteristics of triazine-resistant Powell amaranth (*Amaranthu powellii*) in Idaho. *Weed Science* 40(4), 507–512. DOI: 10.1017/S0043174500058045.

EFSA (2006) Guidance document of the Scientific Panel on Genetically Modified Organisms for the risk assessment of genetically modified plants and derived food and feed. *European Food Safety Authority Journal* 99, 1–100.

Ervin, D.E. and Frisvold, G.B. (2016) Community-based approaches to herbicide-resistant weed management: lessons from science and practice. *Weed Science* 64, 609–626. DOI: 10.1614/WS-D-15-00122.1.

Evans, J.A., Williams, A., Hager, A.G., Mirsky, S.B., Tranel, P.J. *et al*. (2018) Confronting herbicide resistance with cooperative management. *Pest Management Science* 74(11), 2424–2431. DOI: 10.1002/ps.5105.

Fénart, S., Austerlitz, F., Cuguen, J. and Arnaud, J.-F. (2007) Long distance pollen-mediated gene flow at a landscape level: the weed beet as a case study. *Molecular Ecology* 16(18), 3801–3813. DOI: 10.1111/j.1365-294X.2007.03448.x.

Forcella, F. (1985) Final distribution is related to rate of spread in alien weeds. *Weed Research* 25(3), 181–191. DOI: 10.1111/j.1365-3180.1985.tb00634.x.

Friesen, L.F., Beckie, H.J., Warwick, S.I. and Van Acker, R.C. (2009) The biology of Canadian weeds. 138. *Kochia scoparia* (L.) Schrad. *Canadian Journal of Plant Science* 89(1), 141–167. DOI: 10.4141/CJPS08057.

Gaines, T.A., Zhang, W., Wang, D., Bukun, B., Chisholm, S.T. *et al*. (2010) Gene amplification confers glyphosate resistance in *Amaranthus palmeri*. *Proceedings of the National Academy of Sciences of the United States of America* 107(3), 1029–1034. DOI: 10.1073/pnas.0906649107.

Garrison, E. and Marth, G. (2012) Haplotype-based variant detection from short-read sequencing. Available at: https://arxiv.org/abs/1207.3907 (accessed 1 July 2020).

Geddes, C.M., Ostendorf, T.E., Gulden, R.H., Jones, T. and Leeson, J.Y. (2019) Rapid spread of glyphosate-resistant kochia [*Bassia scoparia* (L.) A.J. Scott] in Manitoba. *Proceedings Canadian Weed Science Society*. Available at: http://www.weedscience.ca (accessed 1 June 2020)

Giacomini, D., Westra, P. and Ward, S.M. (2014) Impact of genetic background in fitness cost studies: an example from glyphosate-resistant Palmer amaranth. *Weed Science* 62(1), 29–37. DOI: 10.1614/WS-D-13-00066.1.

Godar, A.S., Stahlman, P.W., Jugulam, M. and Dille, J.A. (2015) Glyphosate-resistant kochia (*Kochia scoparia*) in Kansas: EPSPS gene copy number in relation to resistance levels. *Weed Science* 63(3), 587–595. DOI: 10.1614/WS-D-14-00145.1.

Hall, L.M., Beckie, H.J., Low, R., Shirriff, S.W., Blackshaw, R.E. *et al*. (2014) Survey of glyphosate-resistant kochia (*Kochia scoparia* L. Schrad.) in Alberta. *Canadian Journal of Plant Science* 94(1), 127–130. DOI: 10.4141/cjps2013-204.

Hardin, G. (1968) The tragedy of the commons. *Science* 162, 1243–1248. DOI: 10.1126/science.162.3859.1243.

Hawkins, N.J., Bass, C., Dixon, A. and Neve, P. (2019) The evolutionary origins of pesticide resistance. *Biological Reviews* 94(1), 135–155. DOI: 10.1111/brv.12440.

Heap, I.M. (2020) International Herbicide-Resistant Weed Database. Available at: http://www.weedscience.org (accessed 1 March 2020).

Hellawell, J.M. (1991) Development of a rationale for monitoring. In: Goldsmith, B. (ed.) *Monitoring for Conservation and Ecology*. Chapman and Hall, London, UK, pp. 1–14.

Holsinger, K.E. and Weir, B.S. (2009) Genetics in geographically structured populations: defining, estimating and interpreting FST. *Nature Reviews Genetics* 10(9), 639–650. DOI: 10.1038/nrg2611.

Jasieniuk, M., Brûlé-Babel, A.L. and Morrison, I.N. (1996) The evolution and genetics of herbicide resistance in weeds. *Weed Science* 44(1), 176–193. DOI: 10.1017/S0043174500093747.

Jugulam, M., Niehues, K., Godar, A.S., Koo, D.-H., Danilova, T. *et al*. (2014) Tandem amplification of a chromosomal segment harboring 5-enolpyruvylshikimate-3-phosphate synthase locus confers glyphosate resistance in *Kochia scoparia*. *Plant Physiology* 166(3), 1200–1207. DOI: 10.1104/pp.114.242826.

Kamvar, Z.N., Tabima, J.F. and Grünwald, N.J. (2018) Poppr : an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction v2.6.1. *PeerJ* 2, e281. DOI: 10.7717/peerj.281.

Kreiner, J.M., Stinchcombe, J.R. and Wright, S.I. (2018) Population genomics of herbicide resistance: adaptation via evolutionary rescue. *Annual Review of Plant Biology* 69, 611–635. DOI: 10.1146/annurev-arplant-042817-040038.

Küpper, A., Manmathan, H.K., Giacomini, D., Patterson, E.L., McCloskey, W.B. *et al*. (2018) Population genetic structure in glyphosate–resistant and –susceptible Palmer amaranth (*Amaranthus palmeri*) populations using genotyping-by-sequencing (GBS). *Frontiers in Plant Science* 9, e29. DOI: 10.3389/fpls.2018.00029.

Lenormand, T. (2002) Gene flow and the limits to natural selection. *Trends in Ecology & Evolution* 17(4), 183–189. DOI: 10.1016/S0169-5347(02)02497-7.

Liu, J., Davis, A.S. and Tranel, P.J. (2012) Pollen biology and dispersal dynamics in waterhemp (*Amaranthus tuberculatus)*. *Weed Science* 60(3), 416–422. DOI: 10.1614/WS-D-11-00201.1.

Llewellyn, R.S. and Allen, D.M. (2006) Expected mobility of herbicide resistance via weed seeds and pollen in a Western Australian cropping region. *Crop Protection* 25(6), 520–526. DOI: 10.1016/j.cropro.2005.08.007.

Lu, Y.-Q., Baker, J. and Preston, C. (2007) The spread of resistance to acetolactate synthase inhibiting herbicides in a wind borne, self-pollinated weed species, *Lactuca serriola* L. *Theoretical and Applied Genetics* 115(4), 443–450. DOI: 10.1007/s00122-007-0576-3.

Malinsky, M., Trucchi, E., Lawson, D.J. and Falush, D. (2018) RADpainter and fineRADstructure: population inference from RADseq data. *Molecular Biology and Evolution* 35(5), 1284–1290. DOI: 10.1093/molbev/msy023.

Mallory-Smith, C., Hall, L.M. and Burgos, N.R. (2015) Experimental methods to study gene flow. *Weed Science* 63(Sp. Iss.), 12–22. DOI: 10.1614/WS-D-13-00064.1.

Martin, S.L., Benedict, L., Sauder, C.A., Wei, W., da Costa, L.O. *et al*. (2017) Glyphosate resistance reduces kochia fitness: comparison of segregating resistant and susceptible F2 populations. *Plant Science* 261, 69–79. DOI: 10.1016/j.plantsci.2017.04.010.

Martin, S.L., Benedict, L., Wei, W., Sauder, C.A., Beckie, H.J. *et al*. (2020) High gene flow maintains genetic diversity following selection for high EPSPS copy number in kochia. *Scientific Reports* 10(1), 18864. DOI: 10.1038/s41598-020-75345-6.

Mengistu, L.W. and Messersmith, C.G. (2002) Genetic diversity of kochia. *Weed Science* 50(4), 498–503. DOI: 10.1614/0043-1745(2002)050[0498:GDOK]2.0.CO;2.

Messer, P.W. and Petrov, D.A. (2013) Population genomics of rapid adaptation by soft selective sweeps. *Trends in Ecology & Evolution* 28(11), 1–22. DOI: 10.1016/j.tree.2013.08.003.

Molin, W.T., Wright, A.A., VanGessel, M.J., McCloskey, W.B., Jugulam, M. *et al*. (2018) Survey of the genomic landscape surrounding the 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) gene in glyphosate-resistant *Amaranthus palmeri* from geographically distant populations in the USA. *Pest Management Science* 74(5), 1109–1117. DOI: 10.1002/ps.4659.

Murray, B.G., Morrison, I.N. and Friesen, L.F. (2002) Pollen-mediated gene flow in wild oat. *Weed Science* 50(3), 321–325. DOI: 10.1614/0043-1745(2002)050[0321:PMGFIW]2.0.CO;2.

Nei, M. (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89(3), 583–590. DOI: 10.1093/genetics/89.3.583.

Nielsen, R., Williamson, S., Kim, Y., Hubisz, M.J., Clark, A.G. *et al*. (2005) Genomic scans for selective sweeps using SNP data. *Genome Research* 15(11), 1566–1575. DOI: 10.1101/gr.4252305.

Norsworthy, J.K., Smith, K.L., Steckel, L.E. and Koger, C.H. (2009) Weed seed contamination of cotton gin trash. *Weed Technology* 23(4), 574–580. DOI: 10.1614/WT-08-146.1.

Okada, M., Hanson, B.D., Hembree, K.J., Peng, Y., Shrestha, A. *et al*. (2013) Evolution and spread of glyphosate resistance in *Conyza canadensis* in California. *Evolutionary Applications* 6(5), 761–777. DOI: 10.1111/eva.12061.

Orr, H.A. and Unckless, R.L. (2014) The population genetics of evolutionary rescue. *PLoS Genetics* 10(8), 1–9. DOI: 10.1371/journal.pgen.1004551.

Owen, M.J., Walsh, M.J., Llewellyn, R.S. and Powles, S.B. (2007) Widespread occurrence of multiple herbicide resistance in Western Australian annual ryegrass (*Lolium rigidum*) populations. *Australian Journal of Agricultural Research* 58(7), 711–718. DOI: 10.1071/AR06283.

Patterson, E.L., Saski, C.A., Sloan, D.B., Tranel, P.J., Westra, P. *et al*. (2019) The draft genome of *Kochia scoparia* and the mechanism of glyphosate resistance via transposon-mediated EPSPS tandem gene duplication. *Genome Biology and Evolution* 11(10), 2927–2940. DOI: 10.1093/gbe/evz198.

Pembleton, L.W., Cogan, N.O.I. and Forster, J.W. (2013) StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources* 13(5), 946–952. DOI: 10.1111/1755-0998.12129.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S. and Hoekstra, H.E. (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE* 7(5), 1–11. DOI: 10.1371/journal.pone.0037135.

Przeworski, M. (2002) The signature of positive selection at randomly chosen loci. *Genetics* 160(3), 1179–1189. DOI: 10.1093/genetics/160.3.1179.

R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.r-project.org (accessed 1 March 2020).

Ravet, K., Patterson, E.L., Krähmer, H., Hamouzová, K., Fan, L. *et al*. (2018) The power and potential of genomics in weed biology and management. *Pest Management Science* 74(10), 2216–2225. DOI: 10.1002/ps.5048.

Sarangi, D., Tyre, A.J., Patterson, E.L., Gaines, T.A., Irmak, S. *et al*. (2017) Pollen-mediated gene flow from glyphosate-resistant common waterhemp (*Amaranthus rudis* Sauer): consequences for the dispersal of resistance genes. *Scientific Reports* 7, e44913. DOI: 10.1038/srep44913.

Shields, E.J., Dauer, J.T., VanGessel, M.J. and Neumann, G. (2006) Horseweed (*Conyza canadensis*) seed collected in the planetary boundary layer. *Weed Science* 54(6), 1063–1067. DOI: 10.1614/WS-06-097R1.1.

Sosnoskie, L.M., Webster, T.M., Kichler, J.M., MacRae, A.W., Grey, T.L. *et al*. (2012) Pollen-mediated dispersal of glyphosate-resistance in Palmer amaranth under field conditions. *Weed Science* 60(3), 366–373. DOI: 10.1614/WS-D-11-00151.1.

Stallings, G.P., Thill, D.C., Mallory-Smith, C.A. and Lass, L.W. (1995a) Plant movement and seed dispersal of Russian thistle (*Salsola iberica*). *Weed Science* 43(1), 63–69. DOI: 10.1017/S0043174500080838.

Stallings, G.P., Thill, D.C., Mallory-Smith, C.A. and Shafii, B. (1995b) Pollen-mediated gene flow of sulfonylurea-resistant kochia (*Kochia scoparia*). *Weed Science* 43(1), 95–102. DOI: 10.1017/S0043174500080887.

Stephenson, G.R., Dykstra, M.D., McLaren, R.D. and Hamill, A.S. (1990) Agronomic practices influencing triazine-resistant weed distribution in Ontario. *Weed Technology* 4(1), 199–207. DOI: 10.1017/S0890037X00025227.

Tafoya-Razo, J.A., Núñez-Farfán, J. and Torres-García, J.R. (2017) Migration by seed dispersal of ACCase-inhibitor-resistant *Avena fatua* in north-western Mexico. *Pest Management Science* 73(1), 167–173. DOI: 10.1002/ps.4282.

Thill, D.C. and Mallory-Smith, C.A. (1997) The nature and consequence of weed spread in cropping systems. *Weed Science* 45(3), 337–342. DOI: 10.1017/S004317450009295X.

Thyer, R. (2019) Area-Wide Approach to Weed Management Investigated in New Project. Grains Research and Development Corporation, Canberra, Australia. Available at: https://groundcover.grdc.com.au/story/6425537/project-launched-to-test-broad-scale-weed-management-approach/ (accessed 1 March 2020).

Varanasi, V.K., Godar, A.S., Currie, R.S., Dille, A.J., Thompson, C.R. *et al*. (2015) Field-evolved resistance to four modes of action of herbicides in a single kochia (*Kochia scoparia* L. Schrad.) population. *Pest Management Science* 71(9), 1207–1212. DOI: 10.1002/ps.4034.

Vila-Aiub, M.M., Neve, P. and Powles, S.B. (2009) Fitness costs associated with evolved herbicide resistance alleles in plants. *The New Phytologist* 184(4), 751–767. DOI: 10.1111/j.1469-8137.2009.03055.x.

Wiersma, A.T., Gaines, T.A., Preston, C., Hamilton, J.P., Giacomini, D. *et al*. (2015) Gene amplification of 5-enol-pyruvylshikimate-3-phosphate synthase in glyphosate-resistant *Kochia scoparia*. *Planta* 241(2), 463–474. DOI: 10.1007/s00425-014-2197-9.

Yang, G.-Q., Chen, Y.-M., Wang, J.-P., Guo, C., Zhao, L. *et al*. (2016) Development of a universal and simplified ddRAD library preparation approach for SNP discovery and genotyping in angiosperm plants. *Plant Methods* 12(1), 1–17. DOI: 10.1186/s13007-016-0139-1.

7 Modeling Gene Flow from Genetically Modified Plants

Wei Wei[1]*, Jun-Ming Wang[2], Xiang-Cheng Mi[1], Yan-Da Li[3] and Yan-Ming Zhu[1]
[1]State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing, China; [2]Climate and Atmospheric Science Section, Illinois State Water Survey, Prairie Research Institute, University of Illinois at Urbana-Champaign, Champaign, IL, USA; [3]Institute of Agricultural Engineering, Jiangxi Academy of Agricultural Sciences, Nanchang City, Jiangxi Province, China

Abstract

Gene flow from genetically modified (GM) plants is concerning because of its ecological risks. In modeling studies, these risks may be reduced by altering crop management while taking environmental conditions into account. Gene flow modeling should consider many field aspects, both biological and physical. For example, empirical statistical models deduced from experimental data simulate gene flow well only under limited conditions (similar to experimental conditions). Mechanistic models, however, offer a potentially greater predictive ability. Gene flow models from GM crops to non-GM crops are used to simulate field conditions and minimize the adventitious presence of transgenes to meet certain threshold levels. These models can be adapted to simulate gene flow from GM crops to crop wild relatives using parameters of sexual compatibility and growth characteristics of the wild plants. Currently, modeling gene flow from herbicide-resistant weeds has become very important in light of the increased application of herbicides and widely evolved resistance in weeds.

Keywords: airborne pollen; empirical model; herbicide resistance; hybridization; mechanistic model; mitigation; particles diffusion; pollen dispersal; seed spreading; simulation; transgene escape

7.1 Introduction

With the expanded releases of genetically modified (GM) crops, potential risks to the environment have raised concerns for human health and environmental safety, among which, gene flow from GM crops and its related consequences have become hot topics of study, especially in the crops' centers of origin (Hokanson *et al*., 2018). Gene flow from crops may result in the adventitious presence (AP) of transgenes or their products in conventional crops when GM and non-GM crops coexist in an agricultural system and may also cause transgene escape into wild relative plants occurring in sympatric distributions (Dong *et al*., 2018). Crop wild relatives are resources for crop improvement, thus the gene flow from crops to wild relatives has become a focus of study for risk assessment (Eastham and Sweet, 2002; Liu

*Corresponding author: weiwei@ibcas.ac.cn

 Gene Flow: Monitoring, Modeling and Mitigation (eds Wei Wei and C. Neal Stewart, Jr)
DOI: 10.1079/9781789247480.0007

et al., 2013; Ellstrand, 2014). Gene flow from GM crops to non-GM crops and wild relatives could happen via transfer of pollen, seeds, and vegetative propagules, potentially affecting biodiversity conservation (Lal *et al*., 2020).

With the expanded release area of GM crops, the AP of GM components from gene flow may affect the coexistence between GM and conventional crops. In the European Union (EU), 0.9% of the AP of authorized genetically modified organisms (GMOs) in non-GM crop products is allowable. Reforming the GMO regulation system (for example, changing the threshold of AP and isolation distance) was proposed regarding the application of new biotechnology (e.g. gene editing) in crop breeding (Eriksson *et al*., 2020) along with the coincidental increased market share of non-GM food products in both Europe and the USA (Castellari *et al*., 2018). In the USA, the adoption rate of GM crops ranks first in the world; however, it seems that only a few GM products (e.g. maize) could enter the food supply chain (USDA, 2015). Regulation in the USA appears to be less strict than that in the EU, and the AP of GM components does not seem likely to prevent certified organic farmers (or processors) from selling their products as USDA organic. Yet many organic producers support the low tolerance of the AP standard (Huffman and Strzok, 2017).

Models have been developed to simulate gene flow in coexisting farmlands to meet the standard of the AP level in the food supply chain, especially in EU countries. Employing those gene flow models, studies have been conducted and measures have been analyzed to develop strategies for controlling the risk of gene flow to reduce the opportunity of transgene escape. There are many reasons for these efforts, including improving field management to mitigate the impact on nature and contributing to decision making regarding environmental release and commercialization of GM crops to achieve perceived increases in sustainable agriculture (e.g. Meillet *et al*., 2015; Ridley and Alexander, 2016). Detecting and monitoring the presence of transgenes in nature may help in estimating levels and the strength of gene flow, while modeling of gene flow will contribute to predicting gene flow quantitatively, with opportunities to decrease it by altering certain parameters and introducing appropriate interventions.

When the crop wild relative receiving the transgene via gene flow is a resource plant, issues could emerge in its future application in crop breeding. When the crop wild relative is a weed itself, the introgression of an advantageous transgene could enhance the performance of the individual plants in nature and result in weed control difficulties. Even the introgressed individual or population could spread transgenes to other plants of the same or other related species. In addition, the worldwide increased application of herbicides has likely contributed to a resistance evolution in weeds. The USA likely has the most serious problem of herbicide-resistant weeds in the world (Heap, 2021).

This chapter focuses on modeling studies of gene flow via pollen and seeds, from GM crops to non-GM crops, from GM crops to their wild relatives, and from resistant weeds to susceptible populations of the same or related species. The chapter aims to review past progress in gene flow modeling studies and provides a reference for further research and for gene flow risk management.

7.2 Gene Flow Modeling Based on the Regression Analysis of Pollen Dispersal

In the middle of the last century, scientists started to assess gene flow in crops by empirical modeling (e.g. Bateman, 1947a, b). These studies focused mainly on the distance of pollen dispersal, considering wind, pollen movement, isolation, and other factors to eliminate contamination in seed crops. Bateman (1947c) provided three exponential regression equations to model gene flow mediated by pollen (wind and insect pollination) but combined them into one (Eqn. 7.1) as he assumed that the contamination rate might be very low in the field:

$$F = ye^{-KD}/D \qquad \text{(Eqn. 7.1)}$$

where F = proportion of contamination, D = distance, y = contamination rate at zero isolation distance and k = rate of decrease of contamination with distance.

In general, the pollen-mediated gene flow rate is assumed to decrease with increased distance fitting an exponential function, so an

isolation distance was normally recommended to avoid contaminated crop seeds. However, there are other factors in the environment that also affect gene flow levels, such as wind, temperature, and humidity. In addition, the scale of the release area (pollen source) will play an important role in pollen dispersal mediated by either wind or insects. For a release area of 70 m^2 of GM oilseed rape (OSR) (*Brassica napus*), the gene flow rate was reduced to 0.00033% at 47 m from the GM field, while the rate was still more than 10 times higher at 200 m when the pollen source increased to 400 m^2 (Reviewed in Wei *et al.*, 1999). For an even larger area (100 ha) of GM OSR, the pollen could disperse to more than 3000 m from the pollen source (Rieger *et al.*, 2002). For another insect-pollinated outcrossing crop, alfalfa (*Medicago sativa*), the transgene flow frequency for large fields was nearly 10 times greater than that of research-sized plots, and a minimum isolation distance of 1557 m was suggested to prevent gene flow in alfalfa (Amand *et al.*, 2000). With the global adoption of GM crops, however, it could be difficult to predict the dispersal of GM pollen in fields. In 2019, the total global area of GM crops was more than 190 million hectares (including cotton *Gossypium hirsutum*, soybean *Glycine max*, maize *Zea mays*, OSR, alfalfa and many other crops) (ISAAA, 2019). By monitoring gene flow from GM alfalfa fields to conventional alfalfa fields and modeling the relationship between the adventitious presence (AP) proportion and distance to GM source fields (Eqn. 7.2), with the threshold permission of 0.9% AP, the minimum isolation distance was proposed as 330 m (Kesoju *et al.*, 2019). However, a 0.03% level of AP still existed at an average distance of 8800 m from the GM fields.

$$\text{logit}(p) = -3.775 - 0.0034[\log(\text{distance})] - 0.0798[\log(\text{distance})]^2 \quad \text{(Eqn. 7.2)}$$

where logit(p) is the logit of the AP proportion and log(distance) is the log of distance from the nearest GM field (Kesoju *et al.*, 2019).

Cultivated rice (*Oryza sativa*) is a wind-pollinated and selfing crop. Rong *et al.* (2010) also used an exponential model to simulate rice pollen dispersal and revealed a pollen source size effect. However, they suggested that the size effect leveled off if the pollen source size was large enough. They stated in their study that a pollen source of 1 ha enabled the gene flow frequency to approach the upper limit, but Hu *et al.* (2020) reported that rice pollen deposition no longer increased once the source area exceeded 100 ha. Rong *et al.* (2012) argued that even with the increase in pollen source scale, if pollen travels a longer time to reach the recipients at a longer distance from the increased pollen sources size, the pollen may lose viability and have to compete with pollen of local plants.

The above-mentioned pollen source size effect was generally based on only one pollen source. However, in a smallholder farming system in some countries (e.g. in China), the size of some small farmlands may be less than 0.005 ha (50 m^2) (Fig. 7.1), which combined may form a larger complex farmland landscape that could be more than 100 ha (10^6 m^2) in size. Once a GM crop is commercialized, the owners of each small farmland can decide which crop they like to plant. Thus plenty of pollen sources could exist in the fields across the landscape, producing ample pollen to flow to their neighbors' adjacent non-GM crop fields. Consequently, managing a certain level of AP of GM component(s) in farmers' non-GM crop products could be very challenging. Meillet *et al.* (2015) developed a visualizing tool using Bayesian models to support farmers in choosing crops to plant regarding the expected AP level of GM components and the crops planted in adjacent fields. This required a minimal amount of input data and information on neighbors' field planning in advance.

7.3 Gene Flow Modeling Based on Elements of the Transgene Dispersal Process

7.3.1 Modeling gene flow mediated by pollen

In general, many gene flow models based on regression analysis that require measuring pollen-mediated gene flow frequency in the recipient plants are focused on the movement of pollen grains near the ground at a height of no more than 5 m for most food crops, depending on the plant height (Aylor *et al.*, 2003). It has been suggested that the pollen of some plants might

Fig. 7.1. A field map of small farmlands (607 individual farmlands; the largest one is less than 0.15 ha) in Qinghuapu Village, Ningxiang County in Hunan Province, China, produced by CAD (computer-aided design) (1 km × 1 km). Each small enclosed irregular grid represents a small farmland, and each of the long open strips represents a small path in the field.

have a greater concentration at a higher vertical level; for example, the mean count of pollen grains of *Populus* plants is greater at a height of 55 m than at 10 m (Bryant *et al.*, 1989). Huang *et al.* (2015) reported that the pollen of a small herbaceous plant could be found at heights of 80–100 m above the source plants and might result in long-distance dispersal. For food crops, pollen grain concentrations could not be very large at a higher vertical level. For example, maize pollen concentrations decreased nearly 100-fold between heights at 1 m above the tassels and 60 m above the crop (Aylor *et al.*, 2003). In a wheat (*Triticum aestivum*) field in the Huanghuai River region in China, a dynamic vertical distribution pattern was reported for wheat pollen: the highest accumulative density of pollen was at a height of 1 m at the pollen source, while at 185 m from the pollen source the highest density was detected at a height of 3 m (Sun *et al.*, 2015). Pollen dispersal and distribution patterns might be affected by complex meteorological conditions, especially wind speed, wind direction, and climate variability. Because of the vertically distributed airborne pollen and its sensitivity to climate and environmental conditions (Wang and Yang, 2010), scientists were able to develop mechanistic gene flow models that were originally established to simulate air pollution and particle diffusion in the atmosphere by integrating physical and biological characteristics of pollen and the dispersal process, including Gaussian plume models, gradient diffusion models, and Lagrangian random-flight models (Lu *et al.*, 2019). These mechanistic models could be established only

with known data from the literature or by measuring *in situ*, or both, without actual data from field experiments, but field experimental data were still needed for validation.

A gene flow model based on a Lagrangian model for GM maize was developed and validated by Wang and Yang (2010) and could also be applied for other wind-pollinated crops. This model consists of four major sub-models: pollen release, dispersion, deposition, and final outcrossing. By inputting species, field dimension, and weather data, the model predicts the dynamic three-dimensional pollen concentrations and two-dimensional depositions and the final two-dimensional outcrossing at the downwind female plants. This works well with local pollen dispersal (< 1 km). For long-distance dispersal (1–100 km), the 'Hybrid Single-Particle Lagrangian Integrated Trajectory' (HYSPLIT) model can be employed to predict puff advection for various particle sizes (including pollen size) (Wang *et al.*, 2011). The HYSPLIT model could account for the effects of plant particle attributes in long-distance dispersal, such as the initial release height and the falling velocity of particles (Liu *et al.*, 2018).

By replacing the exponential function in Rong *et al.* (2010) with the adjusted inverse Gaussian function, Wang *et al.* (2016) established a new pollen-mediated gene flow model and found a high level of goodness-of-fit by validating their model with other reported experimental data on crop-to-crop gene flow (in maize, OSR, and wheat) and crop-to-wild-relatives (for OSR and rice) gene flow frequencies. Based on this quasi-mechanistic pollen-mediated gene flow model for wind-pollination species, Wang and Lu (2017) further developed a tool to calculate gene flow frequency and isolation distance using biological and wind speed parameters that can be measured in the field, greenhouse or lab, or obtained from published data (Fig. 7.2). This tool is easy to download, install, and operate; it is released in the public domain and only requires simple training to input collected data.

Compared with the gene flow model of wind-pollination species, which is straightforward and uses assumptions of airborne particulate dispersal, modeling the pollination process of insects is much more complex (Ramsay, 2005). Vallaeys *et al.* (2017) suggested that the current models for pollen dispersal by insect pollinators rely on insect Brownian motion, an assumption that likely underestimates GM pollen outcrossing in conventional crops. Thus, they developed a fractional-order diffusion model based on Lévy flights of pollinating insects to provide a much more accurate prediction of pollen dispersal. To better model gene flow at the landscape scale, Ramsay (2005) proposed to break down the whole process into several discrete components of insect-mediated pollen dispersal, including

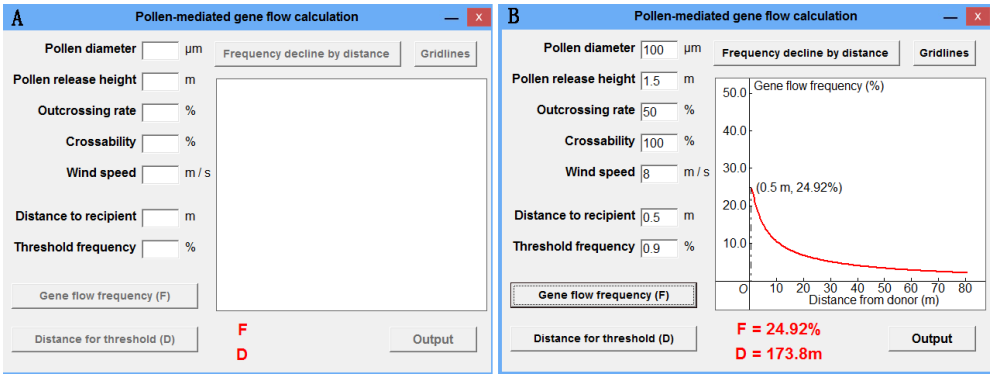


Fig. 7.2. Calculating gene flow frequency in maize crops by the tool developed by Wang and Lu (2017). (A) The blank screen interface of the tool. (B) Calculated results for maize gene flow in China. The data were adopted from Wang *et al.* (2016) at a moderate wind speed (8 m/s) in the maize growing season (Zhang *et al.*, 2020). The simulation results show that the gene flow frequency (F) is 24.92% at 0.5 m and declines with distance. A minimum isolation distance of 173.8 m is requested to keep an AP threshold of 0.9% at these given circumstances.

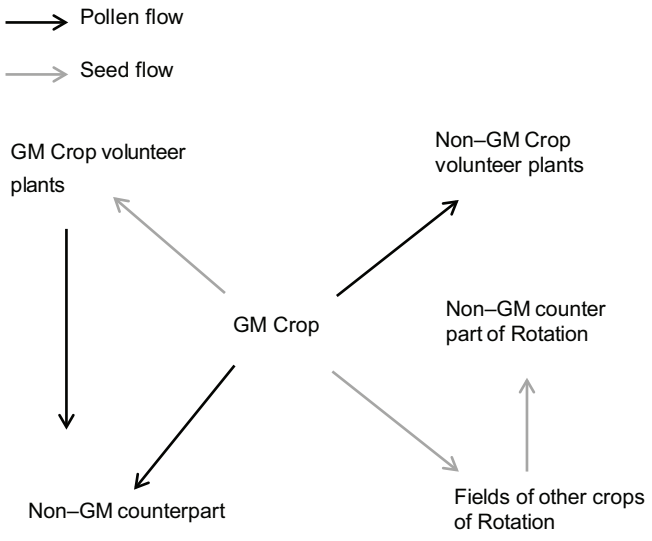


Fig. 7.3. Gene flow in agricultural fields.

elements and events at local-scale, regional-scale and landscape-scale levels.

7.3.2 Modeling gene flow with cropping systems data

In the field, gene flow may be mediated by both pollen flow and seed flow. The former could be mediated by wind or insects or both, whereas the latter might be caused by shattering before harvest and seed loss after harvest. The escaped seeds may survive along roadsides or grow in other crop fields in the following growth season and serve as a new pollen source (Fig. 7.3). Beckie *et al*. (2019) opined that seed dispersal has the potential to impact gene flow for weeds on a much larger scale than pollen flow, while Colbach (2008) assumed that seed-mediated gene flow might be a larger contributor to AP than pollen-mediated gene flow in larger crop fields. Colbach *et al*. (2005) suggested that all relevant factors should be accounted for, including cropping systems and genetic characteristics as well as pollen and seed dispersal. The gene flow model GeneSys was developed by Colbach *et al*. (2001a, b), who included all known pertinent factors in their model. This model was developed to simulate the gene flow of OSR pollinated by wind and insects and might be useful for developing a similar model for other crops, e.g. sugar beet (Colbach *et al*., 2001b). However, the GeneSys model was not useful for maize, which normally does not have volunteer plants in the field (Colbach *et al*., 2005).

The GeneSys model was originally developed based on field conditions in regions of France and other parts of Europe, and the parameters need to be revised and readjusted under new environmental circumstances outside Europe. Some attempts were made to adapt the GeneSys model to Chinese cropping systems, of which two case studies are given below.

7.3.2.1 Case study 1. Soil seed movement by tillage

The Yangtze River Basin region is the main production area for winter OSR in China. Crop rotation is favorable in this area; for example, in a typical cropping system, rice is planted after the OSR harvest in early May, and OSR will be planted after harvesting rice in September. During this rotation process, plowing would be conducted before planting each of the two crops, a paddy-rice–upland-crop rotation system that is one of the most popular in the Yangtze River Basin region.

Table 7.1. Evaluation of mungbean seeds movement model performance by plowing and analyzing the prediction accuracy of lateral displacement and final vertical coordinates in soils in Jinjing and Jinghuapu, Hunan Province, China. Adapted from Zhu (2012) with modifications.

Location	Plowing depth (cm)	Plow width (cm)	Number of points	Evaluated output variable	Mean of residuals (cm)	Modeling efficiency (r^2)	Mean Squared Error of Prediction (MSEP, cm)
Jijing	14–20	29	41	Lateral displacement	1.232	0.778	1.015
				Final vertical coordinate	0.779	0.648	0.268
Jinghuapu	8–12	26	35	Lateral displacement	1.360	0.450	10.985
				Final vertical coordinate	–0.873	0.342	5.548

Seedling emergence depends on the vertical distribution of plant seeds (Gruber *et al*., 2010), while plowing influences seed movement and distribution in soil (Colbach *et al*., 2000). A submodel (Roger-Estrade *et al*., 2000) was tested to estimate vertical and lateral seed displacement during plowing in a field trial in Jinghuapu and Jinjing, Hunan Province (Zhu, 2012). Mungbeans purchased from supermarkets were painted with six colors and buried in six layers at different depths and lateral positions. This study simulated the final vertical and lateral seed displacement during soil tillage as a function of soil structure, plowing width and depth, and initial seed position. Modeling efficiency (r^2) and the mean squared error of prediction (MSEP) were employed to evaluate model performance as suggested by Colbach *et al*. (2000). The results showed that the modeling efficiency in Jinghuapu was only slightly higher than 0.3 with a large mean squared error of prediction, while the r^2 reached a value as high as 0.778 in Jinjing with quite a small mean squared error of prediction (Table 7.1). Although the modeling efficiency in Jinjing was not high enough for a good performance, it showed better fit than that in Jinghuapu. The results suggested that deeper plowing (15 cm on average) might result in better performance of the model. Further work might be needed to adjust the model parameter to adapt to Chinese cropping systems. In this area, soil compaction might have been increased due to long-term cultivation of paddy rice (Yi *et al*., 2020).

7.3.2.2 Case study 2. Seed loss at harvest and persistence in soil (Zhu et al., 2012)

Large harvest seed losses in OSR result in a persistent volunteer seedbank in the soil and thus contribute significantly to gene flow via seeds. In China, harvesting is usually done by hand, and the recommended standard for harvesting time is when two-thirds of seeds become dark brown. At the beginning of harvest, OSR plant stalks are usually cut directly and are collected into bundles. These bundles of stalks are then dried in direct sunlight for 2–4 days, followed by seed shedding in the field. Following this process, studies were established to investigate the seed loss during harvest. Findings showed that seed losses were greatest during the sun-drying of harvested plants in the field, accounting for about three-quarters of the total seed losses. Seed losses during harvest could be reduced from 3.6–7.6% to 0.7–1.1% by weight of total seed production by placing a plastic membrane under the cut plants during drying.

A significant amount of seeds adds to the seedbank, thus it is necessary to check the fate of escaped seeds in soil. The subsequent crop following OSR is usually rice or cotton planted during late May and harvested in September or October in the Yangtze River Basin region. Seed burial assays showed that the proportion of germinating seeds relative to the initially buried seeds was significantly higher in the unirrigated versus irrigated fields and decreased

significantly with burial depth in the unirrigated fields but not in the irrigated field. No volunteer plants germinated in the irrigated field after 3 months in waterlogged conditions. Two years of field volunteer plant surveys showed that no OSR volunteers could be found in the flood-irrigated fields, while in the unirrigated fields significantly more OSR volunteers were found, from 0.75 ± 0.19 to 2.59 ± 0.47 plants/m^2 in different fields and in different years. It seems that the cropping system of winter OSR cultivation followed by paddy rice in autumn could reduce the gene flow risk via seedbank in the Yangtze River Basin region (Zhu *et al.*, 2012).

7.4 Modeling Gene Flow to and from Wild Relatives (and Weeds)

7.4.1 Modeling gene flow from crops to wild relatives

To model gene flow from GM rice to wild common rice, the same model for gene flow between GM and non-GM rice crops can be used but with the application of lower sexual cross-compatibility (Rong *et al.*, 2010). Thus, a single model can be used for both directions of gene flow with adjusted parameters if necessary. However, the modeling of gene flow from crops to wild relatives should take into account growth dynamics and the fate of the hybrids in natural ecosystems. In modeling gene flow from crop rice to the same species of weedy rice, Dauer *et al.* (2018) used characteristics of weedy rice as parameters in their simulation model, such as outcrossing rate, dormancy, germination, seedling survival, seed production, and seed shattering. To simulate the herbicide resistance evolution, survival after herbicide use was also considered. In addition, hybrids that formed between crops and their wild relatives were found to backcross to the wild maternal plants, and the escaped transgene could be expressed normally in the backcrossed progenies (Cao *et al.*, 2014). Hence, the fate and fitness of the advanced generation should be taken into account. Both empirical and mechanistic models could be employed to model gene flow from crops to their wild relatives.

Sester *et al.* (2008) adapted the GeneSys model to simulate gene flow from herbicide-resistant sugar beet (*Beta vulgaris*) to wild beet (a weedy form of sugar beet). As wild beet is different from OSR volunteer plants and the wild beet seeds were more dormant and survived even with lower reproductive ability, the parameters were adjusted using data from the field experiments of this weed, and weed management elements were incorporated into modeling. This model could also be used in a simulation study to identify key elements for weed control and gene flow management (Tricault *et al.*, 2009).

Garnier *et al.* (2014) developed a mechanistic crop–wild hybridization model to make predictions about the fate of advantageous crop transgenes in a population of a wild relative species. They used OSR and wild radish (*Raphanus raphanistrum*) as examples for a case study in demonstrating the modeling process and output. This model considered advanced generations of backcrossed progenies and the demographic status of wild populations. The study suggested that accounting for uncertainty in setting the parameters could be crucial in avoiding obtaining a misleading prediction.

7.4.2 Modeling gene flow from herbicide-tolerant weeds

Single herbicide tolerance (mainly to glyphosate) in GM crops was consistently a dominant engineered trait until about 2018, when the adoption of stacked traits (insect resistance and herbicides resistance, IR/HR) continually increased (ISAAA, 2019). Although areas of herbicide-resistant GM crops were reduced in 2019, areas of stacked trait (IR/HR) crops continued to grow, and GM crops tolerant to new herbicides have been approved (ISAAA, 2019). Worldwide glyphosate use has risen almost 15-fold since GM glyphosate-tolerant crops were introduced in 1996 (Benbrook, 2016). Widespread use of the herbicide and adoption of glyphosate-resistant crop species have led to strong selection pressure for and emergence of weed species with glyphosate resistance (Beckie *et al.*, 2019; see also Chapter 6 of this book). Globally, 53 weed species had developed glyphosate resistance by 2020 (Heap, 2021). With the gene flow via pollen and seeds, herbicide resistance spread very

quickly. Glyphosate-resistant kochia (*Kochia scoparia*) was first reported in 2007 in Kansas, USA, and the resistance extended quickly over ten other western states and three Canadian prairie provinces (see Chapter 6, this book). Studies suggested that under normal farming practices, EPSPS (5-enolpyruvylshikimate-3-phosphate synthase) gene copy numbers are continuing to increase in the field, where tandem gene amplification of four or more EPSPS copies on a single chromosome conferred the glyphosate resistance in kochia. By resampling the wild populations, two of three Alberta populations were found to have significant increases in the average EPSPS copy number (Martin *et al.*, 2017). After receiving a herbicide-resistant transgene or evolving herbicide resistance in weeds, it is important to manage the further flow of herbicide-resistant alleles via pollen or seeds within and among populations of the same or related species (Beckie *et al.*, 2019; Jhala *et al.*, 2021).

Herbicide resistance provided a convenient approach to detect gene flow (Wei *et al.*, 2013). Most hybrid plants formed between GM crops and wild relatives (including weeds) could be identified by spraying herbicides and confirmed by molecular evidence, which would then be used to estimate the pollen-mediated gene flow frequency. Creeping bentgrass (*Agrostis stolonifera*) is a cool-season, wind-pollinated perennial grass used on golf courses that is sometimes considered an economic weed and has been reported as weedy in some countries (Watrud *et al.*, 2004). Glyphosate-resistant GM creeping bentgrass was first tested in the field in 2003 and was deregulated by the US Department of Agriculture (USDA) in 2017 (ISAAA, 2017). This could be the first grass species that was genetically modified to carry herbicide resistance. Like other grasses, the pollen is small, the seed is also very small, and creeping bentgrass is sexually compatible with its many wild relatives (e.g. *A. gigantea*, *A. idahoensis*, *A. pallens* and *A. scabra*) (Banks *et al.*, 2003; Van de Water *et al.*, 2007). All these traits potentially favor dispersal and gene flow. By glyphosate spraying and EPSPS gene amplification in progenies of downwind plants, the distance of pollen-mediated gene flow was recorded as far away as 21 km (Watrud *et al.*, 2004). Although under favorable atmospheric conditions the pollen could theoretically disperse to hundreds and even thousands of kilometers, pollen is only viable for approximately 3 h. Backward trajectory analysis had to be employed to correlate the pollen deposition and fertilization event with the timing of the pollen-shed of creeping bentgrass (Van de Water *et al.*, 2007). The back-trajectory analysis is very useful in tracking particle movement in space along timelines. Bilińska *et al.* (2017) applied this method in tracing the pollen source of common ragweed (*Ambrosia artemisiifolia*) in Poland. With backward trajectory analysis, van de Water *et al.* (2007) employed the HYSPLIT atmospheric model for particle dispersion to simulate the dispersal of creeping bentgrass pollen and revealed that the viable pollen could potentially disperse as far as 75 km downwind of the initial release point within 3 h.

Both common ragweed and giant ragweed (*A. trifida*) have evolved resistance to multiple herbicides, including glyphosate (Heap, 2021). These two species grow in common habitats and have overlapping flowering times. Interspecific gene transfer is possible between the two species (Jhala *et al.*, 2021), which may form hybrids that are resistant to herbicides at more sites of action. As the seeds of giant ragweed are large, pollen-mediated gene flow is particularly important because of the outcrossing nature of this weed species (Ganie and Jhala, 2017). Using collected data from two years of field experiments, Ganie and Jhala (2017) employed a double exponential decay model to reveal the impact of distance and wind direction on gene flow. They found that the maximum distance at which a 90% reduction in gene flow occurred was 49.5–106.5 m, depending on different years and various wind directions. They believed that modeling of gene flow could help in weed resistance management.

Among the three species in the genus *Conyza* that have evolved resistance to herbicides, horseweed (*Conyza canadensis*) has the widest distribution (Heap, 2021). It occurs in more than 20 US states and has also been reported in Canada (Ontario province), Eastern Asia (China, Japan, and South Korea), Europe (Czech Republic, France, Greece, Hungary, Italy, Portugal, and Spain) and South America (Brazil). It was reported that the total number of pollen grains generated from horseweed was $2.3–5.1 \times 10^6$ per plant during a pollination season (Jhala *et al.*, 2021) and that the release

of horseweed pollen reaches 80–100 m high vertically above the plant canopy (Huang *et al*., 2015). In their field study, Huang *et al*. (2015) found that horseweed pollen deposition with distance followed an exponential function. After collecting the field experiment data for horseweed, for example pollen concentration, deposition, source strength and atmospheric parameters, Wang *et al*. (2017) adapted their Lagrangian model of maize pollen dispersion (Wang and Yang, 2010) to simulate pollen flow in horseweed. The model was run to simulate pollen dispersion under different environmental conditions. The authors concluded that while atmospheric conditions cannot be controlled, large fields of crops with a high leaf area density and tall plants can effectively prevent pollen dispersion of horseweed (Wang *et al*. 2017). In addition, Shields *et al*. (2006) reported that a single horseweed plant can produce more than 200,000 seeds and that the seeds could be released to a vertical height of 40–140 m and disperse to more than 500 km. Thus, the gene flow risk of horseweed via seeds is also high. Liu *et al*. (2018) validated the HYSPLIT model to simulate horseweed seed dispersal using experimental data, and then by simulating seed dispersion under weak and strong wind scenarios and for the entire seed-shedding season. They found that horseweed seeds could spread to 36.5 km at a wind speed of 1–4 m/s and to 165 km at a wind speed of 5–7 m/s. The results are consistent with the findings in Shields *et al*. (2006).

7.5 Predicting and Monitoring Gene Flow Using Geographic Information Systems

Although isolation distance is requested for GM crop planting to prevent contaminating local conventional crops and landraces (e.g. Baltazar *et al*., 2015), attention should be paid to transgene introgression in crop wild relatives that are also important resources for crop improvement, especially in the areas of sympatric presence of both GM crops and crop wild relatives (Dong *et al*., 2018) where natural hybridization could happen (Ellstrand, 2014). Sympatry plays an important role in gene flow. Wilkinson *et al*. (2003) assessed hybrids arising from long-distance pollination and those formed at sympatric sites between the *B. napus* crop and wild/weedy *B. rapa* plants. Through population surveys, remote sensing, pollen dispersal profiles, herbarium data, and local flora and other floristic databases, they estimated that plenty of hybrids formed annually, while an earlier study could only predict fewer hybrids due to infrequent local sympatry (Wilkinson *et al*., 2000).

The cultivation areas at regional, national or local levels of target crops could be collected from agricultural yearbooks or annual reports of agricultural departments or both, a relatively easier process compared with locating the distribution of crop wild relatives. The records on naturally occurring crop wild relatives are usually sporadic. Data collected from herbaria and the literature normally describe known occurrences of species, which may be geographically skewed. This type of data has been used in many investigations, being justified by the lack of systematic survey data (Elith and Leathwick, 2009). Species distribution models can be employed to predict the potential distribution of crop wild relatives using natural occurrence information and bioclimatic variables. For example, sympatric occurrence can be calculated by mapping canola cultivation areas and the distribution of its wild relatives using geographic information systems. Dong *et al*. (2018) used this approach to predict the potential distribution of wild brassicas in China, together with the geographic information of national OSR crop cultivation. Sites with a high level of gene flow were identified by data of flowering overlapping with information on sexual compatibility between OSR crops and specific wild relatives. Three main factors were considered and qualitatively scored: (i) cultivation area (CA) with five score levels from 1 to 5 (low to high); (ii) flowering overlap (FO) with three levels of scores from 1 to 3 (low to high); and (iii) sympatric presences (SP) of the crop and wild relatives with scores from 1 to 5 (low to high). The sum of quantitative scores for the three factors (CA, FO, and SP) would then predict the whole gene flow levels (GF).

In their study, three regions of high-level gene flow were identified, among which two situations for high gene flow levels were revealed (Dong *et al*., 2018): (i) the main production regions of OSR with a modest occurrence of wild relative plants; and (ii) the main distribution

areas of wild relatives with a moderate area of the OSR crop. To ensure a low level of gene flow, the former requires frequent monitoring of any changes in functional traits of wild relatives, which may be caused by the transgene escape from the crop; for the latter, the GM crop or any new cultivar should be released only after careful assessment or evaluation or to take necessary measures to avoid or reduce the possibility of any adverse impact of crops posed to wild relatives.

7.6 Perspectives

In cases in which transgene escape could be inevitable, measures and strategies should be taken to mitigate the impacts on natural and agricultural ecosystems (Kareiva *et al.*, 1994). Gene flow models could assist in optimizing best management practices in reducing and mitigating adverse effects. The modeling of gene flow between crops was originally used to reduce contamination of other crops in breeding and is mostly used to minimize the AP of unauthorized GM components in food chains. Reducing the AP of GM components to a certain level in food supply chains could be possible, especially in large field releases, through appropriate field management and control, as optimistically estimated by some authors (e.g. Beckie and Hall, 2008). However, that optimistic idea needs further verification. In addition, multiple pollen and seed sources, which could be present in smallholder farming systems, would pose difficulties and challenges for gene flow modeling and management (Zhang *et al.*, 2018).

Manasse (1992) once found an increase in mean gene dispersal distance with increased distance between clumps or individual plants of the pollen recipients, which the author attributed to pollinating behaviors. Pollinators routinely travel to distant plants over spaces that do not have any attractive plants, and a low density of plants would increase foraging movement. However, pollen-mediated gene flow could also increase to longer distances when pollen recipients in adjusted clumps or populations might be able to serve as new pollen sources for further gene spread. Introgressed wild individuals or populations could act as a bridge, enabling low-level and long-distance transgene flow (Beckie *et al.*, 2019). Thus, gene flow from crops to their wild relatives is vitally important. In modeling this process, it is necessary to account for the fate of hybrid progenies. Once established in wild populations, which is the first step of escape, the transgene could escape further into nature. Studies have already considered the population status of wild maternal plants (e.g. Garnier *et al.*, 2014). Future work on modeling should also consider the growth performance and fitness of hybrids and their advanced generation progenies formed with the wild plants. With the large release of GM crops that are resistant to herbicides, weeds that have evolved resistance to herbicides may also serve as a new source of herbicide-resistant genes to spread in nature via either pollen or seeds. This has become a new hotspot for gene flow modeling studies.

Uncertainty that results from climate change and human activities is an inherent element of scientific analysis and risk assessment (SCBD, 2016) that should be appropriately dealt with in modeling gene flow. For example, climate change may affect the growth (for example, emergence of seedlings) of sugar beet as well as that of weed beet (Lamichhane *et al.*, 2019), which may cause changes in cropping systems and pest control measures and result in different management for gene flow. Bensadoun *et al.* (2014) suggested using the Bayesian method to predict crossing rates using spatial and climatic data to deal with uncertainty in gene flow modeling, which could allow integrating much more additional data such as climatic variables and thus improve the accuracy of model predictions. However, this has to be validated with additional field factors and real agricultural landscape settings.

Acknowledgments

The authors gratefully acknowledge the partial financial support for this work from two research projects of the National Natural Science Foundation of China (project nos 30300044 and 31370357)

and from the Illinois State Water Survey at the University of Illinois at Urbana-Champaign. The opinions expressed are those of the authors and not necessarily those of the Illinois State Water Survey, the Prairie Research Institute, or the University of Illinois at Urbana-Champaign.

References

Amand, P.C.S., Skinner, D.Z. and Peaden, R.N. (2000) Risk of alfalfa transgene dissemination and scale-dependent effects. *Theoretical and Applied Genetics* 101(1-2), 107–114. DOI: 10.1007/s001220051457.

Aylor, D.E., Schultes, N. and Shields, E.J. (2003) An aerobiological framework for assessing cross-pollination in maize. *Agricultural and Forest Meteorology* 119(3-4), 111–129. DOI: 10.1016/S0168-1923(03)00159-X.

Baltazar, B.M., Castro Espinoza, L., Espinoza Banda, A., de la Fuente Martínez, J.M., Garzón Tiznado, J.A. *et al*. (2015) Pollen-mediated gene flow in maize: implications for isolation requirements and co-existence in Mexico, the center of origin of maize. *Plos One* 10(7), e0131549. DOI: 10.1371/journal.pone.0131549.

Banks, P., Branham, B., Harrison, K., Whitson, T. and Heap, I. (2003) Determination of the potential impact from the release of glyphosate- and glufosinate-resistant *Agrostis stolonifera* L. in various crop and non-crop ecosystems. Available at: https://www.aphis.usda.gov/brs/pdf/Resistant_Bentgrass_Final_Report.pdf (accessed 16 March 2021).

Bateman, A.J. (1947a) Contamination of seed crops: I. Insect pollination. *Journal of Genetics* 48(2), 257–275. DOI: 10.1007/BF02989385.

Bateman, A.J. (1947b) Contamination of seed crops: II. Wind pollination. *Heredity* 1, 235–246.

Bateman, A.J. (1947c) Contamination in seed crops: III. Relation with isolation distance. *Heredity* 1, 303–336.

Beckie, H.J. and Hall, L.M. (2008) Simple to complex: modelling crop pollen-mediated gene flow. *Plant Science* 175, 615–628.

Beckie, H.J., Busia, R., Bagavathiannan, M.V. and Martin, S.L. (2019) Herbicide resistance gene flow in weeds: under-estimated and underappreciated. *Agriculture, Ecosystems & Environment* 283(Sp Iss), 106566. DOI: 10.1016/j.agee.2019.06.005.

Benbrook, C.M. (2016) Trends in glyphosate herbicide use in the United States and globally. *Environmental Sciences Europe* 28(1), 3. DOI: 10.1186/s12302-016-0070-0.

Bensadoun, A., Monod, H., Angevin, F., Makowski, D. and Messéan, A. (2014) Modelling of gene flow by a Bayesian approach: a new perspective for decision support. *AgBioforum* 17, 213–220.

Bilińska, D., Skjøth, C.A., Werner, M., Kryza, M., Malkiewicz, M. *et al*. (2017) Source regions of ragweed pollen arriving in south-western Poland and the influence of meteorological data on the HYSPLIT model results. *Aerobiologia* 33(3), 315–326. DOI: 10.1007/s10453-017-9471-9.

Bryant, R.H., Emberlin, J.C. and Norris-Hill, J. (1989) Vertical variation in pollen abundance in North-Central London. *Aerobiologia* 5(2), 123–137. DOI: 10.1007/BF02486510.

Cao, D., Stewart, C.N. Jr., Zheng, M., Guan, Z.-J., Tang, Z.-X. *et al*. (2014) Stable *Bacillus thuringiensis* transgene introgression from *Brassica napus* to wild mustard *B. juncea*. *Plant Science* 227, 45–50. DOI: 10.1016/j.plantsci.2014.06.018.

Castellari, E., Soregarolia, C., Venus, T.J. and Wesseler, J. (2018) Food processor and retailer non-GMO standards in the US and EU and the driving role of regulations. *Food Policy* 78(12), 26–37. DOI: 10.1016/j.foodpol.2018.02.010.

Colbach, N. (2008) Commentary on Beckie & Hall: 'Simple to complex: Modelling crop pollen-mediated gene flow'. *Plant Science* 175(5), 629. DOI: 10.1016/j.plantsci.2008.07.007.

Colbach, N., Roger-Estrade, J., Chauvel, B. and Caneill, J. (2000) Modelling vertical and lateral seed bank movements during mouldboard ploughing. *European Journal of Agronomy* 13(2-3), 111–124. DOI: 10.1016/S1161-0301(00)00069-1.

Colbach, N., Clermont-Dauphin, C. and Meynard, J.M. (2001a) GeneSys: A model of the influence of cropping system on gene escape from herbicide tolerant rapeseed crops to rape volunteers. I. Temporal evolution of a population of rapeseed volunteers in a field. *Agriculture, Ecosystems and Environment* 83, 235–253.

Colbach, N., Clermont-Dauphin, C. and Meynard, J.M. (2001b) GeneSys: A model of the influence of cropping system on gene escape from herbicide tolerant rapeseed crops to rape volunteers. II. Genetic exchanges among volunteer and cropped populations in a small region. *Agriculture, Ecosystems and Environment* 83, 255–270.

Colbach, N., Molinari, N., Meynard, J.-M. and Messéan, A. (2005) Spatial aspects of gene flow between rapeseed varieties and volunteers. *Agronomy for Sustainable Development* 25(3), 355–368. DOI: 10.1051/agro:2005035.

Dauer, J., Hulting, A., Carlson, D., Mankin, L., Harden, J. *et al*. (2018) Gene flow from single and stacked herbicide-resistant rice (*Oryza sativa*): modeling occurrence of multiple herbicide-resistant weedy rice. *Pest Management Science* 74(2), 348–355. DOI: 10.1002/ps.4711.

Dong, J.-J., Zhang, M.-G., Wei, W., Ma, K.-P. and Wang, Y.-H. (2018) GIS assessment of the risk of gene flow from *Brassica napus* to its wild relatives in China. *Environmental Monitoring and Assessment* 190(7), 1–12. DOI: 10.1007/s10661-018-6753-9.

Eastham, K. and Sweet, J. (2002) *Genetically Modified Organisms (GMOs): The Significance of Gene Flow Through Pollen Transfer*. European Environment Agency, Copenhagen.

Elith, J. and Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* 40(1), 677–697. DOI: 10.1146/annurev.ecolsys.110308.120159.

Ellstrand, N.C. (2014) Is gene flow the most important evolutionary force in plants? *American Journal of Botany* 101(5), 737–753. DOI: 10.3732/ajb.1400024.

Eriksson, D., Custers, R., Björnberg, K.E., Hansson, S.O., Purnhagen, K. *et al*. (2020) Options to reform the European Union legislation on GMOs: post-authorization and beyond. *Trends in Biotechnology* 38(5), 465–467. DOI: 10.1016/j.tibtech.2019.12.015.

Ganie, Z.A. and Jhala, A.J. (2017) Modeling pollen-mediated gene flow from glyphosate-resistant to susceptible giant ragweed (*Ambrosia trifida* L.) under field conditions. *Scientific Reports* 7(1), 17067. DOI: 10.1038/s41598-017-16737-z.

Garnier, A., Darmency, H., Tricault, Y., Chevre, A.M. and Lecomte, J. (2014) A stochastic cellular model with uncertainty analysis to assess the risk of transgene invasion after crop–wild hybridization: oilseed rape and wild radish as a case study. *Ecological Modelling* 276, 85–94. DOI: 10.1016/j.ecolmodel.2014.01.004.

Gruber, S., Buhler, A., Mohring, J. and Claupein, W. (2010) Sleepers in the soil – vertical distribution by tillage and long-term survival of oilseed rape seeds compared with plastic pellets. *European Journal of Agronomy* 33(2), 81–88. DOI: 10.1016/j.eja.2010.03.003.

Heap, I. (2021) The international survey of herbicide resistant weeds. Available at: www.weedscience.org (accessed 25 February 2021).

Hokanson, K.E., Ellstrand, N. and Raybould, A. (2018) The integration of science and policy in regulatory decision-making: observations on scientific expert panels deliberating GM crops in centers of diversity. *Frontiers in Plant Science* 9, 1157. DOI: 10.3389/fpls.2018.01157.

Hu, N., Jiang, X.-D., Yuan, Q.-H., Liu, W.-G., Yao, K.-M. *et al*. (2020) Increased pollen source area does not always enhance the risk of pollen dispersal and gene flow in *Oryza sativa* L. *Scientific Reports* 10(1), 6143. DOI: 10.1038/s41598-020-63119-z.

Huang, H.-Y., Ye, R.-J., Qi, M.-L., Li, X.-Z., Miller, D.R. *et al*. (2015) Wind-mediated horseweed (*Conyza canadensis*) gene flow: pollen emission, dispersion, and deposition. *Ecology and Evolution* 5(13), 2646–2658. DOI: 10.1002/ece3.1540.

Huffman, W. and Strzok, J. (2017) *The Economics of Organic and GMO Farming Systems (in the US): Interactions and How They Might Co-exist*, Economics Working Papers, 17035. Iowa State University, Ames, Iowa. Available at: https://lib.dr.iastate.edu/econ_workingpapers/36 (accessed 14 July 2021)

ISAAA (2017) *Global Status of Commercialized Biotech/GM Crops in 2017: Biotech Crop Adoption Surges as Economic Benefits Accumulate in 22 Years*. ISAAA Brief No. 53. International Service for the Acquisition of Agri-biotech Applications, Ithaca, New York.

ISAAA (2019) *Global Status of Commercialized Biotech/GM Crops: 2019. ISAAA Brief No. 55*. International Service for the Acquisition of Agri-biotech Applications, Ithaca, New York.

Jhala, A.J., Norsworthy, J.K., Ganie, Z.A., Sosnoskie, L.M., Beckie, H.J. *et al*. (2021) Pollen-mediated gene flow and transfer of resistance alleles from herbicide-resistant broadleaf weeds. *Weed Technology* 35(2), 173–187. DOI: 10.1017/wet.2020.101.

Kareiva, P., Morris, W. and Jacobi, C.M. (1994) Studying and managing the risk of cross-fertilization between transgenic crops and wild relatives. *Molecular Ecology* 3(1), 15–21. DOI: 10.1111/j.1365-294X.1994.tb00037.x.

Kesoju, S.R., Greene, S.L., Martin, R.C., Kramer, M., Walsh, D.B. *et al*. (2019) Isolation distances for transgenic alfalfa seed production in the Pacific Northwest. *Crop Science* 59(4), 1701–1708. DOI: 10.2135/cropsci2018.07.0414.

Lal, M., Bhardwaj, E., Chahar, N., Dangwal, M. and Sandip Das, S. (2020) Trans)Gene flow: mechanisms, biosafety concerns and mitigation for containment. In: Tandon, R., Shivanna, K.R. and Koul, M. (eds) *Reproductive Ecology of Flowering Plants: Patterns and Processes*. Springer, Singapore, pp. 335–394.

Lamichhane, J.R., Constantin, J., Aubertot, J.-N. and Dürr, C. (2019) Will climate change affect sugar beet establishment of the 21st century? Insights from a simulation study using a crop emergence model. *Field Crops Research* 238, 64–73.

Liu, J., Qi, M.L. and Wang, J.M. (2018) Long-distance and dynamic seed dispersal from horseweed (*Conyza canadensis*). *Ecoscience* 25(3), 271–285. DOI: 10.1080/11956860.2018.1455371.

Liu, Y.-B., Wei, W., Ma, K.-P., Li, J.-S., Liang, Y.-Y. *et al*. (2013) Consequences of gene flow between oilseed rape (*Brassica napus*) and its relatives. *Plant Science* 211, 42–51. DOI: 10.1016/j.plantsci.2013.07.002.

Lu, H.-W., Lu, J.-Z. and He, L. (2019) Modeling and estimation of pollen-mediated gene flow at the landscape scale. *Ecological Indicators* 106, 105500. DOI: 10.1016/j.ecolind.2019.105500.

Manasse, R.S. (1992) Ecological risks of transgenic plants: effects of spatial dispersion on gene flow. *Ecological Applications* 2(4), 431–438. DOI: 10.2307/1941878.

Martin, S.L., Benedict, L., Sauder, C.A., Wei, W., da Costa, L.O. *et al*. (2017) Glyphosate resistance reduces kochia fitness: comparison of segregating resistant and susceptible F2 populations. *Plant Science* 261, 69–79. DOI: 10.1016/j.plantsci.2017.04.010.

Meillet, A., Angevin, F., Bensadoun, A., Huby, G., Monod, H. *et al*. (2015) Design of a decision support tool for managing coexistence between genetically modified and conventional maize at farm and regional levels. *Ecological Informatics* 30(3), 379–388. DOI: 10.1016/j.ecoinf.2015.09.014.

Ramsay, G. (2005) Pollen dispersal vectored by wind or insects. In: . and Poppy, G.M. and Wilkinson, M.J. (eds) *Gene Flow from GM Plants*. Blackwell, Oxford, UK, pp. 43–77. DOI: 10.1002/9780470988497.ch3.

Ridley, C.E. and Alexander, L.C. (2016) Applying gene flow science to environmental policy needs: a boundary work perspective. *Evolutionary Applications* 9(7), 924–936. DOI: 10.1111/eva.12393.

Rieger, M.A., Lamond, M., Preston, C., Powles, S.B. and Roush, R.T. (2002) Pollen-mediated movement of herbicide resistance between commercial canola fields. *Science* 296(5577), 2386–2388. DOI: 10.1126/science.1071682.

Roger-Estrade, J., Richard, G., Boizard, H., Boiffin, J., Caneill, J. *et al*. (2000) Modelling structural changes in tilled topsoil over time as a function of cropping systems. *European Journal of Soil Science* 51(3), 455–474. DOI: 10.1046/j.1365-2389.2000.00323.x.

Rong, J., Song, Z.-P., de Jong, T.J., Zhang, X.-S., Sun, S.-G. *et al*. (2010) Modelling pollen-mediated gene flow in rice: risk assessment and management of transgene escape. *Plant Biotechnology Journal* 8(4), 452–464. DOI: 10.1111/j.1467-7652.2009.00488.x.

Rong, J., Wang, F., Song, Z.-P., Su, J., Chen, R. *et al*. (2012) Scale effect on rice pollen-mediated gene flow: implications in assessing transgene flow from genetically engineered plants. *Annals of Applied Biology* 161(1), 3–11. DOI: 10.1111/j.1744-7348.2012.00545.x.

SCBD (2016) *Guidance on Risk Assessment of Living Modified Organisms and Monitoring in the Context of Risk Assessment*. Secretariat of the Convention on Biological Diversity, Montreal, Canada.

Sester, M., Tricault, Y., Darmency, H. and Colbach, N. (2008) GeneSys-Beet: A model of the effects of cropping systems on gene flow between sugar beet and weed beet. *Field Crops Research* 107(3), 245–256. DOI: 10.1016/j.fcr.2008.02.011.

Shields, E.J., Dauer, J.T., VanGessel, M.J. and Neumann, G. (2006) Horseweed (*Conyza canadensis*) seed collected in the planetary boundary layer. *Weed Science* 54(6), 1063–1067. DOI: 10.1614/WS-06-097R1.1.

Sun, A.-Q., Zhang, C.-Q., Wu, C.-L. and Gao, Q.-R. (2015) Pollen flow of wheat under natural conditions in the Huanghuai river wheat region, China. *GM Crops & Food* 6(3), 135–149. DOI: 10.1080/21645698.2015.1006072.

Tricault, Y., Darmency, H. and Colbach, N. (2009) Identifying key components of weed beet management using sensitivity analyses of the GENESYS-BEET model in GM sugar beet. *Weed Research* 49, 581–591.

USDA (2015) *USDA Coexistence Factsheets: Biotech Farming*. United States Department of Agriculture Washington, DC.

Vallaeys, V., Tyson, R.C., Lane, W.D., Deleersnijder, E. and Hanert, E. (2017) A Lévy-flight diffusion model to predict transgenic pollen dispersal. *Journal of The Royal Society Interface* 14(126), 20160889. DOI: 10.1098/rsif.2016.0889.

van de Water, P.K., Watrud, L.S., Lee, E.H., Burdick, C. and King, G.A. (2007) Long-distance GM pollen movement of creeping bentgrass using modeled wind trajectory analysis. *Ecological Applications* 17(4), 1244–1256. DOI: 10.1890/06-0962.

Wang, J.-M, Qi, M.-L., Huang, H.-Y., Ye, R.-J., Li, X.-Z. et al. (2017) Atmospheric pollen dispersion from herbicide-resistant horseweed (*Conyza canadensis* L.). *Aerobiologia* 33, 393–406. DOI: 10.1007/s10453-017-9477-3.

Wang, J.-M. and Yang, X.S. (2010) Development and validation of atmospheric gene flow model for assessing environmental risks from transgenic corn crops. *International Journal of Agricultural and Biological Engineering* 3, 18–30.

Wang, J.-M., Sammis, T.W., Miller, D.R., Hiscox, A.L. and Granucci, D. (2011) Simulated regional PM10 dispersion from agricultural tilling operations using HYSPLIT. *Transactions of the ASABE* 54, 1659–1667.

Wang, L. and Lu, B.-R. (2017) Model-based calculating tool for pollen-mediated gene flow frequencies in plants. *AoB PLANTS* 9, plw086. DOI: 10.1093/aobpla/plw086.

Wang, L., Haccou, P. and Lu, B.-R. (2016) High-resolution gene flow model for assessing environmental impacts of transgene escape based on biological parameters and wind speed. *Plos One* 11(3), e0149563. DOI: 10.1371/journal.pone.0149563.

Watrud, L.S., Lee, E.H., Fairbrother, A., Burdick, C., Reichman, J.R. *et al*. (2004) Evidence for landscape-level, pollen-mediated gene flow from genetically modified creeping bentgrass with CP4 EPSPS as a marker. *Proceedings of the National Academy of Sciences (USA)* 101(40), 14533–14538. DOI: 10.1073/pnas.0405154101.

Wei, W., Qian, Y.-Q. and Ma, K.-P. (1999) Gene flow between transgenic crops and their wild related species. *Journal of Integrative Plant Biology (formerly Acta Botanica Sinica)* 41(4), 343–348. [In Chinese with English abstract.]

Wei, W., Kwit, C., Millwood, R.J., Moon, H.-S. and Stewart, C.N. Jr. (2013) Assessment and detection of gene flow. In: Oliver, M. and Li, Y. (eds) *Plant Gene Containment*. John Wiley & Sons, Hoboken, New Jersey, pp. 27–42.

Wilkinson, M.J., Davenport, I.J., Charters, Y.M., Jones, A.E., Allainguillaume, J. *et al*. (2000) A direct regional scale estimate of transgene movement from genetically modified oilseed rape to its wild progenitors. *Molecular Ecology* 9(7), 983–991. DOI: 10.1046/j.1365-294x.2000.00986.x.

Wilkinson, M.J., Elliott, L.J., Allainguillaume, J., Shaw, M.W., Norris, C. *et al*. (2003) Hybridization between *Brassica napus* and *B. rapa* on a national scale in the United Kingdom. *Science* 302(5644), 457–459. DOI: 10.1126/science.1088200.

Yi, J., Qiu, W.-W., Hu, W., Zhang, H.-L., Liu, M.-X. *et al*. (2020) Effects of cultivation history in paddy rice on vertical water flows and related soil properties. *Soil and Tillage Research* 200, 104613. DOI: 10.1016/j.still.2020.104613.

Zhang, C.-J., Yook, M.-J., Park, H.-R., Lim, S.-H., Kim, J.-W. *et al*. (2018) Assessment of potential environmental risks of transgene flow in smallholder farming systems in Asia: *Brassica napus* as a case study in Korea. *Science of the Total Environment* 640–641, 688–695. DOI: 10.1016/j.scitotenv.2018.05.335.

Zhang, Z.-T., Yang, X.-G., Liu, Z.-J., Bai, F., Sun, S. *et al*. (2020) Spatio-temporal characteristics of agro-climatic indices and extreme weather events during the growing season for summer maize (*Zea mays* L.) in Huanghuaihai region, China. *International Journal of Biometeorology* 64(5), 827–839. DOI: 10.1007/s00484-020-01872-6.

Zhu, Y.-M. (2012) Assessment of potential risk of gene flow through soil seedbank caused by GM OSR in Chinese farming system. PhD thesis. Institute of Botany, Chinese Academy of Sciences, China.

Zhu, Y.-M., Li, Y.-D., Colbach, N., Ma, K.-P., Wei, W. *et al*. (2012) Seed losses at harvest and seed persistence of oilseed rape (*Brassica napus*) in different cultural conditions in Chinese farming systems. *Weed Research* 52(4), 317–326. DOI: 10.1111/j.1365-3180.2012.00929.x.

8 Controlling Transgene Flow from Engineered Crops to Unintended Hosts by Molecular Approaches

C. Neal Stewart Jr*
Department of Plant Sciences; Center for Agricultural Synthetic Biology, University of Tennessee, Knoxville, TN, United States

Abstract

For most transgenic crops, the purported ecological risk from transgenic-host hybridization and introgression to unintended host species is negligible. Nonetheless, there remains a risk-associated focus on the potential for gene flow in the governance and regulation of crop biotechnology. Because of uncertainties in the large world of biology as well as regulatory certainties (regulations will likely not diminish), researchers and stakeholders have a great interest in eliminating or substantially decreasing gene flow from transgenic crops. To that end, numerous approaches have been investigated for limiting transgene flow via hybridization and introgression to unintended hosts. While such bioconfinement may be accomplished by ecological and management strategies as discussed elsewhere in this book, this chapter focuses on mitigating unintended gene flow from engineered crops by way of genetic engineering itself. The chapter will mainly discuss the manipulation of relatively simple means to alter plant sexual reproduction and plant growth and development to control transgene flow, with the desired outcome being the prevention of transgenes from moving and/or introgression into free-living unintended hosts. These approaches include: (i) decreasing or delaying flowering; (ii) eliminating pollen production via male sterility or selective male sterility; (iii) removing transgenes from pollen or eggs by gene use restriction technologies; and (iv) kill switches. Emerging synthetic biology approaches that may be used for transgene bioconfinement are explored. Taken together, the same molecular biology strategies that are used to improve crops can also help assure their biosafety.
Keywords: bioconfinement; gene use restriction technologies; introgression; synthetic biology

8.1 Introduction

Most commodity crops that have been genetically engineered are relatively safe with regard to potential gene flow and other ecological effects (NASEM, 2016). Much of this absence of transgene flow may be associated with the biology of crops (for example, they are highly domesticated and mainly selfing) as well as having a substantial absence of wild relatives in major cultivation regions. These factors essentially act to bioconfine transgenes

*nealstewart@utk.edu

 Gene Flow: Monitoring, Modeling and Mitigation (eds Wei Wei and C. Neal Stewart, Jr)
DOI: 10.1079/9781789247480.0008

to their intended crop hosts. In these cases, transgenes typically do not travel to weedy and wild relatives and, when they do, they seldom introgress to form stably engineered populations in the unintended host taxa (Stewart *et al.*, 2003; Kwit *et al.*, 2011; Ellstrand *et al.*, 2013). Indeed, the only documented case of transgene introgression into a weedy relative species in an agricultural field as documented in the peer-reviewed literature involves genetically engineered canola (*Brassica napus*) harboring a herbicide resistance transgene that passed the transgene to *Brassica rapa*. Introgression occurred, presumably, when the hybrids and then advanced hybrids formed by repeated subsequent backcrosses to *B. rapa* facilitated the loss of *B. napus* genes, resulting in the transgene persisting in the *B. rapa* genetic background (Warwick *et al.*, 2008). In this case, only a few individual transgenic herbicide-resistant plants were found in a field where transgenic canola had been cultivated. This particular case in a Canadian agricultural field in Quebec is heuristic in that while several transgenic 'weeds' were found over a 6-year period on the site, only one individual transgenic introgressed plant persisted in the field at the end of the study; that plant had notably decreased fitness (Warwick *et al.*, 2008). Nonetheless, this example shows that *Brassica*-to-*Brassica* transgene flow and introgression is possible in the field. As reviewed elsewhere, there are potential cases of such introgression occurring in *Sorghum bicolor* (sorghum) to weedy wild relatives of sorghum, such as Johnsongrass (*Sorghum halepense* (Stewart *et al.*, 2003; Ohadi *et al.*, 2017) as well as interesting cases of persistence of herbicide-resistant creeping bentgrass (*Agrostis stolonifera*) within a major turfgrass breeding area in Oregon, USA (Watrud *et al.*, 2004; Reichman *et al.*, 2006).

8.2 Preventing Flowering and Gamete Production Using Biotechnology

While most major transgenic row crops do not seem to be at risk for sourcing transgene flow and introgression, we may imagine several different scenarios in which an additional layer of bioconfinement may be of value, for example in transgenic molecular farming (Clark and Maselko, 2020) and perennial crops for bioenergy and bioproducts (Kwit *et al.*, 2011; Sang *et al.*, 2013). In advanced biomass and bioenergy feedstock crops, which are mostly perennial grasses and trees, there is generally no need for flowering at all, at least from the production perspective. Researchers have been investigating methods of manipulating flowering and reproduction in trees (Fritsche *et al.*, 2018), which include altering the expression of flower transition genes such as *LEAFY* (Klocko *et al.*, 2018) and ablating pollen using bacterial ribonuclease (barnase) systems (Klocko *et al.*, 2016). In herbaceous plants, such as switchgrass (*Panicum virgatum*), flower development may be severely delayed by overexpressing miR156, which is a regulator of vegetative-to-floral transition. In one experiment, while flowering could be eliminated, severely decreasing (but not eliminating) flowering also resulted in significantly higher biomass in switchgrass (Johnson *et al.*, 2017). In some cases, complete gamete ablation, such as the result of using the barnase system, is not desired. In field experiments in tobacco (*Nicotiana tabacum*), complete selective pollen ablation was observed in plants expressing the restriction endonuclease gene coding for *EcoRI* (Millwood *et al.*, 2016). In selective male sterility, only pollen carrying the transgenes is ablated; thus plants are fertile, but do not spread transgenes via pollen. In such cases the manipulation of sexual reproduction, gamete production, and the resultant biocontainment may be desired not only for biosafety, but for product-tracking as well. These relatively simple systems of manipulating plant development genes or destroying gametes are useful, but something of a 'blunt hammer' with regard to controlling gene flow in space and time. Nonetheless, in many envisioned cases, these strategies will be sufficient to bioconfine transgenes. Indeed, studies to date demonstrate that there are several genes and mechanisms that facilitate molecular-based biocontainment of transgenes. Clearly, the genomics revolution has illuminated many

genes whose expression can be altered for bioconfinement.

8.3 Gene Use Restriction Approaches for Bioconfinement and their Mechanisms

Conditional control of sexual reproduction may be theoretically conveyed by gene use restriction technologies (GURTs). GURTs were originally innovated to protect transgenic intellectual property in plants. In the original 'GURT' patents, growers would not be able to save and plant proprietary and transgenic seed (Oliver *et al.*, 1998, 1999a, b). There was, however, significant negative public reaction to the idea of such restrictions being made on farmers' rights to save and plant seed, which led to vocal critics coining the label 'terminator technology,' which helped to further promote public outcry (Lombardo, 2014).

Nonetheless, GURTs could be of tremendous use in bioconfinement of engineered crops (Lombardo, 2014), especially to enable the bioeconomy, for example advanced biofuel feedstocks (Sang *et al.*, 2013), and plants engineered for pharmaceutical production (molecular farming) (Clark and Maselko, 2020). An additional utility of GURTs was to delimit unintended gene flow for regulatory and ecological purposes. Indeed, these applications were a mere glimmer in the eye of innovators when GURTs were first labeled as terminator technology (Lombardo, 2014).

In one embodiment, GURTs can be used to make unintended sexual reproduction impossible in transgenic crops by using site-specific recombination technologies that could be precisely induced. Transgenes would be 'clipped out' of a transgenic plant genome when induced by a stimulus designed to activate the SSR system by way of a gene cascade.

GURTs have been categorized into two classes: V-GURTs (variety-related GURTs) and T-GURTs (trait-related GURTs) (Visser *et al.*, 2001; Van Acker *et al.*, 2007; Hills *et al.*, 2007; Lombardo, 2014). V-GURTs are the basis of the infamous terminator technology, in which all genetic materials contained in an entire plant variety would be protected by a company, which would 'activate' seeds for one-time use prior to sale. These seeds would give rise to plants and, in turn, the next generation of seeds, but these seeds would be sterile. Farmers could not save and reuse engineered seeds. In contrast, T-GURTs restrict only the flow of transgenes and the next generation of seeds would be fertile, but no longer genetically engineered. Given that the readers of this book are more interested in bioconfinement than intellectual property protection, T-GURTs and aspects of controlling gene flow will be emphasized in this chapter even though V-GURTs may be used for bioconfinement applications as well.

For bioconfinement, the simplest GURTs would employ multiple genetic components engineered into plants; for example: (i) target gene endowing a trait; (ii) an excisor gene, such as one encoding a site-specific recombinase or CRISPR-Cas9 or similar designer nuclease system; (iii) an inducible and/or tissue-specific promoter controlling the expression of the excisor gene; and (iv) recognition sites for the SSR or nuclease that flank all the other components (Fig. 8.1). In its simplest embodiment, once the promoter controlling the excisor is activated, the transgene construct is simply clipped out of the tissue of interest. For example, transgenes may be removed from entire seedlings (and subsequent plants), or simply from pollen. In the case of pollen, a pollen-specific promoter has been used to induce transgene excision in pollen at over 99% efficiency (Moon *et al.*, 2010, 2011). For such relatively simple systems in which tissue-specific promoters are available (typically taxa-specific), high levels of biocontainment may be attained to prevent inadvertent gene flow though gametes (Lombardo, 2014).

8.4 Synthetic Biology for Engineering Circuits for Bioconfinement

Another set of forward-looking approaches fall under the umbrella of synthetic biology of plants. As noted in other reviews (Liu and Stewart, 2015, 2016), synthetic biology is an exciting new field in plant biology and one that is complementary to biotechnology. Synthetic biology is at the interface of innovations in

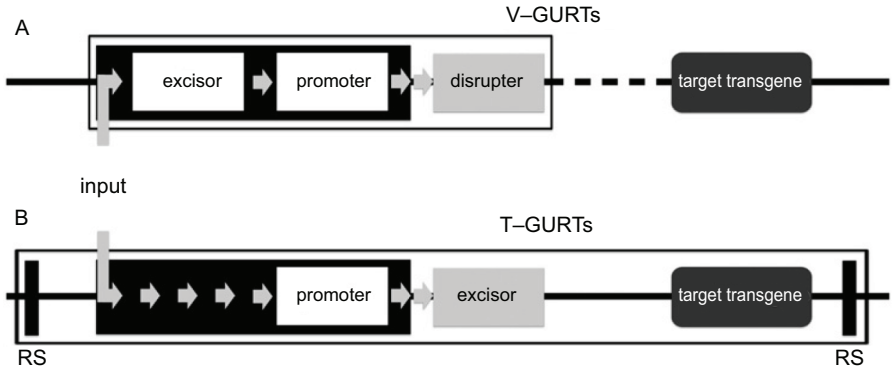


Fig. 8.1. Comparison of (A) variety-related GURTs (V-GURTs) and (B) trait-related GURTs (T-GURTs) for bioconfinement with minimal components shown. An input or stimulus starts a signal transduction that may activate a disruptor and/or excisor gene to act on target transgene disruption or removal in the case of (B), where two recognition sites (RS) flank the DNA to be removed. (Adapted from Sang *et al*., 2013)

computational design and engineering biology to design, build, and test novel DNA in 'chassis' organisms, in this case plants. Synthetic biology strives to implement orthogonal parts to create interchangeable components to produce novel biological effects or even novel organisms.

Essentially, GURTs employ the concept of circuit engineering. Synthetic circuits have been explored in plants, but there are few examples of applications (Kassaw *et al*., 2018; Andres *et al*., 2019). Synthetic regulatory circuits are akin to electronic circuits except that the materials to construct the former are relatively nascent. They require synthetic promoters, transcriptional activators, and other genes to be assembled into functional circuits for applications (Liu and Stewart, 2016; Kassaw *et al*., 2018; Andres *et al*., 2019). The design and characterization of genetic parts for plant engineering has made great strides in the past few years. Of special interest is the development of orthogonal regulatory systems. One such system consists of a library of synthetic transcriptional regulators (Belcher *et al*., 2020). Together, they comprise synthetic promoters, activators, and repressors that can be used to vary transcriptional output for constructing synthetic circuits. While the Belcher *et al*. (2020) parts have been tested only in model plants (stably in *Arabidopsis* and transiently in agroinfiltrated *Nicotiana benthamiana*), such circuits may be envisioned as being powerful in transgene bioconfinement.

8.5 Kill Switches

A particularly useful GURT may act as a 'kill switch' to control inadvertent transgene persistence. While kill switches – indeed, toggle switches in general (Gardner *et al*., 2000) — have been pursued with some enthusiasm in synthetic biology research (Stirling *et al*., 2017), the plant community has been surprisingly slow to engineer systems with kill switches as a means of assuring failsafe biocontainment in crops. Perhaps the absence of rhetoric on the topic goes back to the terminator technology controversies, because the original concept did build in kill switches in the form of seed sterility (Lombardo, 2014). The original concept, however, appeared to be targeted mainly towards applications in annual row crops such as maize and soybean, with intellectual property protection as the main motivation. Nonetheless, niche crops of high value, such as molecular farming crops, could benefit from engineering a kill switch into the host plant. One could imagine some sort of newly discovered problem with the particular pharmaceutical that the transgenic plant produces, and that the owner of the plants wishes

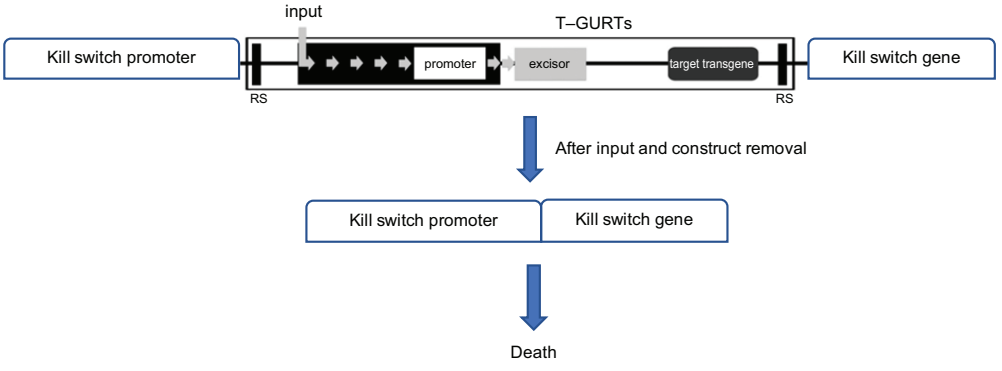


Fig. 8.2. Schematic showing how a T-GURT system may be used to activate a kill switch upon an input or stimulus. When the DNA is removed between the two recognition sites (RS), a promoter controlling the kill switch gene expression is placed upstream of the kill switch gene.

to kill the particular plant organs, propagules, or even the whole plant. In Fig. 8.2, a variant of the scheme in Fig. 8.1 is shown to illustrate how a GURT circuit may be made into a kill switch. In this example, upon a specific stimulus (the input), the transgene cassette flanked by recognition site would be excised, which would place a kill switch gene and a promoter to control its expression in close proximity. The spatial and temporal expression of the kill switch would be exclusively controlled by its promoter. A constitutive promoter would enable the entire plant to be killed. A tissue-specific promoter could be used to ablate an organ or a cell type (such as pollen), whereas an inducible promoter may be used for adding another layer or signal for kill switch activation.

One could also envision a kill switch being used to kill progeny of hybridization to a wild or weedy relative of a transgenic crop plant. Such an approach could be used to assure that no interspecific transgenic hybrids survive. Instead of placing all the components in one plant, the controllability could be realized by a hybridization event, i.e. it would operate in trans (Oliver *et al.*, 1998, 1999a, b). In this case, the target gene and its promoter, as well as the blocking sequence in most cases, are constructed in one plant ('maintainer' line) and the trait switch gene is constitutively expressed in another ('inducer' line or 'activator' line) so that the GURT system would be triggered for the first time in F_1 hybrid plants. This design also contains a dual switch. In the age of synthetic biology, only the crop plant would be engineered. The inducer for a kill switch could be a transcription factor that is exclusively on the wild relative. The crop genome could be engineered to have a synthetic inducible promoter that would be activated only by the wild relative transcription factor. Upon hybridization, the kill switch on the hybrid would prevent persistence in the field.

8.6 Perspectives

As we see throughout this book, the view of gene flow from transgenic crops is a complex one. It spans governance, sociology, ecology, business, and agriculture. One view is that if biotechnology is at the heart of the matter, then perhaps biotechnology can also provide ecological and biosafety solutions in the form of transgene bioconfinement. After working in this field of transgenic plant gene flow and biosafety for more than 25 years, the author has become convinced that even the safest and best science for ecological safety may not be enough to assuage the fears of a sizeable populace who are skeptical of agricultural technology. There are many reasons for this skepticism that go far beyond this book. But, if we are mainly interested in actual assurances of safety, synthetic biology will provide many answers for advanced design in food systems in agriculture. Since the dawn of plant biotechnology, now spanning nearly four decades, engineered crops

have proven to be as safe as those that are not engineered (NASEM, 2016). Essentially, methods to transform them have been unchanged for decades and today's transgenic crops are relatively simple with regards to their novel gene constructs: typically a single transgene of interest and a selectable marker gene. Methods have long been available to create marker-free transgenic plants (Yau and Stewart, 2013). Synthetic biology will soon be developed to engineer more complex constructs in plants, such as synthetic circuits. The same technology that can yield increased function in plants via synthetic circuits for metabolic engineering can be used to make kill switches and other synthetic circuits for biosafety.

References

Andres, J., Blomeier, T. and Zurbriggen, M.D. (2019) Synthetic switches and regulatory circuits in plants. *Plant Physiology* 179(3), 862–884. DOI: 10.1104/pp.18.01362.

Belcher, M.S., Vuu, K.M., Zhou, A., Mansoori, N., Agosto Ramos, A. *et al*. (2020) Design of orthogonal regulatory systems for modulating gene expression in plants. *Nature Chemical Biology* 16(8), 857–865. DOI: 10.1038/s41589-020-0547-4.

Clark, M. and Maselko, M. (2020) Transgene biocontainment strategies for molecular farming. *Frontiers in Plant Science* 11, 210. DOI: 10.3389/fpls.2020.00210.

Ellstrand, N.C., Meirmans, P., Rong, J., Bartsch, D., Ghosh, A. *et al*. (2013) Introgression of crop alleles into wild or weedy populations. *Annual Review of Ecology, Evolution, and Systematics* 44(1), 325–345. DOI: 10.1146/annurev-ecolsys-110512-135840.

Fritsche, S., Klocko, A.L., Boron, A., Brunner, A.M. and Thorlby, G. (2018) Strategies for engineering reproductive sterility in plantation forests. *Frontiers in Plant Science* 9, 1671. DOI: 10.3389/fpls.2018.01671.

Gardner, T.S., Cantor, C. and Collins, J.J. (2000) Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403(6767), 339–342. DOI: 10.1038/35002131.

Hills, M.J., Hall, L., Arnison, P.G. and Good, A.G. (2007) Genetic use restriction technologies (GURTs): strategies to impede transgene movement. *Trends in Plant Science* 12(4), 177–183. DOI: 10.1016/j.tplants.2007.02.002.

Johnson, C.R., Millwood, R.J., Tang, Y., Gou, J., Sykes, R.W. *et al*. (2017) Field-grown miR156-transgenic switchgrass reproduction, yield, global gene expression analysis, and bioconfinememt. *Biotechnology for Biofuels* 10, 255. DOI: 10.1186/s13068-017-0939-1.

Kassaw, T.K., Donayre-Torres, A.J., Antunes, M.S., Morey, K.J. and Medford, J.I. (2018) Engineering synthetic regulatory circuits in plants. *Plant Science* 273, 13–22. DOI: 10.1016/j.plantsci.2018.04.005.

Klocko, A.L., Brunner, A.M., Huang, J., Meilan, R., Lu, H. *et al*. (2016) Containment of transgenic trees by suppression of LEAFY. *Nature Biotechnology* 34(9), 918–922. DOI: 10.1038/nbt.3636.

Klocko, A.L., Lu, H., Magnuson, A., Brunner, A.M., Ma, C. *et al*. (2018) Phenotypic expression and stability in a large-scale field study of genetically engineered poplars containing sexual containment transgenes. *Frontiers in Bioengineering and Biotechnology* 6, 100. DOI: 10.3389/fbioe.2018.00100.

Kwit, C., Moon, H.S., Warwick, S.I. and Stewart, C.N. Jr. (2011) Transgene introgression in crop relatives: molecular evidence and mitigation strategies. *Trends in Biotechnology* 29(6), 284–293. DOI: 10.1016/j.tibtech.2011.02.003.

Liu, W. and Stewart, C.N. Jr. (2015) Plant synthetic biology. *Trends in Plant Science* 20(5), 309–317. DOI: 10.1016/j.tplants.2015.02.004.

Liu, W. and Stewart, C.N. Jr. (2016) Plant synthetic promoters and transcription factors. *Current Opinion in Biotechnology* 37, 36–44. DOI: 10.1016/j.copbio.2015.10.001.

Lombardo, L. (2014) Gene use restriction technologies: a review. *Plant Biotechnology Journal* 12(8), 995–1005. DOI: 10.1111/pbi.12242.

Millwood, R.J., Moon, H.S., Poovaiah, C.R., Muthukumar, B., Rice, J.H. *et al*. (2016) Engineered selective plant male sterility through pollen-specific expression of the *Eco*RI restriction endonuclease. *Plant Biotechnology Journal* 14(5), 1281–1290. DOI: 10.1111/pbi.12493.

Moon, H.S., Li, Y. and Stewart, C.N. (2010) Keeping the genie in the bottle: transgene biocontainment by excision in pollen. *Trends in Biotechnology* 28(1), 3–8. DOI: 10.1016/j.tibtech.2009.09.008.

Moon, H.S., Abercrombie, L.L., Eda, S., Blanvillain, R., Thomson, J.G. *et al*. (2011) Transgene excision in pollen using a codon optimized serine resolvase CinH-*RS2* site-specific recombination system. *Plant Molecular Biology* 75(6), 621–631. DOI: 10.1007/s11103-011-9756-2.

NASEM (2016) *Genetically Engineered Crops: Experiences and Prospects*. National Academies of Sciences, Engineering, and Medicine (Committee on Genetically Engineered Crops). The National Academies Press, Washington, DC.

Ohadi, S., Hodnett, G., Rooney, W. and Bagavthiannan, M. (2017) Gene flow and its consequences in *Sorghum* spp. *Critical Reviews in Plant Sciences* 36(5-6), 367–385. DOI: 10.1080/07352689.2018.1446813.

Oliver, M.J., Quisenberry, J.E., Trolinder, N. and Keim, D.L. (1998) Control of plant gene expression. US patent 5723765.

Oliver, M.J., Quisenberry, J.E., Trolinder, N. and Keim, D.L. (1999a) Control of plant gene expression. US patent 5925808.

Oliver, M.J., Quisenberry, J.E., Trolinder, N. and Keim, D.L. (1999b) Control of plant gene expression. US patent 5977441.

Reichman, J.R., Watrud, L.S., Lee, E.H., Burdick, C.A., Bollman, M.A. *et al*. (2006) Establishment of transgenic herbicide-resistant creeping bentgrass (*Agrostis stolonifera* L.) in nonagronomic habitats. *Molecular Ecology* 15(13), 4243–4255. DOI: 10.1111/j.1365-294X.2006.03072.x.

Sang, Y., Millwood, R.J. Neal Stewart, C. and Stewart, C.N. Jr. (2013) Gene use restriction technologies for transgenic plant bioconfinement. *Plant Biotechnology Journal* 11(6), 649–658. DOI: 10.1111/pbi.12084.

Stewart, C.N., Halfhill, M.D. and Warwick, S.I. (2003) Transgene introgression from genetically modified crops to their wild relatives. *Nature Reviews Genetics* 4(10), 806–817. DOI: 10.1038/nrg1179.

Stirling, F., Bitzan, L., O'Keefe, S., Redfield, E., Oliver, J.W.K. *et al*. (2017) Rational design of evolutionarily stable microbial kill switches. *Molecular Cell* 68(4), 686–697. DOI: 10.1016/j.molcel.2017.10.033.

Van Acker, R.C., Szumgalski, A.R. and Friesen, L.F. (2007) The potential benefits, risks and costs of genetic use restriction technologies. *Canadian Journal of Plant Science* 87(4), 753–762. DOI: 10.4141/CJPS06033.

Visser, B., Meer, Vder, I., Louwaars, N., Beekwilder, J. and Eaton, D. (2001) The impact of 'terminator' technology. *Biotechnology and Development Monitor* 48, 8–12.

Warwick, S.I., Legere, A., Simard, M.J. and James, T. (2008) Do escaped transgenes persist in nature? The case of an herbicide resistance transgene in a weedy *Brassica rapa* population. *Molecular Ecology* 17(5), 1387–1395. DOI: 10.1111/j.1365-294X.2007.03567.x.

Watrud, L.S., Lee, E.H., Fairbrother, A., Burdick, C., Reichman, J.R. *et al*. (2004) Evidence for landscape-level, pollen-mediated gene flow from genetically modified creeping bentgrass with CP4 EPSPS as a marker. *Proceedings of the National Academy of Sciences of the United States of America* 101(40), 14533–14538. DOI: 10.1073/pnas.0405154101.

Yau, Y.Y. and Stewart, C.N. Jr. (2013) Less is more: strategies to remove marker genes from transgenic plants. *BMC Biotechnology* 13(1), 36. DOI: 10.1186/1472-6750-13-36.

9 Gene Flow Mitigation by Ecological Approaches

Yong-Bo Liu* and Xin-Yu Wang
State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing, China

Abstract

With an increased area of cultivating genetically modified (GM) plants worldwide, the ecological risks of transgenic plants released into the environment have caused concern. One of the risks is the occurrence of gene flow between GM plants and non-GM plants, including their wild relatives. Gene flow data from oilseed rape (*Brassica napus*), cotton (*Gossypium hirsutum*), maize (*Zea mays*), soybean (*Glycine max*), rice (*Oryza sativa*), and wheat (*Triticum aestivum*) indicate that the frequency of pollen-mediated gene flow is negatively related with distance between donor and recipient plants, and the frequency is relatively high in closely related species. We discuss five main ecological approaches to mitigate gene flow from GM plants to non-GM plants, including distance isolation, border or trap crops, barrier crops, agricultural practices, and through biological means. The required isolation distance has been adopted in managing GM crops in some countries, and cultivating tall crops, or border or trap crops, can decrease the requisite isolation distance to mitigate gene flow. Combining several approaches is more effective than a single approach in mitigating gene flow, because the frequency of pollen-mediated gene flow depends on plant genotype, flowering time, wind speed and direction, and other factors. Thus, in the framework of biosafety assessment of GM plants, mitigating the occurrence of gene flow between GM and non-GM plants is a key step to decrease the ecological risk of post-commercial cultivation of GM plants.
Keywords: ecological risk; gene flow; transgenic plants; wild relatives

9.1 Introduction

With the rapid development of transgenic biotechnology, genetically modified (GM) crops have been released into the environment since the 1990s. The area of cultivating GM plants worldwide was approximately 200 million hectares in 2018 (ISAAA, 2019). Currently, the most widespread commercial GM crops engineered for herbicide tolerance, virus resistance, and/or insect resistance are maize (*Zea mays*), soybean (*Glycine max*), cotton (*Gossypium hirsutum*), and oilseed rape (*Brassica napus*) (Kumar *et al.*, 2020).

As GM crops are cultivated worldwide, the ecological risks of GM plants released into the environment have led to widespread concern, particularly for potential gene flow between GM crops and their relatives (Liu and Stewart, 2019). Hybridization is a prerequisite to gene flow occurring, but not all hybridizations result in gene flow. The success of hybridization depends on the

*Corresponding author: liuyb@craes.org.cn; liu.yongbo@yahoo.com

 Gene Flow: Monitoring, Modeling and Mitigation (eds Wei Wei and C. Neal Stewart, Jr)
DOI: 10.1079/9781789247480.0009

sexual compatibility of crops and relatives (weedy, wild or crop) in the same area, with overlapping flowering period, appropriate pollen and seed dispersal, and successful fertilization. Many crops are domesticated from one or several wild relatives, so that most transgenic crops can hybridize with their wild relatives (Ellstrand *et al.*, 1999). Gene flow between transgenic plants and their wild relatives can sometimes result in the introgression of transgenes into the progeny (Ellstrand *et al.*, 2013; Liu *et al.*, 2013), which depends mainly on the fate of chromosome transmission and the relative fitness and competitiveness of progenies (Liu *et al.*, 2013).

Transgenes in crops may replace wild-type alleles in wild populations, which may reduce the genetic diversity of wild populations (Ellstrand, 2018). In September 2001, for example, a transgene from GM maize was found in native maize (Dalton and Diego, 2001), which aroused widespread concern. If hybrids have higher fitness than their parent species, hybrids could have increased invasiveness (Tiedje *et al.*, 1989; Li *et al.*, 2020). The hybrid and backcross generations formed between transgenic crops and their relatives have high fitness and seed dormancy, which could further increase the occurrence of gene flow (Di *et al.*, 2009). Backcross generations are often formed and easily establish as a false feral crop population (Liu *et al.*, 2010). The formed backcross generations can result in the introgression of transgenes into non-GM populations (Hansen *et al.*, 2001; Halfhill *et al.*, 2004). Studies have found molecular evidence of introgression from conventional maize to teosinte, the wild ancestor of maize (Doebley, 1984, 1990). Therefore, analyzing gene flow between transgenic crops and wild relatives is key for the ecological risk assessment and regulation of transgenic crops. Here, we describe some possible ecological approaches in mitigating gene flow from GM plants to non-GM plants, including their wild relatives.

9.2 Gene Flow Between Transgenic Plants and Their Relatives

Under natural conditions, the frequency of gene flow between transgenic oilseed rape (*B. napus*) and conventional oilseed rape is 0.2%–4.8%, and the maximum occurrence distance of gene flow is up to 3000 m (Scheffler *et al.*, 1993; Morris *et al.*, 1994; Scheffler *et al.*, 1995; Staniland *et al.*, 2000; Rieger *et al.*, 2002; Beckie *et al.*, 2003; Pu *et al.*, 2005b; Wei *et al.*, 2007; Cai *et al.*, 2008). The frequency of gene flow between transgenic oilseed rape and its wild relatives *Brassica juncea* and *B. rapa* is 0.2–16.17% and 7–37.2%, respectively, and the maximum occurrence distance is up to 400 m (Jørgensen *et al.*, 1994; Bing *et al.*, 1996; Halfhill *et al.*, 2002; Warwick *et al.*, 2003; Halfhill *et al.*, 2004; Pu *et al.*, 2005a, b; Song *et al.*, 2007; Liu *et al.*, 2010; Séguin-Swartz *et al.*, 2013; Liu *et al.*, 2018).The frequency of natural gene flow between transgenic *B. napus* and *Raphanus raphanistrum*, *Brassica carinata*, *Brassica nigra*, *Sinapis arvensis* and *Brassica oleracea* is less than 1% (Darmency *et al.*, 1998; Chèvre *et al.*, 2000), with the maximum occurrence distance up to 150 m (Pu *et al.*, 2005b; Séguin-Swartz *et al.*, 2013).

The frequency of natural gene flow between transgenic cotton and conventional cotton is from 0.03% to 19%, and the maximum occurrence distance is up to 1625 m (Umbeck *et al.*, 1991; Llewellyn and Fitt, 1996; Zhang *et al.*, 1997; Zhang and Guo, 2000; Shen *et al.*, 2001; Van Deynze *et al.*, 2005; Wang *et al.*, 2007). The frequency of gene flow between transgenic maize and conventional maize ranges from 0.318% to 82% under natural conditions, and the maximum occurrence distance is up to 300 m (Jemison and Vayda, 2001; Weekes *et al.*, 2007; Van De Wiel *et al.*, 2009; Viljoen and Chetty, 2011; Baltazar *et al.*, 2015; Liu *et al.*, 2015). Soybean is a self-pollinating crop and the natural outcrossing frequency is well below 1% (Chiang and Kiang, 1987; Ray *et al.*, 2003); the frequency of gene flow between GM soybean and conventional soybean is from 0.05% to 0.52%, ranging from 0% to 0.88% between GM soybean and wild soybean, with a maximum occurrence distance of 29 m (Nakayama and Yamaguchi, 2002; Ray *et al.*, 2003; Yoshimura *et al.*, 2006; Abud *et al.*, 2007; Kuroda *et al.*, 2008; Liu *et al.*, 2008; Wang and Li, 2011; Liu *et al.*, 2012).

The outcrossing frequency between GM rice and conventional rice is below 0.8% at close distances, and transgene frequencies are reduced to 0.01% with increasing distance from GM rice (e.g. 10 m) (Messeguer *et al.*, 2001, 2004; Rong *et al.*, 2005, 2007). The frequency

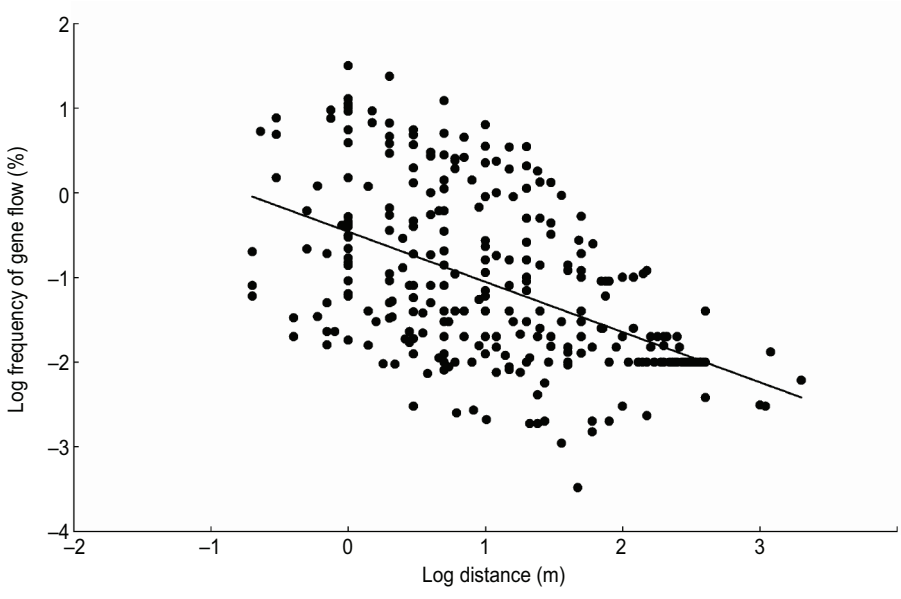


Fig. 9.1. Logarithmic correlation between frequency of gene flow and distance between six transgenic species and their relatives ($y = -0.59 \times -0.46$, $R^2 = 0.25$, $p < 0.01$). (Adapted from Wang and Liu, 2020)

of natural gene flow between GM rice and wild rice (*O. rufipogon* and *O. nivara*) ranged from 1% to 18% within 43.2 m and the maximum occurrence distance is up to 250 m (Song *et al.*, 2003; Chen *et al.*, 2004; Wang *et al.*, 2006). The frequency of gene flow between GM rice and weedy rice (*O. sativa* f. *spontanea*) ranged from 0.01% to 0.30% within 14 m under natural conditions (Zhang *et al.*, 2003; Chen *et al.*, 2004; Messeguer *et al.*, 2004; Shivrain *et al.*, 2007; Sun *et al.*, 2015). Hybrids between *O. sativa* and *Echinochloa caudata* were produced by hand pollination (Li *et al.*, 2003), but there was no natural hybridization observed between them. The frequency of natural gene flow between GM wheat (*Triticum aestivum*) and conventional wheat ranged from 0.01% to 8.5%, the main occurrence distance is within 80 m (Gatford *et al.*, 2006; Gaines *et al.*, 2007; Willenborg *et al.*, 2009; Beckie *et al.*, 2011; Rieben *et al.*, 2011; Miroshnichenko *et al.*, 2016).

Using the data of six crop species, the frequency of pollen-mediated gene flow was negatively related with distance between donor and recipient plants ($y = -0.59^{*} \times -0.46$, $R^2 = 0.25$, $p < 0.01$) (Fig. 9.1) (Wang and Liu, 2020), and the frequency was high between closely related species. The occurrence distance of gene flow depends mainly on plant genotype, flowering time, wind direction and speed, and planting density.

9.3 Ecological Approaches for Reducing Gene Flow

With the cultivation of GM plants worldwide, the ecological risks of possible gene flow from GM crops to their relatives or non-GM crops have raised concerns. Molecular technologies or biological methods to mitigate the occurrence of gene flow have been studied and reviewed (Liu *et al.*, 2013; Clark and Maselko, 2020). Ecological approaches may provide an alternative to biological methods, which may be more economical and environmentally friendly strategies to limit transgene flow. There are two paths for gene flow from crops to wild relatives: (i) seed movement; and (ii) pollen dispersal (Ellstrand, 2003). Seed-mediated gene flow occurs through seed movement by wind, water, animals, or other means during crop harvest and seed transportation. Pollen-mediated gene

flow occurs through pollination by bees and other pollination insects. Thus, to reduce the occurrence of gene flow from GM plants to their wild relatives or to non-GM plants, ecological approaches must focus on decreasing the likelihood of seed movement and pollen dispersal from GM plants.

9.3.1 Distance isolation

The most frequently used ecological approach to mitigate gene flow is to provide physical isolation between GM plants and non-GM plants (including wild relatives). Isolation of populations by various distances is used to decrease or prevent pollination from GM to non-GM plants. The distance requirement depends on the effective dispersal distance of pollen of GM plants, which varies by species and micro-environment. In China, the legal isolation distance for maize and oilseed rape is 300 m and 1000 m, respectively. In Europe, the legal isolation distance for maize and oilseed rape is 200 m and 400 m, respectively.

To know the exact distance required for isolation, numerous studies of the frequency of gene flow between GM plants and non-GM plants have been conducted under different environmental conditions. Most of the studies showed that the frequency of pollen-mediated gene flow is negatively related to the distance between pollen donor and recipient plants (Messeguer *et al.*, 2004; Van Deynze *et al.*, 2005; Weekes *et al.*, 2007; Cai *et al.*, 2008; Beckie *et al.*, 2011). For example, Scheffler *et al.* (1993) found that the natural frequency of gene flow from transgenic oilseed rape (64 m^2) to conventional oilseed rape was 1.5% at a distance of 1 m and decreased to 0.00033% at 47 m. To evaluate the effectiveness of imposing 200 m or 400 m isolation distances, Scheffler *et al.* (1995) conducted a small-scale trial in a 400 m^2 field of transgenic oilseed rape, and found that the frequency of hybridization with non-transgenic oilseed rape was 0.0156% at 200 m and 0.0038% at 400 m, respectively. In Europe, the allowable threshold of transgenic material in supposed non-transgenic seed is 0.9%. The isolation distance of 20 m between GM and conventional maize fields is enough to keep the adventitious presence of GM maize below the 0.9% threshold (Messeguer *et al.*, 2006).

While the imposition of distances between GM plants and non-GM plants can decrease gene flow, it is not easy in practice to mandate an isolation distance of dozens of meters or even hundreds of meters from GM plants, because this practice results in a waste of tens of thousands of square meters of field cultivation area. In addition, pollen dispersal is affected by climate conditions (wind direction and speed), the species of local pollinators, mating systems of target plants, and other micro-environmental conditions. At a landscape-level trial on 25–100 ha, the frequency of gene flow did not decrease with the distance from the GM pollen donor, showing a random pattern, and the frequency of gene flow was detected at 3000 m away from the pollen source (Rieger *et al.*, 2002). Implementing isolation distance is more effective in reducing gene flow in small fields than in large fields, particularly under commercial agricultural conditions (Hüsken and Dietz-Pfeilstetter, 2007).

In fact, there is no cut-off distance where the frequency of gene flow reaches zero (Hüsken and Dietz-Pfeilstetter, 2007). The isolation distance is likely to meet certain threshold levels (e.g. seed purity of 0.9%). Thus, other methods should be adopted to decrease the frequency of gene flow from GM plants to non-GM plants.

9.3.2 Border/trap crop

Border or trap crops may reduce the frequency of gene flow from transgenic plants to non-transgenic plants. A border of non-transgenic plants that are parental lines of transgenic plants is cultivated around a field of transgenic plants. These non-transgenic parental plants are called a border or trap crop and are used as a physical pollen barrier because they could compete with the transgenic pollen through dispersing non-transgenic pollen. To avoid producing transgenic seeds on non-transgenic plants, the border or trap crop will be destroyed after transgenic plants have flowered.

Non-transgenic oilseed rape (*Brassica napus*) has been used (planting width 15–30 m)

in Canada to reduce the spread of transgenic herbicide-resistant *B. napus* pollen, and results indicated that border crops are effective in reducing pollen-mediated gene flow in small fields (Staniland *et al.*, 2000). Creating a buffer zone with a trap crop is more effective in limiting gene flow than a short isolation distance (Morris *et al.*, 1994). Reboud (2003) found that planting several rows of non-transgenic oilseed rape plants between transgenic pollen donor and recipient fields was an efficient way to constrain gene flow. The trap crop plants have to be removed after flowering.

A border of non-transgenic crops may be useful to reduce the required isolation distance of transgenic crops. For example, in fields of cultivated non-GM and GM crops in Slovakia, Mihalčík *et al.* (2012) suggested that a 200 m isolation distance could be replaced by a border of 100 rows of non-GM plants. Quedas and Carvalho (2012) concluded that 24 rows of border non-GM plants could replace a 200 m isolation distance of GM plants in Portugal. Llewellyn and Fitt (1996) found that 20 m buffer zones limited the dispersal of transgenic cotton pollen in a small-scale field. However, some studies found that non-GM border crops could not be recommended as a reliable means to reduce the isolation distance of GM plants. Langhof *et al.* (2015) investigated the effectiveness of border crops and isolation distances in reducing gene flow frequency for securing the coexistence of non-GM and GM maize, and suggested that the isolation distance was the main reason for the reduction in frequency of gene flow. Thus, conducting a suitable isolation distance is effective in limiting the gene flow between GM and non-GM crops in small fields. Cultivating non-GM border crops may be a relatively cost-effective method in limiting gene flow if creating the isolation distance between pollen donor and recipient fields is impossible.

9.3.3 Barrier crop

Besides isolation distance and border crops, planting crops that differ to the GM crop could shorten the effective isolation distance in reducing the frequency of gene flow from GM plants to non-GM plants. Tall crops are often used as a pollen barrier to decrease the dispersal distance of pollen from GM plants.

The tall crop *Sorghum* as a barrier was effective in reducing gene flow between yellow-seed and white-seed maize: the frequency of gene flow was lower in a sorghum barrier site than in an open control site, with a gene flow frequency of 1.04% and 9.35%, respectively (Liu *et al.*, 2015). This indicates that a tall crop could be used as a barrier to decrease the frequency of gene flow from transgenic to conventional crops. Separation nets and separation crops, *Sorghum bicolor*, *Zea mays* and *Lycopersicon esculentum*, were employed as pollen barriers to reduce gene flow between transgenic and non-transgenic cotton, and the frequency of gene flow was the lowest with a separation net of 90 holes/cm^2 and *S. bicolor* as a separation crop (Yan *et al.*, 2018). However, tall crops are not always effective in limiting pollen dispersal and gene flow. For example, with a height up to 4 m, a hemp strip was used as a pollen barrier but was ineffective in limiting the pollen dispersal of transgenic beet (Saeglitz *et al.*, 2010). Comparing the effectiveness of a tall sunflower crop (*Helianthus annuus* L.) and a short clover–grass crop (*Trifolium pratense* L. and *Lolium* spp.) in reducing gene flow between transgenic and non-transgenic maize, Langhof *et al.* (2008) found no difference in gene flow frequency and concluded that employing sunflower as a tall crop barrier was not an appropriate strategy.

The height of a tall crop as a barrier might be the key factor in reducing the pollen-mediated gene flow. In addition, the effectiveness of tall crops might be affected by environmental conditions. For example, changes of wind intensity and speed strongly affect the effectiveness of tall crops as pollen barriers in limiting gene flow (Devos *et al.*, 2005), because pollen can be carried by wind flow to high-altitude areas and fall on plants far away from the pollen donor of transgenic plants. Employing large areas of tall crops could reduce gene flow frequency effectively, but it might be unrealistic to plant such large areas of tall crops in agricultural production because the borders of tall crops are larger in area than the fields of transgenic crops (Hokanson *et al.*, 1997).

9.3.4 Agricultural practices

Appropriate agricultural practices are effective in decreasing the frequency of gene flow from GM plants to non-GM plants. For example, delayed planting of crops or planting in advance can separate the flowering time between transgenic crops and non-transgenic crops or wild relatives so that the frequency of gene flow would be decreased. The frequency of gene flow from transgenic crops to non-transgenic maize was reduced by 50% when the sowing time of transgenic and non-transgenic maize differed by 1 week; changing sowing dates by 3 weeks reduced gene flow frequency by 75% (Brookes and Barfoot, 2004; Ortega-Molina, 2004). However, it is difficult to separate the flowering time of crops that are photoperiod-sensitive if the changed time of sowing crops is not long enough.

Moreover, the alteration of sowing time or delaying or advancing the planting date may run the risk of experiencing poor weather and decreased yields. For example, maize may be damaged by frost if sown too early. Hence, delays in sowing seeds or sowing in advance might be at the expense of crop yield. In addition, the alteration of sowing time may affect the economic income of farmers, because the market price varies with the harvesting season of crops.

Besides pollen-mediated gene flow, seed-mediated gene flow is another pathway for introducing transgenes into other fields through seed movement by animals or machines. The soil bank of transgenic seeds in fields is also a pathway of seed dispersal in time, because seeds could survive several years in soil (Liu *et al.*, 2013). Seed-mediated gene flow could be controlled by agricultural practices, such as completely cleaning harvest machines and loading areas of trucks and ships (Pascher *et al.*, 2017), conducting crop rotation, delaying soil cultivation, and regular tillage. Models simulated the effects of cropping systems with different agricultural practices on gene flow and population dynamics for oilseed rape volunteers (Colbach, 2009) and transgenic hybrids (Hooftman *et al.*, 2008), and suggested that the agricultural practices significantly constrain the occurrence of transgenic volunteers in fields and reduce the likelihood of seed-mediated gene flow.

9.3.5 Biological means

Besides the physical means discussed above, biological means have been developed to decrease the occurrence of risk of gene flow from GM plants to non-GM plants (Gressel, 2000; Daniell, 2002), including chloroplast transformation, male sterility, apomixis, cleistogamy, and ploidy level.

Chloroplast transformation was first developed for tobacco (Svab *et al.*, 1990) and subsequently used in other crops. It can greatly reduce the occurrence of pollen-mediated gene flow, because chloroplasts are maternally inherited in angiosperms. This technology can enable high-level and stable expression of transgenes in plants (Ye *et al.*, 2000). However, the maternal inheritance of chloroplasts occurs rarely in some species and chloroplast genes sometimes transfer into the nuclear genome, although the transfer ratio is extremely low (Huang *et al.*, 2003).

Male sterility is a relatively common means because it can greatly reduce the pollen-mediated gene flow from GM crop to non-GM crop or to wild relatives. Most male sterility systems are not always perfect, because some systems usually have a certain percentage of fertile pollen (Hvarleva *et al.*, 2009), which leads to the stigmas of non-GM crops or wild relatives being pollinated by the fertile pollen of GM crops (Chèvre *et al.*, 2000).

Apomixis can be used as a biological confinement method of pollen-mediated gene flow, because apomictic plants do not require pollination to produce progenies. But the application of apomixes in most transgenic crops commercially will still take many years.

Cleistogamy is an effective tool to reduce pollen-mediated gene flow (Colbach *et al.*, 2000), but it is difficult to create cleistogamous plants for allogamous or cross-fertilized crops (e.g. maize) (Lu, 2003), and pollinators can encourage plant petals to open and contribute to successful pollination (Leflon *et al.*, 2010).

If GM plants and recipient non-GM plants are of different ploidy, the risks of transgenes

spreading to non-GM plants, including wild relatives, are relatively low (Lu and Xia, 2011). Genomes with different ploidy are generally incompatible but the incompatibility is not strict.

The means discussed above have advantages and weaknesses in reducing gene flow from GM plants to non-GM plants, although some of these means has been tested for long-term effectiveness in reducing gene flow frequency. The availability of these means could contribute to further regulation of GM crops. The right combination of several physical and biological means will be more effective than one sole method.

9.4 Conclusion

The frequency of pollen-mediated gene flow is negatively related with distance between donor and recipient plants, and it is affected by plant genotype, flowering time, wind speed and direction, and other factors. The results support the design of isolation distance to effectively reduce the frequency of gene flow. However, the isolation distance is not always effective in controlling the occurrence of gene flow. It is still necessary to study multi-factor gene flow prediction models for reducing the frequency of gene flow between transgenic plants and their relatives. In addition, in the context of ecological risk assessment and regulation of transgenic crops, policy makers could consider a 'partition management' strategy based on the frequency of gene flow between GM crops and their relatives (Dong *et al*., 2018). Using a systematized database, Sánchez *et al*. (2016) conducted a national-scale study of the likelihood of outcrossing between 11 GM crops and vascular plants in Chile. Ecological approaches should be considered in the environment risk assessment for GM crops released into the environment.

References

Abud, S., De Souza, P.I.M., Vianna, G.R., Leonardecz, E., Moreira, C.T. *et al*. (2007) Gene flow from transgenic to nontransgenic soybean plants in the Cerrado region of Brazil. *Genetics and Molecular Research* 6(2), 445–452. DOI: 10.1590/S1415-47572007000600029.

Baltazar, B.M., Castro Espinoza, L., Espinoza Banda, A., de la Fuente Martínez, J.M., Garzón Tiznado, J.A. *et al*. (2015) Pollen-mediated gene flow in maize: implications for isolation requirements and coexistence in Mexico, the center of origin of maize. *PloS one* 10(7), e0131549. DOI: 10.1371/journal.pone.0131549.

Beckie, H.J., Warwick, S.I., Nair, H. and Séguin-Swartz, G. (2003) Gene flow in commercial fields of herbicide-resistant canola (*Brassica napus*). *Ecological Applications* 13, 1276–1294. DOI: 10.1890/02-5231.

Beckie, H.J., Warwick, S.I., Sauder, C.A., Hall, L.M. and Harker, K.N. (2011) Pollen-mediated gene flow in commercial fields of spring wheat in western Canada. *Crop Science* 51, 306–313. DOI: 10.2135/cropsci2010.03.0176.

Bing, D.J., Downey, R.K. and Rakow, G.F.W. (1996) Hybridizations among *Brassica napus*, *B. rapa* and *B. juncea* and their two weedy relatives *B. nigra* and *Sinapis arvensis* under open pollination conditions in the field. *Plant Breeding* 115, 470–473.

Brookes, G. and Barfoot, P. (2004) Co-existence of GM and non GM crops: case study of maize grown in Spain. *International Pest Control* 46, 20–21.

Cai, L., Zhou, B.-W., Guo, X.-L., Dong, C.-H. and Hu, X.-J. (2008) Pollen-mediated gene flow in Chinese commercial fields of glufosinate-resistant canola (*Brassica napus*). *Chinese Science Bulletin* 53, 2333.

Chen, L.-J., Lee, D.-S., Song, Z.-P., Suh, H.-S. and Lu, B.-R. (2004) Gene flow from cultivated rice (*Oryza sativa*) to its weedy and wild relatives. *Annals of Botany* 93(1), 67–73. DOI: 10.1093/aob/mch006.

Chèvre, A.M., Eber, F., Darmency, H., Fleury, A. and Picault, H. (2000) Assessment of interspecific hybridization between transgenic oilseed rape and wild radish under normal agronomic conditions. *Theoretical and Applied Genetics* 100, 1233–1239. DOI: 10.1007/s001220051429.

Chiang, Y.-C. and Kiang, Y.-T. (1987) Geometric position of genotypes, honeybee foraging patterns and outcrossing in soybean. *Botanical Bulletin of Academia Sinica* 28, 1–11.

Clark, M. and Maselko, M. (2020) Transgene biocontainment strategies for molecular farming. *Frontiers in Plant Science* 11, 210–210. DOI: 10.3389/fpls.2020.00210.

Colbach, N. (2009) How to model and simulate the effects of cropping systems on population dynamics and gene flow at the landscape level: example of oilseed rape volunteers and their role for coexistence of GM and non-GM crops. *Environmental Science and Pollution Research* 16(3), 348–360. DOI: 10.1007/s11356-008-0080-6.

Colbach, N., Meynard, J., Clermontdauphin, C. and Champolivier, J. (2000) GeneSys-Colza: a model for the medium-term and long-term effects of cultivation systems on the flow of genes between rape fields and new growth in a given agricultural area. *OCL – Oilseeds and fats, Crops and Lipids* 7, 329–340.

Dalton, R. and Diego, S. (2001) Transgenic corn found growing in Mexico. *Nature* 413(6854), 337–337. DOI: 10.1038/35096714.

Daniell, H. (2002) Molecular strategies for gene containment in transgenic crops. *Nature Biotechnology* 20(6), 581–586. DOI: 10.1038/nbt0602-581.

Darmency, H., Lefol, E. and Fleury, A. (1998) Spontaneous hybridizations between oilseed rape and wild radish. *Molecular Ecology* 7(11), 1467–1473. DOI: 10.1046/j.1365-294x.1998.00464.x.

Devos, Y., Reheul, D. and De Schrijver, A. (2005) The co-existence between transgenic and non-transgenic maize in the European Union: a focus on pollen flow and cross-fertilization. *Environmental Biosafety Research* 4(2), 71–87. DOI: 10.1051/ebr:2005013.

Di, K., Stewart, C.N., Wei, W., Shen, B.-cheng., Tang, Z.-X. *et al*. (2009) Fitness and maternal effects in hybrids formed between transgenic oilseed rape (*Brassica napus* L.) and wild brown mustard [*B. juncea* (L.) Czern et Coss.] in the field. *Pest Management Science* 65(7), 753–760. DOI: 10.1002/ps.1749.

Doebley, J.F. (1984) Maize introgression into teosinte – A reappraisal. *Annals of the Missouri Botanical Garden* 71(4), 1100. DOI: 10.2307/2399247.

Doebley, J. (1990) Molecular evidence for gene flow among Zea species. *BioScience* 40(6), 443–448. DOI: 10.2307/1311391.

Dong, J.-J., Zhang, M.-G., Wei, W., Ma, K.-P. and Wang, Y.-H. (2018) GIS assessment of the risk of gene flow from *Brassica napus* to its wild relatives in China. *Environmental Monitoring and Assessment* 190(7), 405. DOI: 10.1007/s10661-018-6753-9.

Ellstrand, N.C. (2003) Current knowledge of gene flow in plants: implications for transgene flow. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 358(1434), 1163–1170. DOI: 10.1098/rstb.2003.1299.

Ellstrand, N.C. (2018) Does introgression of crop alleles into wild and weedy living populations create cryptic insitu germplasm banks? *Molecular Ecology* 27(1), 38–40. DOI: 10.1111/mec.14390.

Ellstrand, N.C., Prentice, H.C. and Hancock, J.F. (1999) Gene flow and introgression from domesticated plants into their wild relatives. *Annual Review of Ecology and Systematics* 30(1), 539–563. DOI: 10.1146/annurev.ecolsys.30.1.539.

Ellstrand, N.C., Meirmans, P., Rong, J., Bartsch, D., Ghosh, A. *et al*. (2013) Introgression of crop alleles into wild or weedy populations. *Annual Review of Ecology, Evolution, and Systematics* 44(1), 325–345. DOI: 10.1146/annurev-ecolsys-110512-135840.

Gaines, T.A., Byrne, P.F., Westra, P., Nissen, S.J., Henry, W.B. *et al*. (2007) An empirically derived model of field-scale gene flow in winter wheat. *Crop Science* 47(6), 2308–2316. DOI: 10.2135/cropsci2007.05.0248.

Gatford, K.T., Basri, Z., Edlington, J., Lloyd, J., Qureshi, J.A. *et al*. (2006) Gene flow from transgenic wheat and barley under field conditions. *Euphytica* 151(3), 383–391. DOI: 10.1007/s10681-006-9160-1.

Gressel, J. (2000) Molecular biology of weed control. *Transgenic Research* 9(4-5), 355–382. DOI: 10.1023/a:1008946628406.

Halfhill, M.D., Millwood, R.J., Raymer, P.L. and Stewart, C.N. Jr. (2002) *Bt*-transgenic oilseed rape hybridization with its weedy relative, *Brassica rapa*. *Environmental Biosafety Research* 1, 19–28. DOI: 10.1051/ebr:2002002.

Halfhill, M.D., Zhu, B., Warwick, S.I., Raymer, P.L., Millwood, R.J. *et al*. (2004) Hybridization and backcrossing between transgenic oilseed rape and two related weed species under field conditions. *Environmental Biosafety Research* 3(2), 73–81. DOI: 10.1051/ebr:2004007.

Hansen, L.B., Siegismund, H.R. and Jørgensen, R.B. (2001) Introgression between oilseed rape (*Brassica napus* L.) and its weedy relative *B. rapa* L. in a natural population. *Genetic Resources and Crop Evolution* 48(6), 621–627. DOI: 10.1023/A:1013825816443.

Hokanson, S.C., Grumet, R. and Hancock, J.F. (1997) Effect of border rows and trap/donor ratios on pollen-mediated gene movement. *Ecological Applications* 7(3), 1075–1081. DOI: 10.1890/1051-0761(1997)007[1075:EOBRAT]2.0.CO;2.

Hooftman, D.A.P., Oostermeijer, J.G.B., Marquard, E. and den Nijs, H.C.M. (2008) Modelling the consequences of crop–wild relative gene flow: a sensitivity analysis of the effects of outcrossing rates and hybrid vigour breakdown in *Lactuca*. *Journal of Applied Ecology* 45(4), 1094–1103. DOI: 10.1111/j.1365-2664.2008.01508.x.

Huang, C.Y., Ayliffe, M.A. and Timmis, J.N. (2003) Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature* 422(6927), 72–76. DOI: 10.1038/nature01435.

Hüsken, A. and Dietz-Pfeilstetter, A. (2007) Pollen-mediated intraspecific gene flow from herbicide resistant oilseed rape (*Brassica napus* L.). *Transgenic Research* 16(5), 557–569. DOI: 10.1007/s11248-007-9078-y.

Hvarleva, T., Hristova, M., Bakalova, A., Hristov, M. and Atanassov, I. (2009) CMS lines for evaluation of pollen flow in sunflower relevance for transgene flow mitigation. *Biotechnology & Biotechnological Equipment* 23, 1309–1315. DOI: 10.1080/13102818.2009.10817659.

ISAAA (2019) *Global Status of Commercialized Biotech/GM Crops: 2019*. ISAAA Brief No. 55. International Service for the Acquisition of Agri-biotech Applications, Ithaca, New York.

Jemison, J.M. and Vayda, M.E. (2001) Cross pollination from genetically engineered corn: wind transport and seed source. *Agbioforum* 4, 87–92.

Jørgensen, R.B., Andersen, B., Landbo, L. and Mikkelsen, T.R. (1994) Spontaneous hybridization between oilseed rape (*Brassica napus*) and weedy relatives. *ISHS Brassica Symposium-IX Crucifer Genetics Workshop* 407, 193–200.

Kumar, K., Gambhir, G., Dass, A., Tripathi, A.K., Singh, A. *et al*. (2020) Genetically modified crops: current status and future prospects. *Planta* 251(4), 1–27. DOI: 10.1007/s00425-020-03372-8.

Kuroda, Y., Kaga, A., Tomooka, N. and Vaughan, D.A. (2008) Gene flow and genetic structure of wild soybean (*Glycine soja*) in Japan. *Crop Science* 48(3), 1071–1079. DOI: 10.2135/cropsci2007.09.0496.

Langhof, M., Hommel, B., Hüsken, A., Schiemann, J., Wehling, P. *et al*. (2008) Coexistence in maize: do nonmaize buffer zones reduce gene flow between maize fields? *Crop Science* 48(1), 305–316. DOI: 10.2135/cropsci2007.04.0189.

Langhof, M., Gabriel, D. and Rühl, G. (2015) Combination approach of border rows and isolation distance for securing coexistence of non-genetically modified and genetically modified maize. *Crop Science* 55(4), 1818–1826. DOI: 10.2135/cropsci2014.07.0523.

Leflon, M., Hüsken, A., Njontie, C., Kightley, S., Pendergrast, D. *et al*. (2010) Stability of the cleistogamous trait during the flowering period of oilseed rape. *Plant Breeding* 129(1), 13–18. DOI: 10.1111/j.1439-0523.2009.01645.x.

Li, H.-M., Zhao, F.-W., Li, A.-G., Bai, L. and Zhang, J. (2003) Study on seed set and heteroses of distant cross of upland rice (*Oryza sativa*) × long-awn-barnyard-grass (*Echinochloa caudata*). [In Chinese with English abstract]. *Acta Agriculturae Nucleatae Sinica* 17, 11–15. DOI: 10.3969/j.issn.1000-8551.2003.01.003.

Li, Y., Hallerman, E.M., Wu, K. and Peng, Y. (2020) Insect-resistant genetically engineered crops in China: development, application, and prospects for use. *Annual Review of Entomology* 65, 273–292. DOI: 10.1146/annurev-ento-011019-025039.

Liu, J., Zhou, B., Yang, C.-Y., Li, Y.-R. and Jiang, L.-X. (2012) Gene flowing of genetically modified glyphosate-resistant soybean with *EPSPS*. [In Chinese with English abstract.] *Soybean Science* 31(517–521). DOI: 10.3969/j.issn.1000-9841.2012.04.002.

Liu, Q., Li, X., Liu, Z. and Lei, B. (2008) Study on gene flow of roundup soybean with *CP4 EPSPS*. [In Chinese with English abstract]. *Jiangsu Journal of Agriculture Science* 28, 84–87.

Liu, Y.-B. and Stewart, C.N. Jr. (2019) An exposure pathway-based risk assessment system for GM plants. *Plant Biotechnology Journal* 17(10), 1859–1861. DOI: 10.1111/pbi.13146.

Liu, Y.-B., Wei, W., Ma, K.-P. and Darmency, H. (2010) Backcrosses to *Brassica napus* of hybrids between *B. juncea* and *B. napus* as a source of herbicide-resistant volunteer-like feral populations. *Plant Science* 179(5), 459–465. DOI: 10.1016/j.plantsci.2010.07.005.

Liu, Y.-B., Wei, W., Ma, K.-P., Li, J., Liang, Y. *et al*. (2013) Consequences of gene flow between oilseed rape (*Brassica napus*) and its relatives. *Plant Science* 211, 42–51. DOI: 10.1016/j.plantsci.2013.07.002.

Liu, Y.-B., Chen, F.J., Guan, X. and Li, J.S. (2015) High crop barrier reduces gene flow from transgenic to conventional maize in large fields. *European Journal of Agronomy* 71, 135–140. DOI: 10.1016/j.eja.2015.09.005.

Liu, Y.-B., Stewart, C.N. Jr., Li, J. and Wei, W. (2018) One species to another: sympatric *Bt* transgene gene flow from *Brassica napus* alters the reproductive strategy of wild relative *Brassica juncea* under herbivore treatment. *Annals of Botany* 122(4), 617–625. DOI: 10.1093/aob/mcy096.

Llewellyn, D. and Fitt, G. (1996) Pollen dispersal from two field trials of transgenic cotton in the Namoi Valley, Australia. *Molecular Breeding* 2, 157–166. DOI: 10.1007/BF00441430.

Lu, B.-R. (2003) Transgene containment by molecular means – is it possible and cost effective? *Environmental Biosafety Research* 2(1), 3–8. DOI: 10.1051/ebr/2003000.

Lu, B.-R. and Xia, H. (2011) Environmental biosatety of transgenic plants: research and assessment of transgene escape and its potential ecological impacts. [In Chinese with English abstract.] *Chinese Bulletin of Life Science* 23, 186–194. DOI: CNKI:SUN:SMKX.0.2011-02-009.

Messeguer, J., Fogher, C., Guiderdoni, E., Marfa, V. and Catala, M.M. (2001) Field assessments of gene flow from transgenic to cultivated rice (Oryza sativa L.) using a herbicide resistance gene as tracer marker. *Theoretical and Applied Genetics* 103, 1151–1159. DOI: 10.1007/s001220100713.

Messeguer, J., Marfa, V., Catala, M.M., Guiderdoni, E. and Mele, E. (2004) A field study of pollen-mediated gene flow from Mediterranean GM rice to conventional rice and the red rice weed. *Molecular Breeding* 13, 103–112. DOI: 10.1023/B:MOLB.0000012285.39859.9d.

Messeguer, J., Peñas, G., Ballester, J., Bas, M., Serra, J. *et al*. (2006) Pollen-mediated gene flow in maize in real situations of coexistence. *Plant Biotechnology Journal* 4(6), 633–645. DOI: 10.1111/j.1467-7652.2006.00207.x.

Mihalčík, P., Hrčková, K., Singer, M., Plačková, A. and Kraic, J. (2012) Effect of MON 810 cultivation and prevention to adventitious presence in non-GM fields: a case study in Slovakia. *Plant Protection Science* 48, S11–S15. DOI: 10.17221/58/2012-PPS.

Miroshnichenko, D., Pushin, A. and Dolgov, S. (2016) Assessment of the pollen-mediated transgene flow from the plants of herbicide resistant wheat to conventional wheat (*Triticum aestivum* L.). *Euphytica* 209, 71–84. DOI: 10.1007/s10681-016-1637-y.

Morris, W.F., Kareiva, P.M. and Raymer, P.L. (1994) Do barren zones and pollen traps reduce gene escape from transgenic crops? *Ecological Applications* 4, 157–165. DOI: 10.2307/1942125.

Nakayama, Y. and Yamaguchi, H. (2002) Natural hybridization in wild soybean (*Glycine* ssp. *soja*) by pollen flow from cultivated soybean (*Glycine* ssp. *max*) in a designed population. *Weed Biology and Management* 2(1), 25–30. DOI: 10.1046/j.1445-6664.2002.00043.x.

Ortega-Molina, J. (2004) *Results of the studies into the coexistence of genetically modified and conventional maize*. COPA-COGECA colloquy on the co-existence and thresholds of adventitious presence on GMOs in conventional seeds. Committee of Professional Organisations of the European Union (COPA) / General committee of Agricultural Coopeeration of the European Union (COGECA), Brussels.

Pascher, K., Hainz-Renetzeder, C., Gollmann, G. and Schneeweiss, G.M. (2017) Spillage of viable seeds of oilseed rape along transportation routes: ecological risk assessment and perspectives on management efforts. *Frontiers in Ecology and Evolution* 5, 1–6. DOI: 10.3389/fevo.2017.00104.

Pu, H.-M., Qi, C.-K., Zhang, J.-F., Fu, S.-Z. and Gao, J.-Q. (2005a) The studies on gene flow from GM herbicide-tolerant rapeseed to cruciferous weeds. *Acta Ecologica Sinica* 25, 910–916. [In Chinese with English abstract.] DOI: 10.3321/j.issn:1000-0933.2005.04.037.

Pu, H.-M., Qi, C.-K., Zhang, J.-F., Fu, S.-Z. and Gao, J.-Q. (2005b) Studies on the gene flow from herbicide-tolerant GM rapeseed to its close relative crops. *Acta Ecologica Sinica* 25, 581–588. [In Chinese with English abstract.] DOI: 10.3321/j.issn:1000-0933.2005.03.027.

Quedas, F. and Carvalho, P. (2012) A quinquennium of coexistence in Portugal. *AgBioForum* 15, 1–9.

Ray, J.D., Kilen, T.C., Abel, C.A. and Paris, R.L. (2003) Soybean natural cross-pollination rates under field conditions. *Environmental Biosafety Research* 2, 133–138. DOI: 10.1051/ebr:2003005.

Reboud, X. (2003) Effect of a gap on gene flow between otherwise adjacent transgenic *Brassica napus* crops. *Theoretical Applied Genetics* 106(6), 1048–1058. DOI: 10.1007/s00122-002-1142-7.

Rieben, S., Kalinina, O., Schmid, B. and Zeller, S.L. (2011) Gene flow in genetically modified wheat. *PLoS ONE* 6(12), e29730. DOI: 10.1371/journal.pone.0029730.

Rieger, M.A., Lamond, M., Preston, C., Powles, S.B. and Roush, R.T. (2002) Pollen-mediated movement of herbicide resistance between commercial canola fields. *Science* 296(5577), 2386–2388. DOI: 10.1126/science.1071682.

Rong, J., Song, Z.-P., Su, J., Xia, H. and Lu, B.-R. (2005) Low frequency of transgene flow from *Bt/CpTI* rice to its nontransgenic counterparts planted at close spacing. *New Phytologist* 168, 559–566. DOI: 10.1111/j.1469-8137.2005.01539.x.

Rong, J., Lu, B.-R., Song, Z.P., Su, J. and Snow, A.A. (2007) Dramatic reduction of crop-to-crop gene flow within a short distance from transgenic rice fields. *New Phytologist* 173, 346–353. DOI: 10.1111/j.1469-8137.2006.01906.x.

Saeglitz, C., Pohl, M. and Bartsch, D. (2010) Monitoring gene flow from transgenic sugar beet using cytoplasmic male-sterile bait plants. *Molecular Ecology* 9, 2035–2040. DOI: 10.1046/j.1365-294X.2000.01109.x.

Sanchez, M.A., Cid, P., Navarrete, H., Aguirre, C., Chacon, G. *et al*. (2016) Outcrossing potential between 11 important genetically modified crops and the Chilean vascular flora. *Plant Biotechnology Journal* 14(2), 625–637. DOI: 10.1111/pbi.12408.

Scheffler, J.A., Parkinson, R. and Dale, P.J. (1993) Frequency and distance of pollen dispersal from transgenic oilseed rape (*Brassica napus*). *Transgenic Research* 2, 356–364. DOI: 10.1007/BF01976177.

Scheffler, J.A., Parkinson, R. and Dale, P.J. (1995) Evaluating the effectiveness of isolation distances for field plots of oilseed rape (*Brassica napus*) using a herbicide-resistance transgene as a selectable marker. *Plant Breeding* 114, 317–321. DOI: 10.1111/j.1439-0523.1995.tb01241.x.

Séguin-Swartz, G., Beckie, H.J., Warwick, S.I., Roslinsky, V., Nettleton, J.A. *et al*. (2013) Pollen-mediated gene flow between glyphosate-resistant *Brassica napus* canola and *B. juncea* and *B. carinata* mustard crops under large-scale field conditions in Saskatchewan. *Canadian Journal of Plant Science* 93(6), 1083–1087. DOI: 10.4141/cjps2013-129.

Shen, F.-F., Yu, Y.-J., Zhang, X.-K., Bi, J.-J. and Yin, C.-Y. (2001) *Bt* gene flow of transgenic cotton [In Chinese with English abstract.]. *Acta Genetica Sinica* 28, 562–567. DOI: CNKI:SUN:YCXB.0.2001-06-010.

Shivrain, V.K., Burgos, N.R., Anders, M.M., Rajguru, S.N. and Moore, J. (2007) Gene flow between Clearfield™ rice and red rice. *Crop Protection* 26, 349–356. DOI: 10.1016/j.cropro.2005.09.019.

Song, Z.-P., Lu, B.-R., Zhu, Y.-G. and Chen, J.-K. (2003) Gene flow from cultivated rice to the wild species *Oryza rufipogon* under experimental field conditions. *New Phytologist* 157(3), 657–665. DOI: 10.1046/j.1469-8137.2003.00699.x.

Song, X.-L., Huangpu, C.-H. and Qiang, S. (2007) Gene flow from transgenic glufosinate- or glyphosate-tolerant oilseed rape to wild rape. [In Chinese with English abstract.] *Journal of Plant Ecology [Chinese Version]* 31, 729–737.

Staniland, B.K., McVetty, P.B., Friesen, L.F., Yarrow, S. and Freyssinet, G. (2000) Effectiveness of border areas in confining the spread of transgenic *Brassica napus* pollen. *Canadian Journal of Plant Science* 80, 521–526. DOI: 10.4141/P99-117.

Sun, G., Dai, W., Cui, R., Qiang, S. and Song, X. (2015) Gene flow from glufosinate-resistant transgenic hybrid rice Xiang 125S/Bar68-1 to weedy rice and cultivated rice under different experimental designs. *Euphytica* 204, 211–227. DOI: 10.1007/s10681-015-1370-y.

Svab, Z., Hajdukiewicz, P. and Maliga, P. (1990) Stable transformation of plastids in higher plants. *Proceedings of the National Academy of Sciences* 87(21), 8526–8530. DOI: 10.1073/pnas.87.21.8526.

Tiedje, J.M., Colwell, R.K., Grossman, Y.L., Hodson, R.E. and Lenski, R.E. (1989) The planned introduction of genetically engineered organisms: ecological considerations and recommendations. *Ecology* 70, 298–315. DOI: 10.2307/1937535.

Umbeck, P.F., Barton, K.A., Nordheim, E.V., McCarty, J.C. and Parrott, W.L. (1991) Degree of pollen dispersal by insects from a field test of genetically engineered cotton. *Journal of Economic Entomology* 84, 1943–1950. DOI: 10.1093/jee/84.6.1943.

Van De Wiel, C.C.M., Groeneveld, R.M.W., Dolstra, O., Kok, E.J. and Scholtens, I.M.J. (2009) Pollen-mediated gene flow in maize tested for coexistence of GM and non-GM crops in the Netherlands: effect of isolation distances between fields. *NJAS-Wageningen Journal of Life Sciences* 56, 405–423. DOI: 10.1016/S1573-5214(09)80007-9.

Van Deynze, A.E., Sundstrom, F.J. and Bradford, K.J. (2005) Pollen-mediated gene flow in California cotton depends on pollinator activity. *Crop Science* 45, 1565–1570. DOI: 10.2135/cropsci2004.0463.

Viljoen, C. and Chetty, L. (2011) A case study of GM maize gene flow in South Africa. *Environmental Sciences Europe* 23, 8. DOI: 10.1186/2190-4715-23-8.

Wang, C.Y., Liu, Y., Zhou, J., Chen, J.Q. and Qin, P. (2007) Monitoring of pollen-mediated gene flow from transgenic *Bt* cotton. [In Chinese with English abstract.] *Chinese Journal of Applied Ecology* 18, 801–806.

Wang, F., Yuan, Q.-H., Shi, L., Qian, Q., Liu, W.-G. *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 Biotechnology Journal* 4(6), 667–676. DOI: 10.1111/j.1467-7652.2006.00210.x.

Wang, K.-J. and Li, X.-H. (2011) Interspecific gene flow and the origin of semi-wild soybean revealed by capturing the natural occurrence of introgression between wild and cultivated soybean populations. *Plant Breeding* 130, 117–127. DOI: 10.1111/j.1439-0523.2010.01815.x.

Wang, X.-Y. and Liu, Y.-B. (2020) Research advances in gene flow between transgenic plants and their relatives. [In Chinese with English abstract.]. *Chinese Journal of Applied Ecology* 31, 3207–3215. DOI: 10.13287/j.1001-9332.202009.025.

Warwick, S.I., Simard, M.-J., Légère, A., Beckie, H.J., Braun, L. *et al*. (2003) Hybridization between transgenic *Brassica napus* L. and its wild relatives: *Brassica rapa* L., *Raphanus raphanistrum* L., *Sinapis arvensis* L., and *Erucastrum gallicum* (Willd.) OE Schulz. *Theoretical and Applied Genetics* 107(3), 528–539. DOI: 10.1007/s00122-003-1278-0.

Weekes, R., Allnutt, T., Boffey, C., Morgan, S., Bilton, M. *et al*. (2007) A study of crop-to-crop gene flow using farm scale sites of fodder maize (*Zea mays* L.) in the UK. *Transgenic Research* 16(2), 203–211. DOI: 10.1007/s11248-006-9036-0.

Wei, W., Shen, B.C., Tang, Z.X. and Mi, X.C. (2007) Gene flow from transgenic oilseed rape to Chinese conspecific landrace under field conditions and its implication for China. *12th International Rapeseed Congress II Sustainable Development in Cruciferous Oilseed Crops Production*, 96–98.

Willenborg, C.J., Brule-Babel, A.L. and Van Acker, R.C. (2009) Low crop plant population densities promote pollen-mediated gene flow in spring wheat (*Triticum aestivum* L.). *Transgenic Research* 18(6), 841–854. DOI: 10.1007/s11248-009-9267-y.

Yan, S., Zhu, W., Zhang, B., Zhang, X., Zhu, J. *et al*. (2018) Pollen-mediated gene flow from transgenic cotton is constrained by physical isolation measures. *Scientific Reports* 8(1), 2862. DOI: 10.1038/s41598-018-21312-1.

Ye, X., Al-Babili, S., Klöti, A., Zhang, J., Lucca, P. *et al*. (2000) Engineering the provitamin A (β-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* 287(5451), 303–305. DOI: 10.1126/science.287.5451.303.

Yoshimura, Y., Matsuo, K. and Yasuda, K. (2006) Gene flow from GM glyphosate-tolerant to conventional soybeans under field conditions in Japan. *Environmental Biosafety Research* 5(3), 169–173. DOI: 10.1051/ebr:2007003.

Zhang, B.-H. and Guo, T.-L. (2000) Frequency and distance of pollen dispersal from transgenic cotton. [In Chinese with English abstract.]. *Chinese Journal of Applied and Enviromental Biology* 6, 39–42. DOI: 10.3321/j.issn:1006-687X.2000.01.008.

Zhang, C.-Q., Lv, Q.-Y., Wang, Z.-X. and Jia, S.-R. (1997) Frequency of 2, 4-D resistant gene flow of transgenic cotton. [In Chinese with English abstract.] *Scientia Agricultura Sinica* 30, 92–93.

Zhang, N.-Y., Linscombe, S. and Oard, J. (2003) Out-crossing frequency and genetic analysis of hybrids between transgenic glufosinate herbicide-resistant rice and the weed, red rice. *Euphytica* 130(1), 35–45. DOI: 10.1023/A:1022371104679.

10 Containment Strategies for Synthetic Gene Drive Organisms and Impacts on Gene Flow

Lei Pei* and Markus Schmidt
Biofaction KG, Vienna, Austria

Abstract

Gene drives, particularly synthetic gene drives, may help to address some important challenges, by efficiently altering specific sections of DNA in entire populations of wild organisms. Here we review the current development of the synthetic gene drives, especially those RNA-guided synthetic gene drives based on the CRISPR nuclease Cas. Particular focuses are on their possible applications in agriculture (e.g. disease resistance, weed control management), ecosystem conservation (e.g. evasion species control), health (e.g. to combat insect-borne and fungal infections), and for basic research in model organisms (e.g. *Saccharomyces*, fruit fly, and zebra fish). The physical, chemical, biological, and ecological containment strategies that might help to confine these gene drive-modified organisms are then explored. The gene flow issues, those from gene drive-derived organisms to the environment, are discussed, while possible mitigation strategies for gene drive research are explored. Last but not least, the regulatory context and opinions from key stakeholders (regulators, scientists, and concerned organizations) are reviewed, aiming to provide a more comprehensive overview of the field.

Keywords: biosafety; containment; CRISPR-Cas; gene drive; gene flow; regulation

10.1 Introduction

In high schools all over the world, pupils are taught how DNA is inherited from one generation to the next. Mendel's laws of inheritance describe the rules of sexual reproduction; however, there are exceptions to these rules. Natural gene drive has been found in plants, insects, and mammals, as 'a naturally occurring phenomenon in which selfish genetic elements manipulate gametogenesis and reproduction to increase their own transmission to the next generation'(Wedell *et al.*, 2019). A gene drive could refer to 'genetic systems that circumvent the traditional rules of sexual reproduction and increase the odds that a gene will be passed on to offspring, allowing them to spread to all members of a population' (UWE, 2016). In Europe, the European Food Safety Agency (EFSA) defines a gene drive as follows: 'Any genetic element that is inherited at a higher frequency than predicted by Mendelian laws of inheritance can be referred to as a gene drive' (EFSA, 2020).

*Corresponding author: lei.pei@idialog.eu

 Gene Flow: Monitoring, Modeling and Mitigation (eds Wei Wei and C. Neal Stewart, Jr)
DOI: 10.1079/9781789247480.0010

Other than natural drives, a synthetic gene drive was first suggested by Austin Burt, proposing that homing endonucleases could be used to construct gene drives to alter a population in the wild (Burt, 2003). The idea of harnessing gene drives to develop useful applications is not new: they have been studied for years, aiming to develop a system of biased inheritance that enables a dedicated genetic trait to pass to the offspring through sexual reproduction (NASEM, 2016). Usually, by introducing a gene drive into an organism, it would result in a specific genotype that might determine a specific phenotype from one generation to the next, and eventually (ideally) the whole population. Yet the technical constraints of the traditional gene drives, as well as environmental release concerns linked to genetically modified organisms (GMOs), limit their real-world application. The initial research on synthetic gene drives was thus done by harnessing the I-SceI endonuclease gene homing in transgenic mosquitoes with an I-SceI recognition site inserted into the genome (Windbichler *et al.*, 2011). The mitochondrial endonuclease I-SceI is able to recognize and cut an 18 bp restriction site to generate double-strand break (DSB) in fruit flies and was a new tool for studying DSB repair mechanisms in *Drosophila* (Bellaiche *et al.*, 1999). Yet gene drives based on homing endonucleases had difficulty in introducing useful genetic features into wild-type genomes.

There has been more interest in research on gene drives in recent years, since they may be able to address some hard-to-tackle ecological problems by converting entire populations of the targeted organisms, ranging from erasing invasive plant species to eliminating disease-carrying animals, such as mosquitoes. The research on gene drive applications focused mainly on those species that reproduce sexually. Such targets included disease vector organisms (e.g. mosquitoes and flies), agricultural pests (e.g. pigweed, screwworm, and desert locust), and invasive species (e.g. mice, rats, feral cats, cane toads, and some invasive plant species).

Recent advances in synthetic biology have allowed more 'powerful' gene drives to be developed, particularly those developed by new genome modification techniques, such as clustered regularly interspaced short palindromic repeats (CRISPR). CRISPR originates as a natural bacterial immune system against viruses or plasmids. When viruses infect bacteria, some bacteria are able to capture the short fragments of genetic materials from invaded viruses and incorporate these short sequences into their own genomes. This helps bacteria to recognize viruses later on and defend the next round of phage infections. Bacteria achieve defense against the phage infections by producing an RNA molecule as a guide and then with the action of the CRISPR-associated(Cas) dual RNA-guided DNA endonuclease enzyme to break up the viral genome (Bhaya *et al.*, 2011).

Other than a natural immune system against viral infections of bacteria, the unique feature of the CRISPR-Cas system is that it can recognize and cut a dedicated site within a genome, which turns it into a powerful tool for genome modifications. It has now been developed into a gene editing platform in which an endonuclease and a guide RNA are used to introduce double-strand breaks at a specified location within the genome, while further modifications can then be achieved, including site mutation, deletion, and insertion (Jinek *et al.*, 2012; Qi *et al.*, 2012).

Among all the strategies to engineer synthetic gene drive organisms, the CRISPR-Cas approach is the most efficient one so far. The RNA-guided gene drives based on the CRISPR nuclease Cas systems are constructed by introducing a DNA cassette encoding Cas into the genome together with a single-guide RNA (sgRNA) matching either side of the target gene of interest. The consistent expression of Cas and sgRNA would theoretically edit the genome one generation after another, thus spreading dedicated genotypes through wild populations over generations (Esvelt *et al.*, 2014; Akbari *et al.*, 2015). In this way, the change is passed on to up to 100% of offspring (Scudellari, 2019).

It is believed that organisms derived from this technique might pose new opportunities and challenges for research in gene flow. Since 2014, scientists have engineered CRISPR-based gene-drive systems in mice, mosquitoes, fruit flies, and fungi, and are currently developing it for plants (e.g. *Eucalyptus*) and in other mammals (e.g. feral cats). Gene drives have been proposed as a way to reduce or eliminate insect-borne diseases, control invasive species, reverse insecticide resistance in pests, and breed

disease-resistant plants. Yet the real-world applications of these gene-drive modified organisms have been supplanted by other unknowns: the efficiency in the receiving environment, how to test them, and how the releasing organisms should be regulated. To date, no engineered gene drive organism has yet been released into the wild: most of the testing is in laboratories, with some pilot field tests in isolated settings, such as releasing a test batch of mosquitoes that were genetically engineered (a strain of genetically modified sterile male mosquito) but not with gene drive in Burkina Faso – the first such event in the African Continent (Scudellari, 2019).

Unlike any GMOs tested before, gene drives will most likely spread by themselves once released, to fulfill their designed tasks of erasing or limiting the targeted species. Research on gene drive organisms showed that they could convert the target species in laboratory settings up to 100%. Thus, there are concerns on how to control these powerful genetic elements once they are eventually released intentionally for their real-world applications, in case something runs out of control. Theoretically, gene drives are believed to have the potential to alter entire populations and, further, entire ecosystems via gene flow. Therefore, containment strategies have been developed during the development of gene drive organisms, while stakeholders (such as those from the research community and regulatory authorities) have proposed recommendations on how to conduct research on gene drive organisms. We review here the current development of CRISPR-Cas gene drive organisms in agricultural applications, ecosystem conservation, human health, and basic research. The strategies that might help to confine these gene drive-modified organisms will then be explored. Last but not least, we will review the regulatory environment, focusing on the impacts on gene flow.

10.2 Gene Drive Applications

Generally speaking, there are two major types of gene drives in development: self-sustaining, and self-limiting (EFSA, 2020). The self-sustaining type is to introduce desirable genes to increase the frequency in a target population and eventually become dominant in the population. The self-limiting type is to increase the frequency of desirable genes in a target population for a limited number of generations, but subsequently to decrease and eventually become lost from the population. The self-limiting gene drive could either change unwanted characteristics of the target organisms or suppress population density. Examples are daisy-chain drives (Noble *et al.*, 2019) and split killer-rescue drives (Gould *et al.*, 2008). Below, the applications of these gene drives are reviewed based on the aims of these gene drive-modified organisms.

10.2.1 Agricultural applications

Gene drive approaches have progressed rapidly in plant science in recent years, not only for plant breeding but also for pest and weed management (Neve, 2018; Pixley *et al.*, 2019). In Europe, the European Commission's Committee on Gene Drive Research in Non-Human Organisms reviewed the potential agricultural applications (SAM, 2017), including applications that:

- Control or alter organisms that damage crops or carry crop diseases, e.g. fruit flies or aphids;
- Eliminate weedy plants that compete with cultivated crops;
- Eliminate herbicide or pesticide resistance;
- Control or alter organisms that carry infectious diseases that infect farm animals or organisms that directly cause diseases;
- Control or alter organisms that serve as reservoirs of farm animal diseases.

Gene drive approaches would provide novel resilience of endangered or threatened plant species by the spread of fitness-improving traits that would replace the wild counterparts; for example, to introduce disease-resistance genes via gene drives to rescue *Eucalyptus* species that are threatened by fungal infections caused by *Puccinia psidii* (Barrett *et al.*, 2019). Using gene drive modification to suppress specific insect-vectored plant pathogens is also under development. A similar type of gene drive is under investigation that would render amphibians immune to the pathogenic chytrid fungus, which

is responsible for the decline of some amphibian species (Champer *et al.*, 2016). As well as fungal pathogens, a gene drive has been investigated targeting Asian citrus psyllids to make them unable to transmit bacteria causing citrus greening (Champer *et al.*, 2016; Brown, 2017; Gutzmann *et al.*, 2017).

Novel control techniques are needed for agricultural weed management. A concept for CRISPR-Cas9-based gene drives has been developed to control weeds by introducing population-suppression drives to spread deleterious mutations that alter the fitness, or by introducing population-sensitizing drives to turn weeds more sensitive to subsequent management interventions (such as herbicides) (Neve, 2018). Some challenges remain, such as seed production capacity and the self-fertilizing nature of many weed species. The potential application of gene drives for agricultural weeds depends on the breeding systems, genetic parameters, growth conditions, availability of genomic information, and genomic tools. Therefore, to apply gene drives for management of herbicide-resistant weed species, one would have to take all these parameters into consideration. Table 10.1 gives more details on weed species that are subject to gene drive-based weed management (NASEM, 2016; Neve, 2018).

10.2.2 Ecosystem conservation

10.2.2.1 Control of non-indigenous mice on islands

Non-indigenous mice (*Mus musculus*) are one of the leading causes of extinction of native wildlife and plants on islands. To counter the negative effects of these rodents, one option is to reduce or altogether eliminate populations of non-indigenous mice in order to mitigate native biodiversity on islands around the world. The Genetic Biocontrol of Invasive Rodents Programme (GBIRd) (available at: https://www.geneticbiocontrol.org/, accessed 7 April 2021) has worked on gene drive-modified mice to use the technology to eliminate invasive rodents from islands. Releasing the gene drive mice to control the wild rodent population is thought to be more efficient than the rodent-erasing approach by pesticides that is currently used for this purpose. The pesticides are expensive and difficult to use in populated larger islands. In order to control the population, a sex-determining gene drive to produce more male offspring than females has been studied in mice (Cocquet *et al.*, 2012). The molecular mechanism of this sex drive was based on an endogenous region of high meiotic drive in

Table 10.1. Gene drive for agricultural weed control, using *Amaranthus palmeri* as an example. (Adapted from NASEM, 2016, and Neve, 2018)

	Feature	Gene drive as weed control
Mating system	Outcrossing	Gene drive will be able to spread through outcrossing mating system
Fecundity	High	Will take longer for gene drive to suppress population of high fecundity
Resistance	High	Resistant to eight herbicide modes of action (Gaines *et al.*, 2020) and as many as five different modes of action found within a single population (Kumar *et al.*, 2019)[a]
Seed bank	Low	Not persistent in soil, thus a seed repository resistant to gene drive is less likely
Other	Annual, wind pollinated	Relatively fast generation time, less impact on insects even if weeds are eradicated
Genome	421.8 Mb	Under development (Montgomery *et al.*, 2020)[b]

[a]Due to the high resistance to commonly used herbicides, alternative approaches are now under investigations to combat these weeds. One of these approaches is to deploy gene drive modified weed *Amaranthus*, converting the whole population into 'male dominant', thus to cease the reproduction and eventually lead to the population crash.
[b]Full sequence not yet available, but several research groups working on it.

the mouse genome on chromosome 17 called the *t*-complex. The *Sry* gene was introduced into the genome of male mice on chromosome 17 instead of its usual location on the Y chromosome. Upon mating to a wild-type XX female, offspring mice (both XX and XY) would have carried *Sry* on chromosome 17, unlike wild-type on Y chromosome only. The expression of the *Sry* gene would promote male characteristics, thus turning both types of offspring into males. Over a few generations, the population of mice would become male-dominant, theoretically, thus suppressing the mouse population by reducing female mouse numbers (Campbell *et al.*, 2015).

Other than gene drives targeting male mice, super-Mendelian inheritance mediated by CRISPR-Cas9 in the female mouse germ line has also been developed. A genetic element is composed of a guide RNA targeting the mouse tyrosinase gene to be edited by CRISPR-Cas9 when the enzyme is activated in the early stage of mouse embryo or in the developing germ line. It showed that the Cas9 expressed in the female germ line could induce double-stranded breaks to be corrected by homology-directed repair, and be passed to the next generation, suggesting further optimizations would be needed to increase the frequency of gene conversion in both males and females as well as to address issues related to prevalence of drive-resistant alleles to achieve efficient suppression in a wild population (Grunwald *et al.*, 2019).

10.2.2.2 Control of feral cats in Australia

In Australia, feral cats (*Felis catus*) introduced by European settlers have become pests to the ecosystem because of their abilities to adapt and live in the new continent. Many native species have been hunted by these feral cats to extinction, or near-extinction. Traditional animal controls (e.g. poisonous baits) do not work well on the cats. With the development of CRISPR gene drive technology, it is believed that gene drive-modified cats could be introduced into the feral populations to produce male-only offspring; thus the feral cat population would die out (CSIROscope, 2018). Yet there is a knowledge gap to be overcome to fulfill these aims (Moro *et al.*, 2018).

10.2.2.3 Control of invasive knapweed in the USA

Knapweed (*Centaureastoebe* (syn. *Centaurea maculosa*)) is native to Europe but has been considered as an invasive weed in the USA since its introduction in the late 19th century. It will cause erosion on soil and compete against the native plant species once it spreads to new ecosystems. Although different control management strategies have been developed to control this weed (biological control on seed reduction, physical removal, fire and chemical treatments), new approaches are needed, such as deploying gene drives to control the non-indigenous knapweed species. Two gene drive approaches have been developed to limit the spread of knapweed. One is to engineer a suppression gene drive by targeting sex-specific genes to alter the sex ratios and eventually decrease a population. The other is to inactivate the catechin biosynthetic pathway to affect the knapweed's ability to compete against endemic plants. Studies have shown that the ability of knapweed to outgrow the native plants was in part due to allelopathy, the production of a compound (a catechin) that inhibits the germination and growth of native plant species (Pollock *et al.*, 2009).

10.2.3 Human health-related

10.2.3.1 Gene drive-modified mosquito to combat insect-borne infections

Only female mosquitoes can transmit diseases such as malaria, dengue and Zika; therefore, to improve human health, it is important to use gene drives to introduce biased sex ratio in insect offspring to control vector-borne infectious disease. An example of a population reduction strategy by a synthetic gene drive is to reduce mosquito populations by reducing the number of female mosquitoes. This approach has been developed by the research initiative 'Target Malaria', testing on *Anopheles gambiae* mosquitoes, which are an important vector of malaria. The laboratory-grown male *A. gambiae* mosquitoes engineered with a gene drive would be released, and expected to increase the male share in populations over generations; thus these male gene

drive mosquitoes would disrupt the fertility of female mosquitoes or convert the population into a male-dominant one. Therefore, the goal to reduce malaria transmission could be achieved by reducing the numbers of female mosquitoes. In *A. gambiae*, the gene *doublesex* (Agdsx) encoding two alternatively spliced transcripts, dsx-female (AgdsxF) and dsx-male (AgdsxM), can control differentiation of the two sexes. An exon of the female transcript is highly conserved and is targeted for gene editing. CRISPR-Cas9 constructs were developed to disrupt the intron–exon boundary to interrupt the formation of functional AgdsxF. Such a CRISPR-Cas9 gene drive construct was shown to spread rapidly in caged mosquitoes, reaching 100% prevalence within 7–11 generations, resulting in progressively reduced egg production and eventually total population collapse. No selection of alleles resistant to this gene drive was detected in these caged mosquitoes, although Cas9-resistant variants were found in each generation at the target site, but they did not block the spread of the drive (Kyrou *et al.*, 2018). A male-biased sex-distorter gene drive (SDGD) in the human malaria vector *A. gambiae* has been developed, also targeting the *doublesex* gene. A gene drive was composed coupling an X-chromosome-shredding I-PpoI nuclease to a CRISPR construct, and inserted the drive into a conserved sequence of the *doublesex* gene. In modeling of invasion dynamics, SDGD was predicted to have a quicker impact on female mosquito populations than previously developed gene drives targeting female fertility. This SDGD achieved a male-only population after 10–14 generations, starting from a 2.5% allelic frequency, with no known selection resistance (Simoni *et al.*, 2020).

Another approach to population alternation strategy by synthetic gene drive is to create mosquitoes no longer transmitting pathogens such as malaria parasites. Mosquitoes would be genetically engineered to make their offspring incapable of transmitting the malaria parasites to humans (Vernick *et al.*, 2018). In this approach, male mosquitoes with gene drive, grown in the laboratory, would be introduced into wild populations to mate with wild-type females, and the anti-parasite effector genes in female mosquitoes would be replicated by the gene drive. Theoretically, the affected females would survive without malaria parasites and thus be unable to transmit malaria to humans.

Apart from completely reducing or altering mosquito populations, some other strategies of gene drive mosquitoes are under development: for example, to engineer *Aedes aegypti* mosquitoes to express an antibody that protects the insects against all four major strains of dengue. An all-purpose gene drive is also currently under investigation and in this case a toxin will be activated when a virus infects *A. aegypti*, making the insects unable to spread dengue, Zika, chikungunya, and yellow fever, due to the built-in feature of the all-purpose gene drive that will trigger toxin production, which will kill the mosquito in case of any viral infection.

10.2.3.2 Gene drive-modified yeast to combat fungal infections

Some organisms are known to be challenging to manipulate genetically, yet genome modifications are needed for research purposes. For example, a gene drive-based approach has been developed in *Candida albicans*, a common drug-resistant fungal pathogen. The CRISPR-Cas9-based gene drive platform generated mutations in the fungus with close to 100% efficiency (Morio *et al.*, 2020). To develop gene drive for medical mycology study, a construct expressing *CAS9* and two sgRNAs is integrated at the *NEUT5L* locus in a haploid cell. When haploid cells with the gene drive were mated to wild-type cells, the incoming wild-type allele would be modified by the sgRNA module, leading to homozygous diploid double-deletion mutants in *C. albicans* (Shapiro *et al.*, 2018). Thus, this gene drive-based genome modification platform would be further applied for research on other fungal pathogens that are difficult to modify genetically.

10.2.4 Gene drive for basic research

10.2.4.1 Saccharomyces cerevisiae

Inheritance-biasing gene drives based on RNA-guided CRISPR-Cas9 genome editing systems

have been applied in *S. cerevisiae*. As a novel method to alter population, the gene drive can reach a conversion rate above 99% when mated to wild yeast over successive generations. When testing the gene drives in *S. cerevisiae*, both non-essential (ADE2) and essential genes (ABD1) were targeted. The results showed that the efficiency of the gene drives was not related to which gene the drive targeted. In addition, a strategy for a new gene drive to overwrite and reverse the changes made by a previous drive was tested in these engineered yeasts and showed the reversibility of CRISPR-Cas9-based gene drive. These synthetic gene drive yeasts showed that yeast might be use as a platform to test CRISPR gene drive constructs for other organisms, and to facilitate gene drive optimizations (DiCarlo *et al.*, 2015).

10.2.4.2 Fruit fly

The fruit fly *Drosophila* is a model insect that has been used for genetic studies for many years. The first CRISPR-Cas9-derived gene drive was reported in the fruit fly in 2015 (Gantz and Bier, 2015). The method to build this gene drive was termed the 'mutagenic chain reaction' (MCR), which is based on the CRISPR-Cas9 genome editing system to introduce autocatalytic mutations, leading to homozygous loss-of-function mutations. It showed that MCR could convert heterozygous mutations to homozygosity within the fly population. By using the CRISPR-Cas9 construct to introduce mutations that disabled both normal copies of a pigmentation gene on the fruit fly chromosomes, the gene drive could transmit itself to the next generation with 97% efficiency.

10.2.4.3 Zebrafish

As a model organism, the zebrafish (*Danio rerio*) has been used to study gene drive mechanisms in a vertebrate animal. A self-limiting type of gene drive could be developed in zebrafish and, so far, a tissue-specific spatial and temporal system to control gene expression by the presence of tetracycline, known as TetON, has been developed for this model organism (Knopf *et al.*, 2010). These dually inducible TetON systems were based on a tetracycline-inducible transcriptional activator, which could be regulated by tetracycline-responsive promoters upon administration of the appropriate ligands, leading to conditional and tissue-specific control of gene expression. Meanwhile, gene expression induction could be controlled, either reversibly or fine-tuned by the doses of inducers. The fish genome has been fully sequenced (Howe *et al.*, 2013) and gene editing techniques have been applied (Prykhozhij *et al.*, 2016). It can be used to build a model to study basic research questions about gene drives in a vertebrate species (Shah and Moens, 2016). It has been suggested that CRISPR-Cas-based gene drive could be constructed in non-essential gene loci, while using zebrafish as a model organism to test gene drives for other vertebrate animals (NASEM, 2016).

10.3 Containment Strategies for Organisms Modified with Synthetic Gene Drives

Since gene drive-modified organisms are part of GMOs, all the containment strategies foreseen for GMOs would also be applicable to gene drive organisms. We review those strategies here with respect to their applicability for gene drives.

10.3.1 Physical containment

A physical containment can be the confinement within a defined space (e.g. labs, bioreactors, greenhouses, fences, cages, nets, blocks), defined conditions (e.g. temperature, light, fire), or removal by physical means (e.g. catching and hunting for animals and removal of plants). Because gene drive organisms can be microbes, plants, or animals, the physical containment systems applied should be decided by the nature of the organisms and the intended uses. For example, when testing gene drives in insects, physical confinement would be applied at the first step in performing trials in large outdoor cages to prepare for subsequent open-field releasing, where escape is less likely (WHO,

2014). In general, physical containment is the preferred choice in the testing stage.

10.3.2 Chemical containment

Chemical containment refers to the application of chemical agents that can limit growth or kill the gene drive organisms. It consists mainly of chemical reagents that are toxic to organisms. These chemical reagents include herbicides for plants, poisons for animals, insecticides for insects, and disinfectants and anti-microbials for bacteria and yeasts.

10.3.3 Biological containment

Most physical and chemical containment strategies are not suitable for application when a gene drive-modified organism is released into the environment. Therefore, non-physical/chemical containment methods should be employed to prevent the creation of an RNA-guided gene drive capable of spreading beyond its payload in the ecosystem. To safeguard gene drive organisms, different biological containments have been developed, aiming to control, counter, and reverse gene drives to mitigate the potential for harm, if any. In addition, due to the nature that gene drives are promoted sexually, this feature would pose greater risks if it escapes. Therefore, when a gene drive experiment is done on model organism such as fruit fly, working on a species with a genetic background that is less likely to escape the laboratory and mate, such as wingless flies, may provide an additional layer of biological containment.

10.3.3.1 Safeguarding gene drives in animals

A self-limiting drive known as a daisy drive has been developed. The drive is engineered to lose its genetic payloads one at a time, like plucking a daisy flower, until all payloads run out over several generations. The so-called daisy-chain gene drive has been constructed to achieve local restriction of a drive within a mouse population (Noble *et al.*, 2019). The CRISPR-Cas9 components were split up and distributed throughout the genome. The components for the drive were arranged in a linear daisy-chain: one component at the base promoted the next component of the drive and so on to promote the next higher component. Since the gene drive was activated via chain reactions, it would most likely be lost successively: after a certain number of generations, the gene drive would be lost from the population. The release ratio and the number of links to the daisy chain at the beginning of the release would then decide how long the gene drive would function.

A so-called killer-rescue (toxin-antidote based) system has also been developed, although not specifically for CRISPR gene drives. For example, an inverse maternal effect dominant embryonic arrest (Medea) system has been based on rescuing a toxin in the zygote unless it receives a maternally delivered antidote (Marshall and Hay, 2011). A toxin-antidote type of CRISPR-Cas9-based gene drive has also been developed by targeting a recessive lethal gene with a recoded sequence to rescue only drive-carrying individuals, with computational modeling predicting that such a drive would be threshold-dependent, spreading conditionally when introduced above a frequency threshold (Champer *et al.*, 2019).

10.3.3.2 Safeguarding gene drives in yeast

To prevent the accidental escape of gene drives into wild yeast, strategies have been developed to contain gene drive yeast *Saccharomyces cerevisiae*. One approach was to split the CRISPR-Cas9-based gene drive system into two separate genetic locations: an episomal plasmid carrying the Cas9 gene and a drive element encoding the guide RNA located in selected genomic loci, thus preventing the spread of a self-sufficient inheritance-biasing drive into the unintended wild-type yeast. It was suggested that even if gene drive yeast were to escape into the wild, the required Cas9 episomal plasmid would be segregated away from the drive element (in the genome), rendering the drive inoperative (DiCarlo *et al.*, 2015). In addition to molecular confinement for the gene drive in yeast, the same research group developed an approach to overwrite

the changes made by an earlier gene drive, as an additional safeguard strategy for gene drive-modified yeast (DiCarlo *et al.*, 2015). In those types of applications where a population is targeted for extinction, obviously such second-order gene drives would come too late.

10.3.3.3 Inhibition by anti-CRISPR-Cas proteins

Given the widespread use and application of the CRISPR-Cas gene editing system, countermeasures to control its gene editing activities have also been developed. For example, the discovery of a new class of 'anti-CRISPR' proteins, evolved from bacteriophages in response to the prokaryotic nuclease-based immune system, provides a new platform for control over genome editing. This type of control for the CRISPR-Cas system would also be useful to control gene drives derived from this system. It has been shown that two of the anti-CRISPR molecules, the AcrIIA2 and AcrIIA4 proteins, could inhibit active gene drive systems in budding yeast (Basgall *et al.*, 2018).

10.3.3.4 Control by small molecules

A modified Cas9-based gene drive that could be controlled by a small molecule was developed in the fruit fly, as an additional layer of safety for laboratory studies. A small molecule-controlled gene drive system was engineered to control the propagation capacity of gene drive-modified *Drosophila*, with built-in genetic control elements that could be tightly and conditionally controlled to achieve an inheritance probability between 50% and 100%. This control of efficacy of the gene drive was based on a modified SpCas9 (a Cas9 enzyme isolated from *Staphylococcus aureus*), the expression of which was fine-tuned by a synthetic small molecule trimethoprim. This trimethoprim-induced gene drive activation system might increase the safety for future field trials by enabling spatial and temporal control of gene drive organisms using small molecules (López Del Amo *et al.*, 2020). A requirement for this approach to work is of course the ability to apply the chemicals to the target organisms, something that becomes difficult (if not impossible) once an environmental release takes place.

10.3.3.5 Other CRISPR-Cas inhibitors

Multiple families of anti-CRISPR compounds can be found widely in mobile genetic elements (e. g. phages and conjugative elements) of diverse bacterial species. To improve applications of CRISPR-Cas gene drive approaches, it would be safer to enable the CRISPR-Cas-based gene drive systems with the ability to inhibit Cas function spatially, temporally, or conditionally. Three distinct families of anti-CRISPRs were discovered specifically inhibiting the CRISPR-Cas9 system by binding directly to Cas9, and could be used as potent inhibitors of genome editing by Cas9 in human cells. These CRISPR inhibitors might also enable 'off-switches' for CRISPR-Cas9 activity in gene drive systems and provide a possible countermeasure to inhibit CRISPR-Cas9 gene drive in eukaryotes (Pawluk *et al.*, 2016).

10.3.4 Ecological containment

Built on top of the containment strategies mentioned above, there are additional containment strategies that are useful in order to contain gene drive organisms. One of the additional layers of containment for gene drive research is ecological containment. An ecological containment could be used to conduct research of gene drive organisms in ecological settings that do not harbor or sustain native populations of the target organisms, for example to conduct gene drive studies of mosquito-borne tropical diseases (such as malaria and dengue fever) in temperate regions, where the engineered mosquitoes most likely cannot survive or find mates if they escape. Case-by-case evaluation is necessary here, especially concerning the question of whether the organism could survive a transportation process over large distances by air and water, or piggybacked onto other animals (e.g. migratory birds), or as a blind passenger in human transport.

A special form of ecological containment is to conduct deployment in islands, such as releasing engineered mosquitoes in an island when conducting a field trial. Gene drives can also be designed to spread only within a certain geographic region, under certain conditions, to alter a target population locally. A tethered

homing gene drive system was proposed by Dhole *et al.* (2019). This type of drive was based on driving a genetic load using a homing construct that was anchored to a spatially restricted gene drive. These so-called spatially restricted gene drives were composing of a split homing drive and an engineered under-dominance drive linked to one part of the split drive (e.g. Cas) (Dhole *et al.*, 2019). It has been suggested that the spatially restricted gene drives would not be able to establish themselves at high frequency in neighboring populations when migration rates were low (EFSA, 2013, 2020).

10.4 Current Regulatory Situation and Opinions from Key Stakeholders

It is clear that the enhanced ability to control populations by gene drive technology would offer enhanced agricultural practices, counteract invasive species, and improve human health by reducing vector-borne diseases. It has been shown that gene drive-modified organisms could potentially prevent the spread of disease (e.g. by rodents and insects), reverse pesticide and herbicide resistance in insects and weeds, and control invasive plant species. However, the potential environmental risks for ecosystems need to be taken into consideration. Other than novel containment strategies, the question is whether novel regulation is also needed to control the spread of gene drives and reverse genomic changes.

10.4.1 Opinions from World Health Organization (WHO)

In 2014, the WHO published its guidance on how genetic modified mosquitoes (GMms) should be tested, including the testing of gene drive mosquitoes. Four-phase trials have been suggested to conduct research on GMms:

- Phase 1 for laboratory population study;
- Phase 2 for physically and ecologically confined field trial;
- Phase 3 for staged open field release; and
- Phase 4 for post-implementation surveillance.

Key messages from this report have been to provide: (i) an overview on state-of-the-art research on GMms; (ii) suggestions on efficacy evaluation; (iii) considerations on biosafety for Phase 1, 2, 3 and 4 studies; (iv) outlines on ethics and public engagement; and (v) an overview on regulatory frameworks. This WHO guidance was intended 'to foster quality and consistency in the processes for testing and regulating new genetic technologies', and hoped to 'contribute to comparability of results and credibility of conclusions in addressing the requirements for decision-making by countries interested in the potential use of these technologies' (WHO, 2014).

Gene flow issues were tackled in the WHO guidance. Gene flow was defined as 'the movement (expressed as increase in frequency) of genes or alleles into a population from one or more other populations'. It has been suggested that the likely risk of transgenic gene flow should be studied in Phase 2 to understand the local pattern of gene flow in mosquito population, in Phase 3 to be assessed as one of the additional biosafety considerations, and in Phase 4 to monitor the rate of spread of the gene drive in wild populations and comparison with model and Phase 3 predictions, as well as wide-scale intermittent measurement of GMms dispersal and gene flow (WHO, 2014).

10.4.2 Opinions from European Food Safety Agency (EFSA)

The EFSA has recently published a draft opinion on how to evaluate existing EFSA guidelines for their adequacy for the molecular characterization and environmental risk assessment of genetically modified insects with synthetically engineered gene drives (EFSA, 2020). The background for this scientific opinion is the recent advances in engineering of gene drives that spread genes of interest through interbreeding populations at a high rate. Gene drive insects are the most likely to be deliberately released into the environment. To take a proactive measure and to take into consideration the potential of gene drives to spread within populations, persist in the environment, and the possible irreversible effects on

ecosystems, the EFSA was required to review whether the existing guidelines for the risk assessment of genetically modified animals were adequate to regulate gene drive-modified mosquitoes and agricultural pests for deliberate release into the environment. The two existing EFSA guidelines are:

1. Scientific opinion on guidance on the risk assessment of food and feed from genetically modified animals and animal health and welfare aspects (EFSA, 2012).
2. Guidance on the environmental risk assessment of genetically modified animals (EFSA, 2013).

The draft opinion concluded that it was necessary to revise the existing guidance on molecular characterization, persistence and invasiveness, modeling, and post-market environmental monitoring, taking into account the non-food/feed uses of gene drive insects and the self-replicating nature of these organisms. As for the environmental risk assessment of gene drive insects, an explicit problem formulation should be started so that 'the protection goals, plausible and relevant exposure scenarios and the potential adverse effects from those exposures are identified on a case-by-case basis'. In addition, it was important to enhance dialogue among risk assessors, risk managers, and stakeholders to 'define clear protection goals and decision-making criteria' for the environmental risk assessment (EFSA, 2020).

When conducting a staged approach to the problem of formulation for the identification of hazards associated with the dispersal of genetically modified animals, EFSA provided an example of how probability of gene flow should be formulated for hazard characterization for risk assessment: 'the probability of gene flow = probability of entry into receiving environment x probability of introgression'. Taking the example of gene flow risk assessment on genetically modified fish, two major endpoints should be addressed: '*entry* of sexually mature, fertile GM individuals into a receiving environment; and *introgression* of recombinant DNA genotypes into the gene pool of wild relatives' (EFSA, 2013). This gene flow risk assessment on transgene animals could also be applied to gene flow risk assessment on gene drive organisms.

Horizontal gene transfer (HGT) between insects is only infrequently inferred (Silva *et al.*, 2004). The gene drive systems aim for vertical transfer of genes of interest above the expected Mendelian segregation rate, and are not expected to increase HGT rates. Unlike the case with insects, HGT events could happen in single-cell organisms (such as bacteria) (Keeling, 2009). It is known that natural transformation would facilitate uptake and genomic integration of foreign DNA fragments from higher organisms. Microorganisms, especially bacteria, can act as recipients of genes transferred from eukaryotes (Anderson and Seifert, 2011). Therefore, HGT events from GM insects to bacteria might take place if there were sequences with high similarity to bacterial DNA in the insect genome (Bensasson *et al.*, 2004). Thus, the likelihood of HGT from genome-modified insects into microorganisms should also be considered. Furthermore, it has been pointed out that 'the introduction of intron-free coding sequences in the GM insect genome with high similarity to microbial DNA would increase the probability of recombination and expression, if transferred' (EFSA, 2013).

10.4.3 Opinions from the US National Academies of Sciences, Engineering and Medicine

In their review of gene drives, the US National Academies of Sciences, Engineering and Medicine (NASEM) stated that it was important to continue research to establish the efficacy and safety of gene drive organisms prior to the decision about whether they are suitable for use. NASEM provided several recommendations for the research and development of gene drive (see Box 10.1), suggesting that 'it is crucial to establish a rich understanding of the target organism, its relationship with its environment, and potential unintended consequences' and that 'a phased testing pathway, such as the one outlined by the World Health Organization (WHO) for testing genetically modified mosquitoes, can facilitate a precautionary, step-by-step approach to research on gene drives' (NASEM, 2016).

Box 10.1. Some recommendations for gene drive research from National Academies of Sciences, Engineering, and Medicine study on gene drives (Source: NASEM, 2016).

To reduce potential harms

1. Research should be conducted in a phased testing pathway, a step-by-step framework and continued through monitoring gene-drive modified organisms in the environment, while pre-defined 'go/no-go' decisions should be defined for each phase to determine whether to transition to the next phase, based on experimental evidence
2. Available datasets and models should be applied to evaluate strategies that will help to minimize off-target and non-target effects
3. A split gene drive should be considered in the early phase of study to avoid an unintended consequence from a failure of containment
4. A gene drive should be tagged with a recognizable marker to distinguish modified organisms from wild types and facilitate monitoring
5. A 'reversal' gene drive should not be considered as the only strategy to mitigate the unintended effects of another gene drive

To assess ecological risk

1. Ecological risk assessment should be applied to estimate the probability of immediate and long-term environmental and public health effects
2. Experimental field trials should be designed to validate or improve cause–effect pathways
3. Researchers and risk assessors should collaborate early and often to design studies to provide information needed for risk assessment

10.4.4 Opinions from the gene drive research community

The research community has published several articles on its opinions on gene drive research. One was by Esvelt *et al.* (2014) who proposed possible safeguards and control strategies for gene drive organisms: reversibility, immunization, precisely targeting subpopulations, and limiting population suppression. It was suggested that keeping a certain population immunized by the gene drive would help to limit the gene flow of the drive, protecting some subpopulations from the effects of drives. Thus, even though gene flow might happen, the immunized population would restore the gene drive suppressed population, gradually. In addition, due to the broader consequences of releasing synthetic gene drive organisms into the environment, it has been suggested that research involving gene drives intended for environmental release 'should occur only after a careful and fully transparent review process' (Esvelt *et al.*, 2014). In another opinion article, a larger group of scientists proposed three recommendations on how to safeguard gene drive experiments in the labs (Akbari *et al.*, 2015), as follows.

1. All work involving potential gene drive systems should be preceded by a thorough assessment by the relevant biosafety authorities of the risk of unwanted release from the laboratory. We encourage these authorities to seek guidance from external experts and make their evaluation available to others.
2. All laboratory gene drive experiments should employ at least two stringent confinement strategies (…) whenever possible to minimize the risk of altering wild populations. Using one form of confinement may be justified only if relevant biosafety authorities determine that it will reduce the probability of release to a level that is acceptably low. This probability must be defined on a case-by-case basis. The analyses necessary to confidently predict the efficacy of confinement strategies for gene drive systems are in a nascent form. Therefore, any proposal to use one rather than multiple forms of confinement requires even greater scrutiny and extensive deliberation between regulatory authorities and scientists.
3. Organisms carrying gene drive constructs that could spread if the reproductively

capable life stages were to escape in transit should not be distributed to other institutions until formal biosafety guidelines are established. Whenever possible, laboratories should instead send DNA constructs or information sufficient to reconstruct the gene drive. Protocols for distributing materials should be established in discussion with the wider research community and other relevant stakeholders.

10.4.5 Opinions from biodiversity research community

The biodiversity research community has called for a global moratorium on releasing gene drive organisms, e.g. those from the organization Save Our Seeds (S.O.S., 2020). The call for a moratorium was based on the risks and open questions: risky research; concerns on non-recoverable, uncontrollable cross-border spread; efficiency and accuracy of CRISPR-Cas gene editing; lack of model to predict ecological effects; lack of risk assessment methods and guidelines; possible damage to biodiversity; dual use potential; lack of technology assessment; and lack of global decision-making body. Taking into consideration both the risks and the open questions, a moratorium would be placed until globally binding rules are set and solutions are found.

10.5 Concluding Remarks

Gene drives have been studied for years in laboratory settings, and researchers are now working on them to develop promising applications to tackle some pressing issues in the real world by deploying gene drive-modified organisms, with the new genome editing CRISPR-Cas systems. The potential applications for gene drives range from basic research (in the model organisms), agriculture applications (to control diseases and weeds), and ecosystem conservation (to eliminate invasive plants and animals), to human health (to combat insect-borne diseases).

Along with the advancing CRISPR-Cas genome editing technologies, a synthetic gene drive can now spread a targeted gene through a population nearly 100% within just a dozen generations, as shown in yeast, fruit flies, and mosquitoes. Such a powerful conversion efficiency calls for proper containment strategies and sufficient regulation. Physical, chemical, and ecological containment approaches have played important roles in the early stages of gene drive research. Yet these strategies might be of limited function for those organisms aimed for environmental release. Biological containment strategies are also being developed, which might be of importance for real-world applications of the gene drive-modified organisms.

Regulatory opinions from different stakeholders (regulators, scientists, and concerned organizations) have shown that a comprehensive approach to regulate gene drive-modified organisms would include not only proof of efficacy, but also proof of acceptability and deliverability. Due to the differences between gene drive-modified organisms and those of GMOs, four-phase trials have been proposed by WHO, which would be useful to guide the gene drive research from the laboratory bench to real-world applications. The EFSA has reviewed its existing guidance and laid down its risk assessment opinions. The NASEM in the USA has provided some case studies on gene drive research as well as drafted several recommendations. The scientific research community has also provided its recommendations on how to safeguard gene drive research, while concerned organizations have proposed a moratorium on environmental release of these modified organisms, citing lack of sufficient evidence to support their release.

Gene drives are transmitted by sexual reproduction. Therefore, vertical gene flow is expected. The effectiveness of this transmission will depend on the reproductive characteristics of a species, such as fecundity, mate choice that could bias inherited genes, number of progeny contributed to the next generation, mating system, and generation time, as for gene drive animals. Vertical gene flow to unintended species needs further study. Horizontal gene transfer events of gene drives would likely happen given enough time and scope, most likely from eukaryotic to prokaryotic (such as flows from insects to bacteria). More research would also be needed to gain more knowledge on this biosafety aspect. It has been suggested that,

taking gene flow into consideration, at least some mitigating approaches can be applied; for example, when designing the gene drive constructs, one should minimize homologous sequences of bacteria introduced into gene drives for target hosts (such as in insects). By implementing such principles, the probability of gene flow via HGT mechanism from insects (gene drive-modified organisms) to bacteria (in ecosystems) would be reduced.

Acknowledgement

This work has been supported by European Commission Horizon 2020 Project NEWCOTIANA (EC 760331), MADONNA (H2020-FETOPEN-1-2016-2017-76697), and BIOROBOOST (H2020-NMBP-BIO-CSA-2018-820699).

References

Akbari, O.S., Bellen, H.J., Bier, E., Bullock, S.L., Burt, A. *et al*. (2015) BIOSAFETY. Safeguarding gene drive experiments in the laboratory. *Science* 349(6251), 927–929. DOI: 10.1126/science.aac7932.

Anderson, M.T. and Seifert, H.S. (2011) Opportunity and means: horizontal gene transfer from the human host to a bacterial pathogen. *mBio* 2(1), e00005–00011. DOI: 10.1128/mBio.00005-11.

Barrett, L.G., Legros, M., Kumaran, N., Glassop, D., Raghu, S. *et al*. (2019) Gene drives in plants: opportunities and challenges for weed control and engineered resilience. *Proceedings of the Royal Society B: Biological Sciences* 286(1911), 20191515. DOI: 10.1098/rspb.2019.1515.

Basgall, E.M., Goetting, S.C., Goeckel, M.E., Giersch, R.M., Roggenkamp, E. *et al*. (2018) Gene drive inhibition by the anti-CRISPR proteins AcrIIA2 and AcrIIA4 in *Saccharomyces cerevisiae*. *Microbiology* 164(4), 464–474. DOI: 10.1099/mic.0.000635.

Bellaiche, Y., Mogila, V. and Perrimon, N. (1999) I-SceI endonuclease, a new tool for studying DNA double-strand break repair mechanisms in Drosophila. *Genetics* 152(3), 1037–1044. DOI: 10.1093/genetics/152.3.1037.

Bensasson, D., Boore, J.L. and Nielsen, K.M. (2004) Genes without frontiers? *Heredity* 92(6), 483–489. DOI: 10.1038/sj.hdy.6800451.

Bhaya, D., Davison, M. and Barrangou, R. (2011) CRISPR-Cas Systems in bacteria and Archaea: versatile small RNAs for adaptive defense and regulation. *Annual Review of Genetics* 45(1), 273–297. DOI: 10.1146/annurev-genet-110410-132430.

Brown, Z. (2017) Economic, Regulatory and International Implications of Gene Drives in Agriculture. *Choices. Quarter* 2. Available at: http://www.choicesmagazine.org/choices-magazine/submitted-articles/economic-regulatory-and-international-implications-of-gene-drives-in-agriculture (accessed 5 May 2021).

Burt, A. (2003) Site-specific selfish genes as tools for the control and genetic engineering of natural populations. *Proceedings of the Royal Society B: Biological Sciences* 270(1518), 921–928. DOI: 10.1098/rspb.2002.2319.

Campbell, K.J., Beek, J., Eason, C.T., Glen, A.S., Godwin, J. *et al*. (2015) The next generation of rodent eradications: innovative technologies and tools to improve species specificity and increase their feasibility on islands. *Biological Conservation* 185, 47–58. DOI: 10.1016/j.biocon.2014.10.016.

Champer, J., Buchman, A. and Akbari, O.S. (2016) Cheating evolution: engineering gene drives to manipulate the fate of wild populations. *Nature Review Genetics* 17(3), 146–159. DOI: 10.1038/nrg.2015.34.

Champer, J., Lee, Y.L., Yang, E., Liu, C. and Clark, A.G. (2019) A toxin-antidote CRISPR gene drive system for regional population modification. *bioRxiv*.

Cocquet, J., Ellis, P.J.I., Mahadevaiah, S.K., Affara, N.A., Vaiman, D. *et al*. (2012) A genetic basis for a postmeiotic X versus Y chromosome intragenomic conflict in the mouse. *PLoS Genetics* 8(9), e1002900. DOI: 10.1371/journal.pgen.1002900.

CSIROscope (2018) Gene drive technology: a new hope in the fight against feral cats. Available at: https://blog.csiro.au/gene-drive-technology-a-new-hope-in-the-fight-against-feral-cats/ (accessed 7 April 2021).

Dhole, S., Lloyd, A.L. and Gould, F. (2019) Tethered homing gene drives: a new design for spatially restricted population replacement and suppression. *Evolutionary Applications* 12(8), 1688–1702. DOI: 10.1111/eva.12827.

DiCarlo, J.E., Chavez, A., Dietz, S.L., Esvelt, K.M. and Church, G.M. (2015) Safeguarding CRISPR-Cas9 gene drives in yeast. *Nature Biotechnology* 33(12), 1250–1255. DOI: 10.1038/nbt.3412.

EFSA (2012) Scientific Opinion on the Guidance on the risk assessment of food and feed from genetically modified animals and animal health and welfare aspects. *EFSA Journal* 10(1), 2501. DOI: 10.2903/j.efsa.2012.2501.

EFSA (2013) Guidance on the environmental risk assessment of genetically modified animals. *EFSA Journal* 11, 3200.

EFSA (2020) Evaluation of existing EFSA guidelines for their adequacy for the molecular characterisation and environmental risk assessment of genetically modified insects with synthetically engineered gene drives. *EFSA Journal* 18, 6297.

Esvelt, K.M., Smidler, A.L., Catteruccia, F. and Church, G.M. (2014) Concerning RNA-guided gene drives for the alteration of wild populations. *Elife* 3, e03401. DOI: 10.7554/eLife.03401.

Gaines, T., Slavov, G., Hughes, D., Kuepper, A. and Sparks, C. (2020) *Investigating the origins and evolution of a glyphosate-resistant weed invasion in South America*. Authorea, Hoboken, New Jersey.

Gantz, V.M. and Bier, E. (2015) Genome editing. The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. *Science* 348(6233), 442–444. DOI: 10.1126/science.aaa5945.

Gould, F., Huang, Y., Legros, M. and Lloyd, A.L. (2008) A Killer–Rescue system for self-limiting gene drive of anti-pathogen constructs. *Proceedings of the Royal Society B: Biological Sciences* 275(1653), 2823–2829. DOI: 10.1098/rspb.2008.0846.

Grunwald, H.A., Gantz, V.M., Poplawski, G., Xu, X.-S., Bier, E. *et al*. (2019) Super-Mendelian inheritance mediated by CRISPR-Cas9 in the female mouse germline. *Nature* 566(7742), 105–109. DOI: 10.1038/s41586-019-0875-2.

Gutzmann, N., Elsensohn, J.E., Barnes, J.C., Baltzegar, J., Jones, M.S. *et al*. (2017) CRISPR-based gene drive in agriculture will face technical and governance challenges. *EMBO Reports* 18(9), 1479–1480. DOI: 10.15252/embr.201744661.

Howe, K., Clark, M.D., Torroja, C.F., Torrance, J., Berthelot, C. *et al*. (2013) The zebrafish reference genome sequence and its relationship to the human genome. *Nature* 496(7446), 498–503. DOI: 10.1038/nature12111.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. *et al*. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337(6096), 816–821. DOI: 10.1126/science.1225829.

Keeling, P.J. (2009) Functional and ecological impacts of horizontal gene transfer in eukaryotes. *Current Opinion in Genetics and Development* 19(6), 613–619. DOI: 10.1016/j.gde.2009.10.001.

Knopf, F., Schnabel, K., Haase, C., Pfeifer, K., Anastassiadis, K. *et al*. (2010) Dually inducible TetON systems for tissue-specific conditional gene expression in zebrafish. *Proceedings of the National Academy of Sciences of the United States of America* 107(46), 19933–19938. DOI: 10.1073/pnas.1007799107.

Kumar, V., Liu, R., Boyer, G. and Stahlman, P.W. (2019) Confirmation of 2,4-D resistance and identification of multiple resistance in a Kansas Palmer amaranth (*Amaranthus palmeri*) population. *Pest Management Science* 75(11), 2925–2933. DOI: 10.1002/ps.5400.

Kyrou, K., Hammond, A.M., Galizi, R., Kranjc, N., Burt, A. *et al*. (2018) A CRISPR-Cas9 gene drive targeting doublesex causes complete population suppression in caged *Anopheles gambiae* mosquitoes. *Nature Biotechnology* 36(11), 1062–1066. DOI: 10.1038/nbt.4245.

López Del Amo, V., Leger, B.S., Cox, K.J., Gill, S., Bishop, A.L. *et al*. (2020) Small-molecule control of super-mendelian inheritance in gene drives. *Cell Reports* 31(13), 107841. DOI: 10.1016/j.celrep.2020.107841.

Marshall, J.M. and Hay, B.A. (2011) Inverse Medea as a novel gene drive system for local population replacement: a theoretical analysis. *Journal of Heredity* 102(3), 336–341. DOI: 10.1093/jhered/esr019.

Montgomery, J.S., Giacomini, D., Waithaka, B., Lanz, C., Murphy, B.P. *et al*. (2020) Draft Genomes of *Amaranthus tuberculatus*, *Amaranthus hybridus*, and *Amaranthus palmeri*. *Genome Biology and Evolution* 12(11), 1988–1993. DOI: 10.1093/gbe/evaa177.

Morio, F., Lombardi, L. and Butler, G. (2020) The CRISPR toolbox in medical mycology: state of the art and perspectives. *PLoS Pathogens* 16(1), e1008201. DOI: 10.1371/journal.ppat.1008201.

Moro, D., Byrne, M., Kennedy, M., Campbell, S. and Tizard, M. (2018) Identifying knowledge gaps for gene drive research to control invasive animal species: the next CRISPR step. *Global Ecology and Conservation* 13, e00363. DOI: 10.1016/j.gecco.2017.e00363.

NASEM (2016) *Gene Drives on the Horizon: Advancing Science, Navigating Uncertainty, and Aligning Research with Public Values*. National Academies of Sciences, Engineering, and Medicine. National Academies Press, Washington, DC.

Neve, P. (2018) Gene drive systems: do they have a place in agricultural weed management? *Pest Management Science* 74(12), 2671–2679. DOI: 10.1002/ps.5137.

Noble, C., Min, J., Olejarz, J., Buchthal, J., Chavez, A. *et al*. (2019) Daisy-chain gene drives for the alteration of local populations. *Proceedings of the National Academy of Sciences of the United States of America* 116(17), 8275–8282. DOI: 10.1073/pnas.1716358116.

Pawluk, A., Amrani, N., Zhang, Y., Garcia, B., Hidalgo-Reyes, Y. *et al*. (2016) Naturally occurring off-switches for CRISPR-Cas9. *Cell* 167(7), 1829–1838. DOI: 10.1016/j.cell.2016.11.017.

Pixley, K.V., Falck-Zepeda, J.B., Giller, K.E., Glenna, L.L., Gould, F. *et al*. (2019) Genome editing, gene drives, and synthetic biology: will they contribute to disease-resistant crops, and who will benefit? *Annual Review of Phytopathology* 57(1), 165–188. DOI: 10.1146/annurev-phyto-080417-045954.

Pollock, J.L., Callaway, R.M., Thelen, G.C. and Holben, W.E. (2009) Catechin-metal interactions as a mechanism for conditional allelopathy by the invasive plant *Centaurea maculosa*. *Journal of Ecology* 97(6), 1234–1242. DOI: 10.1111/j.1365-2745.2009.01553.x.

Prykhozhij, S.V., Rajan, V. and Berman, J.N. (2016) A guide to computational tools and design strategies for genome editing experiments in zebrafish using CRISPR/Cas9. *Zebrafish* 13(1), 70–73. DOI: 10.1089/zeb.2015.1158.

Qi, L., Haurwitz, R.E., Shao, W., Doudna, J.A. and Arkin, A.P. (2012) RNA processing enables predictable programming of gene expression. *Nature Biotechnology* 30(10), 1002–1006. DOI: 10.1038/nbt.2355.

SAM (2017) *New Techniques in Agricultural Biotechnology*, Explanatory No. 02. Scientific Advice Mechanism. European Commission, Brussels.

Scudellari, M. (2019) Hijacking evloution. *Nature* 571, 160–162.

Shah, A.N. and Moens, C.B. (2016) Approaching perfection: new developments in zebrafish genome engineering. *Developmental Cell* 36(6), 595–596. DOI: 10.1016/j.devcel.2016.03.004.

Shapiro, R.S., Chavez, A., Porter, C.B.M., Hamblin, M., Kaas, C.S. *et al*. (2018) A CRISPR-Cas9-based gene drive platform for genetic interaction analysis in Candida albicans. *Nature Microbiology* 3(1), 73–82. DOI: 10.1038/s41564-017-0043-0.

Silva, J.C., Loreto, E.L. and Clark, J.B. (2004) Factors that affect the horizontal transfer of transposable elements. *Current Issues in Molecular Biology* 6, 57–72.

Simoni, A., Hammond, A.M., Beaghton, A.K., Galizi, R., Taxiarchi, C. *et al*. (2020) A male-biased sex-distorter gene drive for the human malaria vector *Anopheles gambiae*. *Nature Biotechnology* 38(9), 1054–1060. DOI: 10.1038/s41587-020-0508-1.

S.O.S. (2020) *Why a global moratorium on the release of gene drive organisms is necessary*. Save Our Seeds. Available at: https://www.saveourseeds.org/en/gene-drive-moratorium.html (accessed 7 April 2021).

UWE (2016) *Science for Environment Policy: Synthetic biology and biodiversity*. University of the West of England, Bristol, UK. Available at: http://ec.europa.eu/science-environment-policy (accessed 5 May 2021).

Vernick, K.D., Dong, Y., Simões, M.L., Marois, E. and Dimopoulos, G. (2018) CRISPR/Cas9-mediated gene knockout of *Anopheles gambiae* FREP1 suppresses malaria parasite infection. *PLoS Pathogens* 14, e1006898.

Wedell, N., Price, T.A.R. and Lindholm, A.K. (2019) Gene drive: progress and prospects. *Proceedings of the Royal Society B: Biological Sciences* 286(1917), 20192709. DOI: 10.1098/rspb.2019.2709.

WHO (2014) *Guidance Framework for Testing Genetically Modified Mosquitoes*. World Health Organization, Geneva.

Windbichler, N., Menichelli, M., Papathanos, P.A., Thyme, S.B., Li, H. *et al*. (2011) A synthetic homing endonuclease-based gene drive system in the human malaria mosquito. *Nature* 473(7346), 212–215. DOI: 10.1038/nature09937.

Index

Gene Flow

Monitoring, Modeling and Mitigation

Edited by Wei Wei and C. Neal Stewart, Jr.

Gene flow is a natural process that occurs spontaneously and enables the evolution of life. However, with the release of genetically modified organisms, concerns have focused on introduced foreign transgenes and their dispersal in nature through gene flow. This book examines gene flow of transgenes, such as herbicide resistance genes, with the goal of understanding the factors that may affect the process of gene flow. A greater biological understanding is essential to make sound management regulatory decisions when also taking into consideration the processes that occur in conventional plants. Monitoring, modeling, and mitigation are the three most closely related elements of gene flow. The book includes both scientific reviews and perspectives on gene flow and experimental case studies, including studies of gene flow in soybean and poplar. The authors present diverse views and research methodologies in order to understand transgene flow. This book:

- Focuses on applications of gene flow (monitoring, modeling, and mitigation).
- Includes both review chapters and case studies.
- Is written by an international team of scientists currently working in gene flow.

This book will be valuable for students and researchers in genetics, biotechnology, plant science, and environmental science. It also provides key insights of value to regulators of biotechnology as well as policy makers.