Investigating the presence of transgenic crop constructs in DNA of organisms of aquatic ecosystems

HJ Venter

orcid.org/0000-0002-9957-8464

Thesis submitted in fulfilment of the requirements for the degree *Doctor of Philosophy in Environmental Sciences* at the North-West University

Promoter: Prof CC Bezuidenhout

Graduation May 2019

20654693

PREFACE

And indeed there will be time
For the yellow smoke that slides along the street,
Rubbing its back upon the window-panes;
There will be time, there will be time
To prepare a face to meet the faces that you meet;
There will be time to murder and create,
And time for all the works and days of hands
That lift and drop a question on your plate;
Time for you and time for me,
And time yet for a hundred indecisions,
And for a hundred visions and revisions,
Before the taking of a toast and tea.

(excerpt from *The Love Song of J. Alfred Prufrock*, by T.S. Elliot)

ACKNOWLEDGEMENTS

“It takes a village to raise a child” – so says the proverb. I have found that the same principle applies to my thesis: it has taken a community, with different levels of support and flavours of influence, for my thesis to “grow up”. Thank you to everyone in my village, I could not have done it without you.

A special word of thanks to my supervisor, Carlos Bezuidenhout. You have been the soul of patience: thank you. I have learned a lot from you.

I am very grateful to my parents, Barry, Fiona and Guy, for their unwavering support. Thank you, not only for putting up with me, but also for actively encouraging me in this pursuit. Thanks also to my sisters for their support.

My partner Odin, aside from being a rock of support and reassurance, has also been my IT consultant and crisis manager. Thank you for all your patience. Thanks also to the Nordgård clan at large, for their collective support.

To my friends and colleagues at the NWU, you have all been an amazing support network and you inspire me with your dedication and creativity. In particular I would like to mention Karen, Bianca, Guzéne, Herman, Alewyn, Audrey, Ina and Abraham. It has been a pleasure to learn with you and from you. I am extremely fortunate to have had wonderful colleagues across the world at GenØk as well, who have taught me a lot and helped keep me going, especially in the last months. Special thanks to Thomas (who co-authored one of these chapters with me), Odd-Gunnar, and Flor. Outside of work, a very special mention to Kara, Mariza, Arnoldeen, Marianne and especially Deon, for exceptional support and motivation: thank you!

I gratefully acknowledge Biosafety South Africa for funding (BSA 09-006) and GenØk – Centre for Biosafety (Norad project GLO-3450) for funding. The views expressed in this thesis do not necessarily represent the views of either organization. I would also like to acknowledge FK Norway, for facilitating my research exchange to Norway, and thus the resulting collaboration.

I would also like to thank my anonymous examiners for the time and effort they put into examining this thesis, and for the recommendations they made for its improvement.

ABSTRACT

Genetically modified Bt crops, which produce proteins toxic to certain agricultural pests, have been cultivated for two decades. In that time, several hundred papers examining their interaction with terrestrial ecosystems have been published, while fewer than 50 have been published which deal with aquatic ecosystems. Aquatic organisms are exposed to Bt crop mostly through deposition of crop detritus due to wind, rain, and runoff. Bt transgenes and proteins are released into the water, present in the crop detritus, and present in the food chain as organisms consume Bt plant material and/or other organisms which have consumed the plant material. Interactions between Bt crops and adjacent ecosystems may be positive, due to the replacement of other, more harmful pesticides. Questions have been raised, however about potential effects on non-target organisms, for horizontal gene transfer (HGT) of transgenes to bacteria, and potential effects on ecosystem services and biogeochemical cycling have also been raised. In this study, DNA-based methods were used to better understand the interactions between Bt maize MON810, and the aquatic environment of the Vaalharts Irrigation Scheme, a farming area in the North-West Province with a high rate of Bt maize adoption. The irrigation system comprises a variety of aquatic environments, including dams, rivers, and canals. Macroinvertebrates and water samples were collected from sites spread through the irrigation system. Samples were also collected from the Tshiombo Irrigation System in Venda, which was used as a control site since no Bt maize was grown in that area at the time of sampling. Macroinvertebrates were identified according to morphology. Microorganisms (bacteria, yeast and fungi) were cultured from the water samples on a variety of media. Following DNA isolation, a PCR-based approach using MON810 primers (CM01 and CM02 and Hsp70 and cry1Ab primers being the most successful of the batch) was used to detect transgene DNA in the DNA isolated from the aquatic organisms. Positive results were detected in 7 macroinvertebrates, 56 bacteria isolates, and 20 yeast and fungi isolates. In the case of the macroinvertebrates, this was taken as an indication of exposure to Bt plant material, most likely through diet. In microorganisms, the presence of transgene sequences was seen as a potential occurrence of HGT. A selection of bacterial isolates was chosen for whole genome sequencing (Illumina MiSeq). These bacteria were identified as *Aeromonas veronii, A. salmonicida, Arthrbacter* sp., *Pseudomonas mendocina, P. protegens, Massilia* sp., and *Serratia fonticola.* It was hoped that detection of transgene fragments in the assembled genomes of the selected organisms would provide information regarding the genomic context of the insertion sites and any genes which had been interrupted due to recombination with transgene fragments. However, after scrutinising the genomes and the sequencing reads using a mapping based approach (Daisy and a BWA-MEM), traces of transgene DNA could not be detected in the isolates' draft assemblies. This may be due to a lack of sequencing coverage in some areas of

the genome, or possibly due to loss of the sequences over time. Though HGT was not detected in this study, there is still a need to take the possibility of HGT of transgene constructs borne by genetically modified plants seriously. This study has contributed towards filling the knowledge gap regarding the interaction of Bt crops and aquatic environments by providing information on exposure of aquatic macroinvertebrates to Bt crops, and by surveying the microbial community for potential HGT of transgenic DNA. A workflow which could be used in future studies detecting transfer of short DNA fragments was developed. Recommendations regarding the monitoring of Bt crops and aquatic ecosystems more generally have been made, as well as suggestions for how HGT of transgenic fragments might be detected in a high-throughput, culture-free method in future.

Keywords: Bt crops, aquatic ecosystems, horizontal gene transfer (HGT), PCR, whole genome sequencing

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND PROBLEM STATEMENT

1.1 General Introduction

1.1.1 GM crops: patterns of use

One of the most significant additions to agriculture in the 20th century was the development of genetically modified (GM) crops (Datta, 2013). Genetic engineering techniques were used to insert useful genes from organisms such as bacteria into the genomes of crop plants, thereby passing the function of the gene on to the recipient plant. The transferred gene is known as a transgene, while the organism (in this case a crop plant) into which it was inserted is then referred to as transgenic (Schouten *et al.*, 2006). The first two decades of GM crop cultivation were largely dominated by two commercial transgenic traits: insect resistance (IR), mostly achieved through the expression of insecticidal proteins from the soil bacterium *Bacillus thuringiensis* (known as Bt crops); and herbicide tolerance (HT), most frequently paired with a glyphosate-based herbicide (Barrows *et al.*, 2014). Transgenic crops which contain combinations of multiple transgenes (such as IR/HT combinations) are known as stacked events, or stacks (Taverniers *et al.*, 2008).

More than twenty years have passed since the commercialization of GM crops in 1996. South Africa was an early adopter of the technology, planting its first GM crops (Bt transgenic maize) in 1997, and maintaining a position among the top ten GM crop planting countries ever since (James, 2014). As can be seen in Figure 1.1, the proportion of GM maize to conventional maize increased steadily in the decade between 2004 and 2014. Hectarage under GM crops was estimated to be 2.7 million ha in 2014, 2.08 million ha of which was GM maize (Bt, HT and IR/HT stacked). Of the total commercial maize hectarage, 67.48% contained Bt insect resistance genes either as a single trait (24.12%), or as part of an IR/HT stack (43.36%). The total area of land under maize expressing Bt transgenes was approximately 1.69 million ha. Another important trend can be seen in Figure 1.1: the increasing proportion of the total maize crop which contains multiple (stacked) transgenes (James, 2014).

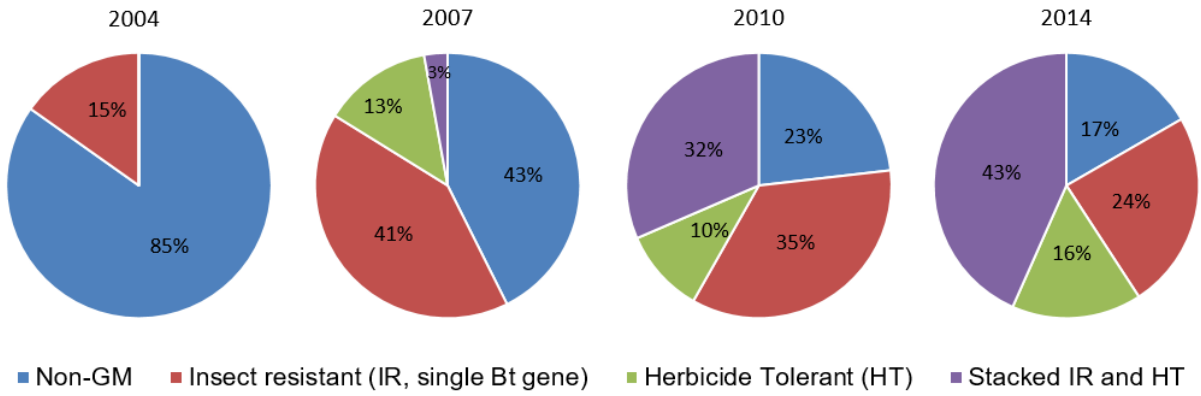


Figure 1.1: **The proportions of maize hectarage as divided between GM maize (insect resistant (IR), herbicide tolerant (HT) and stacked (IR + HT)), and conventional (non-GM maize), in selected years from 2004 to 2014. Graphs compiled from data in ISAAA reports (2004, 2007, 2010, 2014).**

As the first generation of single-trait crops gives way to stacked events and more sophisticated gene editing techniques, knowledge gaps persist, as do questions about environmental impact of GM crops (Hilbeck *et al.*, 2015). Since Bt crops are among the most popular traits, and some (like Bt IR) have been available for two decades, one might expect that Bt crops would be sufficiently well-studied at this point to put to rest uncertainties regarding environmental interactions and impact, however this is not the case (Barrows *et al.*, 2014). Indeed, a large number of studies have looked at the impact on field invertebrates, soil communities, and pollinators such as bees (see these reviews and meta-analyses: Duan *et al.* (2008); Lövei and Arpaia (2005); Marvier *et al.* (2007) Wolfenbarger *et al.* (2008)). Aquatic ecosystems, on the other hand, have experienced consistent neglect as an environmental context in which GM (Bt and others) crops should be studied (Pott *et al*., 2018). This gap is not only reflected in literature, but in risk assessment of GM crops as well. Carstens *et al.* (2012), point out that environmental risk assessment of GM crops in aquatic ecosystems is not well-developed, especially compared to the frameworks in place for terrestrial risk assessment. The South African framework for GM monitoring is discussed in Chapter 3.

1.1.2 Transgenes and their interaction with aquatic ecosystems

Deposition of Bt plant material by wind and rain is the primary route by which Bt transgenes and proteins enter aquatic systems, though runoff and drainage water from fields also contribute (Kratz *et al.*, 2010; Rosi-Marshall *et al.*, 2007; Strain & Lydy, 2015; Tank *et al.*, 2010). Aquatic organisms may thus potentially come into contact with transgenes and Bt proteins through their

diet, in solution in their surrounding water, or bound to sediment, organic particles or algae (Douville *et al.*, 2007; Strain & Lydy, 2015; Wang *et al.*, 2014b).

Most research considering the interaction of Bt crops and surrounding ecosystems (be they aquatic or terrestrial) has focused on the proteins they produce and the possible non-target effects these might have. Interactions between transgenes and the environment have been considered mainly in the context of gene flow (Iversen *et al.*, 2014; Quist & Chapela, 2001) and the possibility of horizontal gene transfer (HGT) of antibiotic resistance marker genes to soil microbes (de Vries & Wackernagel, 2002; de Vries & Wackernagel, 2005; Nielsen *et al.*, 2014). Fewer studies have considered the interactions which the transgenes might have within aquatic ecosystems, or the possibility of HGT of transgenic constructs which do not carry antibiotic resistance marker genes. However, Douville *et al.* (2007; 2009) indicated that transgenes can be used to trace movement of transgenic material through aquatic environments, as well as detect potential incidences of HGT to aquatic organisms. These indicate the ways in which DNA can be used to investigate interactions between aquatic organisms and genetically modified crops. Limited knowledge is available about which aquatic organisms might be vulnerable to the toxins produced by Bt crops (reviewed in Chapter 2). Previous works have measured the amounts of Bt proteins in terrestrial arthropods to gauge which arthropods were exposed to Bt proteins, which also provided information about which life stages experienced the most exposure and elucidated routes multi-trophic transfer of Bt proteins (Harwood *et al.*, 2005; Qing-ling *et al.*, 2013; Yu *et al.*, 2014b). However, by using a PCR-based approach and primers targeted to transgene constructs, the presence of transgene fragments in DNA isolated from aquatic macroinvertebrates can indicate which organisms have been exposed to Bt crop material (Douville *et al.*, 2009). This may indicate which organisms are at risk of non-target effects, and provide some baseline information for future monitoring, as well as identify candidate species for further toxicity testing (of the Bt protein, not the transgene), and is discussed in Chapter 3.

Numerous primer sets have been developed for detecting transgenic constructs, often for the purpose of detecting their presence in food (Kuribara *et al.*, 2002; Matsuoka *et al.*, 2000), though these have also been utilised to track potential incidences of gene flow (Iversen *et al.*, 2014; Quist & Chapela, 2001). Some of these primer sets detect regions such as promoters which are present in a number of transgenic crops: this indicates the presence of transgenic DNA, but cannot determine which crop event it originated from (Kuribara *et al.*, 2002). Other primer sets are event-specific, meaning if those sequences are detected, the crop event it came from will also be known (Hernández *et al.*, 2003; Matsuoka *et al.*, 2000).

Such primer sets may also be used to detect HGT of transgene constructs to aquatic microorganisms, if the constructs can be detected in DNA isolated from these organisms (Douville *et al.*, 2009). Horizontal gene transfer (HGT) is the transfer DNA between organisms, via mechanisms other than those involved in vertical inheritance (Keese, 2008; Ravenhall *et al.*, 2015a). Among prokaryotes, there are three main mechanisms of HGT: transformation, conjugation and transduction. Natural transformation occurs when bacteria take up DNA from their surroundings and integrate it into their genomes. Transformation occurs most frequently (though not exclusively) between related organisms, since the presence of regions of homology between donor and recipient DNA allows easier integration of incoming DNA (Thomas and Nielsen, 2005). The bacterial origin of many transgenes may mean that regions of homology of various lengths exist between the transgenes and the genomes of environmental bacteria, potentially facilitating integration of transgenic constructs into bacterial genomes, a principle which was demonstrated by Kay *et al.* (2002). However, mechanisms such as illegitimate recombination and homology- and micro-homology facilitated illegitimate recombination mean that less-homologous fragments may also be integrated, albeit at lower frequencies than homologous recombination (de Vries & Wackernagel, 2002; Kohli *et al.*, 1999).

Detection of transgenic constructs may be done by PCR-based methods with targeted primer sets as mentioned above. For investigation of horizontal gene transfer, the insertion site is also of interest. Techniques which have been used to characterise gene insertion sites include primer-walking, Southern blotting, TAIL-PCR, and inverse PCR are methods (Domingues *et al.*, 2012a; Hernández *et al.*, 2003; Quist & Chapela, 2001). Next-generation sequencing (NGS) techniques, including whole genome sequencing (WGS), and the wealth of information they provide, have also recently been harnessed for such purposes. WGS, followed by assembly and analyses of these genomes, indicates the context into which sequences have been integrated (Domingues *et al.*, 2012a). This approach may provide information about the insertion site and flanking regions, whether recombinations took place upon insertion, and whether any endogenous genes have been interrupted. Another approach using WGS data, is to view potential insertions (in this case MON810 transgene constructs) as structural variants. Recently, a programme (Daisy) has been developed specifically for the detection of variants caused by HGT (Trappe *et al.*, 2016).

The consequences of HGT are dependant both on the sequence and function of the transferred DNA, as well as on where the DNA is integrated in the recipient organism's genome. HGT may prove beneficial or detrimental to an organism, or may be neutral and have no discernible effect at all (Koonin, 2009). Possible positive effects include gaining useful genes, such those conferring antibiotic resistance (Nielsen & Townsend, 2004; Wellington *et al.*, 2013); elimination of selfish mobile genetic elements (Croucher *et al.*, 2016); and repair of faulty genes (de Vries &

Wackernagel, 2002; Overballe-Petersen *et al.*, 2013), all of which may allow the receiving organism to adapt well to its environment. Possible negative effects, on the other hand, include interruption the recipient cell's own genes by the inserted DNA, leading to alteration of gene products and possibly loss of function of such products; increased metabolic burden; and the insertion of regulatory sequences such as promoters which may bring about changes in the regulation of transcription patterns of endogenous genes (Baltrus, 2013).

As one of the first generation of GM crops, Bt crops can act as a case study of what happens when large quantities of technology-driven genes are released into the environment (Douville *et al.*, 2007). The present study considers transgenes both in terms of potential HGT to aquatic microorganisms, and as a means of investigating exposure of non-target organisms to Bt crops in aquatic ecosystems.

1.2 Problem Statement

The interaction between genetically modified crops and aquatic ecosystems is an area which has received insufficient attention in terms of the environmental impact of GM crops, despite some transgenic crops having been commercially available for over two decades. The role of transgenes in this interaction, and the potential for horizontal gene transfer of transgenic constructs to aquatic organisms, has also received very little attention. In addressing these knowledge gaps, molecular methods targeting transgenic constructs were used both to investigate potential horizontal gene transfer and provide information regarding exposure of non-target aquatic organisms to transgenic Bt crop materials. Next-generation sequencing techniques were used to investigate potential occurrences of HGT to environmental bacterial isolates.

1.3 Research aim and Objectives

Aim:

To determine whether transgene fragments associated with genetically modified maize can be detected in DNA isolated from aquatic organisms, and to annotate sites of insertion in selected isolates if detected, in order to address knowledge gaps related to the interaction between Bt crops and aquatic organisms.

Objectives:

1.) Review the available literature regarding the interaction of GM crops and aquatic organisms and ecosystems and highlight knowledge gaps.
2.) Collect water samples and aquatic macroinvertebrates from the Vaalharts Irrigation Scheme (study site) and Tshiombo Irrigation Scheme (reference site). Use a culture-based approach to cultivate bacteria, yeast and fungi from the water samples.
3.) Isolate DNA from the aquatic organisms collected.
4.) Use a PCR-based approach to detect transgenic constructs in the DNA isolated from aquatic organisms.
5.) Use next generation sequencing (whole genome sequencing) to confirm whether integration of the transgenic constructs into the genomes of selected bacteria has occurred.
6.) If insertion of the transgene has occurred, use bioinformatics software, such as the CLC Genomics workbench, to annotate and characterise the site of insertion into the genomes of recipient organisms.

1.4 Outline of Thesis

Chapter 1 provides a brief introduction to GM crops and their patterns of use, as well as the fate of transgenes in the environment. A problem statement is provided, before the research aims and objectives are given.

Chapter 2 is a review article which focuses on the interaction between Bt crops and aquatic ecosystems. This synthesis analyses the available literature and highlights knowledge gaps.

Title:	Interactions between Bt transgenic plants and aquatic ecosystems
Authors:	Venter, H.J., Bøhn, T.
Journal:	Environmental Toxicology and Chemistry
Manuscript number:	ETCJ-Apr-16-00294

Chapter 3 considers the South African GM crop monitoring framework, and how molecular techniques using DNA (PCR, DNA barcoding, meta-barcoding) might be used to augment current monitoring efforts. Knowledge of the organisms present in the aquatic communities, as well as which are exposed to Bt crop materials, provides a starting point for future biomonitoring endeavours, as well as options for research into the effects of GM crop technology on aquatic species (e.g. selection of organism for toxicity tests). This chapter is partially based on an article (DNA-based identification of aquatic invertebrates – useful in the South African context?, Venter and Bezuidenhout (2016)). A second publication “PCR-based detection of transgenic crop

constructs in DNA isolated from aquatic macroinvertebrates" will also be submitted to the South African Journal of Science.

Title:	DNA-based identification of aquatic invertebrates – useful in the South African context?
Authors:	Venter, H.J., Bezuidenhout, C.C.
Journal:	South African Journal of Science
Manuscript number:	sajs.2016/20150444

Chapter 4 reports on the collection of water samples, cultivation of aquatic microorganisms, and isolation of DNA and detection of transgenic constructs using PCR. Sanger sequencing of PCR amplicons for identification of the organisms, as well as comparison of constructs with reference constructs, was also conducted.

In chapter 5, selected bacterial isolates which gave positive results during standard PCR to detect transgenic constructs, were subjected to whole genome sequencing. Chapter 5 reports on the assembly and scrutiny of the genomes the selected bacteria.

Chapter 6 contains a general discussion. It considers the results of the previous chapters of the study together, and provides context and perspective by comparison with previous studies. It also contains recommendations and the conclusion.

CHAPTER 2

INTERACTIONS BETWEEN BT-TRANSGENIC PLANTS AND AQUATIC ECOSYSTEMS

Abstract

Bt crops collectively refer to crops which have been genetically modified to include a gene (or genes) sourced from *Bacillus thuringiensis* (Bt) bacteria. These genes confer the ability to produce proteins toxic to certain insect pests. The interaction between Bt crops and adjacent aquatic ecosystems has received limited attention in research and risk assessment, despite the fact that some Bt crops have been in commercial use for 20 years. Reports of effects on aquatic organisms such as *Daphnia magna*, *Elliptio complanata* and *Chironomus dilutus* suggest that some aquatic species may be negatively affected; while other reports suggest that decreased use of insecticides precipitated by Bt crops may benefit aquatic communities. In the present study, the literature regarding entry routes and exposure pathways by which aquatic organisms may be exposed to Bt crop material is considered, as well as feeding trials and field surveys which have investigated the effects of Bt-expressing plant material on such organisms. The development of Bt crops beyond single gene events, towards multigene stacked varieties - which often contain herbicide resistance genes in addition to multiple Bt genes - is discussed, as well as how their use (in conjunction with co-technologies such as glyphosate/Roundup) may impact and interact with aquatic ecosystems.

2.1 Introduction: Interactions between Bt transgenic plants and aquatic ecosystems

Aquatic ecosystems are facing significant pressures that threaten natural dynamics, ecological integrity and biodiversity (Ward, 1998). Dominating stressors that reduce biodiversity include land use, homogenization of resources, eutrophication and habitat destruction (Stendera *et al.*, 2012). All these factors are intimately linked to modern agriculture. In fact, agriculture is highlighted as a key driver of environmental change in freshwater ecosystems (Friberg, 2010).

Modern commercial agriculture is predominantly characterized by large-scale monoculture production. In the mid-1990s transgenic crops with internally produced toxins to combat insect pests were introduced to the market. These crops, particularly for maize and cotton, now contribute significantly to the world markets, reflected in the large production volumes of the US, Canada, Brazil, Argentina, Paraguay, India, China, Pakistan and South Africa (James, 2014).

The term ‘Bt crops’ is the collective term for crops which have been genetically modified to include a gene (or genes) sourced from *Bacillus thuringiensis* (Bt) bacteria, which code for insecticidal proteins. Bt genes confer the ability to produce insecticidal proteins to the crop plants themselves, reducing the need to spray chemical insecticides to control pest insects (Marvier *et al.*, 2007; Naranjo, 2009). The *B. thuringiensis* toxins which have been harnessed in Bt crops include parasporal crystal proteins, known as Cry proteins (Cyt proteins when they exhibit cytolytic activity), as well as vegetative insecticidal proteins, known as VIPs (Bravo *et al.*, 2007; Crickmore *et al.*, 1998; Ibrahim *et al.*, 2010; Schnepf *et al.*, 1998). Bt toxin-producing crops have been used for controlling pests of Lepidoptera, Coleoptera, Diptera and Hymenoptera, as well as nematodes (Bravo *et al.*, 2007). *Cry1Ab* is arguably the best-studied of the Bt genes, though the International Service for the Acquisition of the Agri-biotech Applications (ISAAA, 2016) currently lists 21 Bt genes in commercial use (including truncations and other modified versions – for an overview, see Table 2.1), and hundreds of Bt toxins have been reported (Crickmore, 2016).

Single transgene GM crops are being replaced by varieties that combine or “stack” several transgenes, incorporating multiple Bt toxins and/or other traits, such as herbicide tolerance (HT), in the same plants. The genetic modifications of HT crops allow them to be used in tandem with specific herbicide co-technologies: such herbicides can be applied multiple times during a growing season for weed control without major damage to the HT crop plants (Duke, 2005). Crops which stack Bt and HT traits are particularly popular and have overtaken single transgene Bt crops in terms of area planted in recent years (James, 2014). James (2014) further reports that approximately 80 million hectares of crops containing Bt genes were grown worldwide in 2014, with maize and cotton as the dominant crops. Bt crops produce feed, food and fibre, but a large amount of biomass (leaves, stalks, cobs and roots which remain after harvest, as well as pollen) enter local food-web interactions in soil and aquatic ecosystems.

The assessment of environmental safety is crucial and a key element of transgenic crop technology (Naranjo, 2009). Research on potential non-target effects of Bt transgenic plants has focused on terrestrial ecosystems, and investigations have predominantly tested Cry1Ab-toxin and Cry1Ab transgenic crops, while other genes/toxins and stacked events (especially in conjunction with herbicide co-technologies) have received limited attention. Despite growing recognition that aquatic ecosystems near agricultural fields receive significant amounts of run-off and crop residues that contain these toxins (Böttger *et al.*, 2015; Li *et al.*, 2013), environmental risk assessments of transgenic crops tend to neglect aquatic ecosystems as a relevant context for testing.

The present study reviews literature related to exposure, spread, break-down rates and effects of various types of Bt crop material on non-target organisms and aquatic communities. Finally, recommendations for research to fill existing knowledge gaps are made.

2.2 Entry routes and Exposure pathways of Bt toxins and plant material

Aquatic ecosystems receive much of their energy from terrestrial systems. Basic aquatic ecology, e.g. through the 'river continuum concept' (Hynes, 1975), has shown that the energy input into a small stream can be significant – often larger than the local energy production within the stream in shaded areas. This highlights the role which allochthonous input can play in aquatic ecosystems. In agricultural settings, the source of allochthonous input is likely to be crop detritus from the surrounding farmland. Other natural links between terrestrial and aquatic systems are include some insects that spend different life-stages in aquatic and terrestrial environments, fish that feed on terrestrial insects, etc.

Regarding such connection in terms of Bt crop fields and adjacent aquatic ecosystems, Carstens *et al.* (2012) differentiate between entry routes and exposure pathways. The ways and means by which Bt crop materials (including plant material, Bt proteins and transgenes) end up in an aquatic ecosystem are the entry routes. Exposure pathways refer to the routes by which aquatic organisms may be in contact with Bt material and affected by it (Carstens *et al.*, 2012).

2.2.1 Entry routes

The main entry route of Bt materials into aquatic systems, seems to be the deposition of plant debris, including pollen, crop dust, leaves, stalks and post-harvest detritus, facilitated by wind, rain and run-off (Kratz *et al.*, 2010; Rosi-Marshall *et al.*, 2007; Strain & Lydy, 2015; Tank *et al.*, 2010). Among the first to investigate the fate of Cry1Ab/Bt corn by-products in agricultural streams was Rosi-Marshall and co-workers. They investigated the presence of Bt maize stratus in 12 headwater streams of agricultural production areas in the Midwest of the USA (Rosi-Marshall *et al*., 2007). Following this work, 217 Indiana streams were sampled by Tank *et al*. in 2010. It was found that, 6 months after harvest was complete, 67% of the streams had maize leaves in the stream channel, while 86% contained other maize detritus in addition to leaves; and that the average concentration of Cry1Ab in streams which tested positive for the protein was 14 ± 5 ng/L (Tank *et al.*, 2010).

Table 2.1: Compiled from information on the ISAAA GMO Approval Database, this Table indicates the various crop plants which have been modified to express Bt proteins, as well as the target orders of each protein (ISAAA, 2016).

Bt Genes	Crops where present	Listed target organisms
cry1A	Cotton (1 event), maize (1 event)	Lepidoptera
cry1A.105	Maize (18 events), soybean (1 event)	Lepidoptera
cry1Ab	Cotton (8 events), maize (53 events), rice (2 events)	Lepidoptera, particularly European corn borer, African corn borer
cry1Ab (truncated)	Maize (1 event), rice (1 event)	Lepidoptera
cry1Ab-Ac (synthetic fusion gene)	Cotton (2 events)	Lepidoptera
cry1Ac	Cotton (28 events), eggplant (1 event), maize (1 event), poplar (2 events), rice (2 events), soybean (4 events), tomato (1 event)	Lepidoptera
cry1C	Cotton (1 event)	Lepidoptera, particularly Spodoptera
cry1F	Cotton (6 events), maize (4 events), soybean (2 events)	Lepidoptera
cry1Fa2 (synthetic form of *cry1F*)	Maize (45 events)	Lepidoptera
mocry1F (synthetic form of *cry1F*)	Maize (1 event)	Lepidoptera
cry2Ab2	Cotton (10 events), maize (20 events), soybean (1 event)	Lepidoptera
cry2Ae	Cotton (4 events), maize (1 event)	Lepidoptera
cry3A	Potato (30 events)	Coleoptera
cry3Bb1	Maize (18 events)	Coleopterans, particularly corn rootworm
cry9C	Maize (1 event)	Lepidoptera
mcry3A	Maize (30 events)	Coleoptera
ecry3.1Ab	Maize (5 events)	Coleoptera and Lepidoptera (Multiple insect resistance)
cry34Ab1	Maize (35 events)	Coleoptera, particularly corn rootworm
cry35Ab1	Maize (35 events)	Coleoptera, particularly corn rootworm
vip3A(a)	Cotton (9 events)	Lepidoptera
vip3Aa20	Maize (22 events)	Lepidoptera

The amount of crop biomass which reaches aquatic ecosystems will be affected by different agricultural and conservation practices. These include conservation tillage and the adoption of riparian buffers. Conservation tillage, which includes practices such as mulch-till, strip-till, no-till etc., refers to production systems in which at least 30% of crop residues are left on the field to prevent soil erosion and water loss (Fernandez-Cornejo *et al.*, 2013). Other outcomes include improved soil structure and increased nutrient cycling, better drainage, and increased available crop material (Holland, 2004) - some of which will enter into aquatic ecosystems (Tank *et al.*, 2010). Interestingly, a correlation has been shown between the use of HT crops and the adoption of conservation tillage practices (Fernandez-Cornejo *et al.*, 2013). Taken together with the increasing use of Bt/HT crops, this may indicate an entry route for herbicides, in addition to Bt toxins, via herbicide-treated Bt plant materials.

Riparian buffers, however, may help counter the entry of crop debris into aquatic environments (Jensen *et al.*, 2010; Moore & Palmer, 2005). Riparian buffers are zones of vegetation (such as grasses, shrubs or trees) which are planted in order to form a barrier between the fields and streams, to reduce the amounts of sediment, nutrients (such as nitrates and phosphates), and runoff entering streams in order to improve water quality (Jaynes *et al.*, 2014; Tomer *et al.*, 2015). By impeding the flow of run-off, they may limit the transfer of crop material into streams, though this is likely to depend on the type and density of the vegetation making up the buffer. The degree to which riparian buffers impede the entry of Bt crop residues into adjacent streams has not (to our knowledge) been examined.

During storms or floods, the amount of crop plant material brought to a local stream or pond can be massive, though dilution due to a large volume of water should also be taken into account when considering Bt toxins in this context. Small streams or ponds may become densely packed with plant material, as can be seen in Figure 2.1. Conversely, some portion of the deposited crop material may be fine particles, such as those generated when whole maize plants are harvested for silage or methane production. These particles (coarse, >1 mm; fine <1 mm) may be a food source for aquatic invertebrates (Kratz *et al.*, 2010). Furthermore, researchers in Canada found that rivers and streams could be implicated in spreading Bt materials away from the immediate surroundings of maize fields, after they detected transgenic DNA from Bt maize several kilometres downstream from the fields where the maize was grown (Douville *et al.*, 2007).

Figure 2.1: Run-off material of Bt-transgenic maize to local stream after flood, South Africa (photo credit: T. Bøhn, with permission).

Entry routes for Bt proteins include transport from fields into aquatic ecosystems as part of erosion or runoff (Saxena *et al.*, 2002), or via drainage water and tile drains (Tank *et al.*, 2010). Bt proteins are released into soil from living Bt crops via their roots, and also from dead plant tissues which remain on the field (Saxena & Stotzky, 2000; Wang *et al.*, 2013a). Soil properties will influence the amount of protein entering the aquatic system via this route. Clay particles appear to bind strongly to Cry1Ac, Cry1Aa, Cry1Ab and Cry1Ab/1Ac fusion proteins, and reduce biodegradation of these proteins (Helassa *et al.*, 2009; Li *et al.*, 2013; Saxena & Stotzky, 2000; Wang *et al.*, 2013a). This has led some authors to theorise that soils which have high clay content may keep Bt toxins close to the soil surface, and lead to a higher rate of bioactive Bt proteins in the run-off

soil of these systems (Saxena *et al.*, 2002). Recent studies, however, have found that Bt protein concentration in the soil of no-tillage Bt maize fields tended to be low (averaging below 5 ng/g before pollination, peaking at 9-29 ng/g during pollination). Higher concentrations of the Cry1Ab protein were detected in surface water runoff and runoff sediment, which increased during the growing season, peaking during pollination at 130 ng/L and 143 ng/g DW (dry weight) for runoff water and sediment respectively (Strain & Lydy, 2015; Whiting *et al.*, 2014). Interestingly, these studies also detected Cry1Ab protein in the runoff of a non-Bt field located close to the Bt field in question, at an average concentration of 14 ng/L. Given that no Cry1Ab protein was detected in the soil of this field even during pollination, the authors speculated that its presence in the runoff was due to the transfer of plant materials between fields due to rain (Strain & Lydy, 2015).

Subsurface drains (tile drains) may represent an alternative route of entry for Bt proteins which have been desorbed from soil particles, as was found by Tank *et al.* (2010). However, considering that only 2 out of 120 groundwater and pore water samples drained from a Bt field analysed by Strain and Lydy (2015) contained detectable amounts of the protein (17.2 and 21.7 ng/L, respectively), this route may not contribute much Bt protein to aquatic systems.

2.2.2 Exposure pathways

Deposited plant material is itself available for consumption (Axelsson *et al.*, 2011; Böll *et al.*, 2013; Chambers *et al.*, 2010; Kratz *et al.*, 2010), and leaches Bt proteins into the water (Griffiths *et al.*, 2009; Li *et al.*, 2007b; Wang *et al.*, 2013b). The proportion of Bt proteins which remain in the plant tissue versus the amount which leach into the water and/or degrade after exposure to aquatic environments has been documented in a number of studies. Several complicating factors including temperature, type of plant tissue, sediment composition and influence of microbes have been noted (Li *et al.*, 2013; Li *et al.*, 2007b; Strain & Lydy, 2015).

Bt-toxin concentration of plant material in aquatic settings

During a field experiment in which rice stubble was left on the field after harvest, Li *et al.* (2007b) found that the Cry1Ac concentration of Bt rice stalks (originally 1501.3 ± 200.5 ng/g DW) dropped by 50% during the first month after harvest, but that the rate of degradation slowed subsequently. Seven months after harvest, 21.3% (319.8 ± 59.8 ng/g DW) of the original Cry1Ac toxin concentration was still present in stalk tissue. Cry1Ac leaching from rice roots (original concentration 516.1 ± 86.4 ng/g DW) followed a different pattern however, with initial release of the toxin being quite slow: 72.4% of the original concentration was present after 1 month.

The concentration decreased to almost un detectable levels by the end of the experiment, 7 months post-harvest. The authors also noticed that, for both stalks and roots, the winter months brought a reduction in the rate of Cry1Ac degradation. The field experiment did not include measurements of the amount of Cry1Ac present in the field soil or water.

Experiments which tracked the decrease in concentration of Cry1Ac protein in Bt rice plant residues in soil compared to an aquatic milieu have been performed under laboratory (Li *et al.*, 2007b) and field conditions (Xiao *et al.*, 2015). In both cases, degradation was found to be somewhat slower under aquatic conditions, at least initially. Under laboratory conditions, degradation of Cry1Ac protein in soil - despite faster initial degradation - plateaued eventually, leaving 15.3% of the initial concentration in the leaf-soil mixture after 135 d, while none was detectable in water by this point (Li *et al.*, 2007b). In contrast, Xiao *et al.* (2015), were unable to detect the protein in the soil surrounding the litterbags of plant materials in the field. This lack of consistency is thought to reflect different methodologies in terms of sample preparation (Xiao *et al.*, 2015).

In contrast to Cry1Ac rice, Cry1Ab concentrations in Cry1Ab maize plants decline more rapidly under aquatic conditions than in soil or aerobic conditions (Böttger *et al.*, 2015; Douville *et al.*, 2005). It has been shown that within the first hour of aquatic exposure, 61% of the Cry1Ab toxin leached from Bt maize leaves (the Cry1Ab concentration in the water was not determined, however) (Griffiths *et al.*, 2009). Strain and Lydy (2015) similarly found that Cry1Ab had a half-life of approximately 2 hours, but that the concentration of the protein in the water peaked at around 2 days after initial exposure. The proportion of Cry1Ab reported to remain in Bt leaves over time varies between studies. It has been reported as 6% and 20% of the initial concentration after 21 days and 70 days exposure to aquatic conditions, respectively (Böttger *et al.*, 2015). Wandeler *et al.* (2002) reported that after 20 days, one variety of Bt maize experienced a reduction in Cry1Ab concentration of 60%, while another decreased by only 21%. Although the plant material in the Wandeler study was not exposed to an aquatic environment, it reflects the variation which differences in cultivar or environmental conditions can introduce.

2.2.3 Degradation of Bt proteins in aquatic settings

Bt proteins which are leached into the water degrade over time, though there is great variation among the reports of how long this takes to happen. Strain and Lydy found that the proportion of Cry1Ab protein in the water decreased to below reporting limits over approximately 2 weeks (Strain & Lydy, 2015). However, a study of Cry1Ac extracted from cotton seeds found that the Cry1Ac protein was still detectable in water and sediment after 60 days (Li *et al.*, 2013). Prihoda and Coats (2008) found that the half-life of Cry3Bb1 from MON863 Bt corn stalks, leaves and roots was just under 3 days. They were also unable to detect Cry3Bb1 protein in the water or sediment of the microcosm treatments, which the authors attributed to rapid adsorption by organic particles, or swift dissipation. However, Strain *et al.* (2014) suggest that this lack of detection may be due to the methodology used (i.e. not concentrating the water samples before determining concentration).

Differences in cultivar, as well as factors such as water chemistry and temperature may account for differences in the rate of Bt protein loss (Böttger *et al.*, 2015; Strain *et al.*, 2014). Temperature in particular is an important factor for Bt protein longevity (Li *et al.*, 2013; Strain & Lydy, 2015). For instance, when the temperature was kept at 4°C, decline of Cry1Ab concentration in plant material and water was much slower than at warmer temperatures. Cry1Ab concentrations in both matrices dropped below reporting level in approximately 2 weeks when incubation was at 37°C. When temperature was held at 4°C, the average concentration of Cry1Ab in the aquatic milieu after 2 months was 300 ng/L (Strain & Lydy, 2015). This indicates that the stability of Bt proteins could be extended during the cooler winter months (as was also noted for Bt rice by Li *et al.* (2007b)), which is significant because a great deal of plant material is present in aquatic environments during that time (Jensen *et al.*, 2010; Strain & Lydy, 2015; Tank *et al.*, 2010). However, another factor to consider is whether the Bt proteins retain bioactivity after prolonged presence of plant materials in water.

2.2.4 Adsorption of Bt proteins

Bt proteins which are leached from plant material may bind to sediment, especially sediment with a high clay and/or organic matter content (Li *et al.*, 2013; Strain & Lydy, 2015). In 2 month experiment, 20 to 40% of the total Cry1Ab protein present in a system of submerged Bt maize plant material was located in the sediment from the second week onwards (Strain & Lydy, 2015). Adsorption to sediment particles protects the Bt proteins from degradation, and may also allow them to keep their toxic/insecticidal properties (Stotzky, 2005). Cry1Ac leached from cotton persisted in sediment longer than in soil, which the authors attributed to greater amounts of

organic matter in sediment having reduced the bioavailability of Cry1Ac, and thereby reduced its degradation by microbes (Li *et al.*, 2013).

Also, leached Bt proteins may be adsorbed by algae: Cry1Ca protein was detected in cells of the green alga *Chlorella pyrenoidosa*, after it was cultured in media containing leachate from Cry1Ca-expressing rice (Wang *et al.*, 2014c). The amount of Cry1Ca present in the algae cells increased with increasing concentration of the protein in the media, but reached saturation at a concentration of 1000 µg/mL of the media. Interestingly, when the Cry1Ca protein concentration was too low to be detected in the culture medium, it could be detected in the algae (Wang *et al.*, 2014b). Given the rapid adsorption of Bt protein reported in these studies, one may question whether adsorption by algae could affect the measurements of Bt proteins in aquatic field samples? Algae as a potential route of exposure for aquatic organisms has not yet been investigated.

To summarise, organisms inhabiting aquatic environments adjacent to Bt crops will potentially be exposed to Bt-containing plant material and Bt toxins at varying concentrations, depending on their feeding habits, the type of crop and cultivar, the age and breakdown rates of the plant material, and the properties of the water and sediment of the aquatic environment. The timing of feeding in relation to when Bt crop material enters the system and how long it was exposed will also be important. Though some authors have suggested that the concentrations of Bt proteins to which aquatic organisms are exposed are too low to cause concern (Carstens *et al.*, 2012; Wolt & Peterson, 2010), others argue that the continuous input of crop debris, as well as runoff water and sediment, may lead to long term exposure of aquatic organisms which may have chronic effects which warrant further investigation (Böttger *et al.*, 2015; Kratz *et al.*, 2010; Strain & Lydy, 2015; Tank *et al.*, 2010).

2.3 Activation and specificity of Bt toxins

In target insects, the toxicity of Bt proteins has been studied in some detail, and although different models exist for exactly how they cause harm to insects, all the models agreed on the following: after ingestion, solubilization of protoxin form of the Bt toxins in the alkaline midgut was required, proteolytically activating them to their smaller active toxin form (Aimanova *et al.*, 2006; Jurat-Fuentes & Adang, 2006; Tabashnik *et al.*, 2015; Vachon *et al.*, 2012). However, recent studies of resistant target insects have found that protoxin activation was not necessary for an insecticidal effect to occur. In fact, in some cases, the protoxin was more effective than the activated toxin (Gómez *et al.*, 2014; Tabashnik *et al.*, 2015). It is theorised that bacterial production of protoxins is a strategy to impede resistance development (Tabashnik *et al.*, 2015). This pathway has only been described in resistant target orders so far, but considering that protoxin activation has been

linked to Bt toxin specificity, one might enquire whether these findings might have implications for cross-activity within other orders (Li *et al.*, 2007a).

Another important finding highlighted by Tabashnik *et al.* (2015) and Gómez *et al.* (2014), was that the use of mammalian trypsin or chymotrypsin proteases to activate the protoxins does not produce results exactly equivalent to protoxins activated by insect midgut juices, potentially leading to underestimations of toxicity. Additionally, bacterially-produced Bt protoxins may differ from those produced in some Bt crops: MON810 and Bt 11 events produce truncated (65 kDa) toxins, in contrast to the bacterially-produced 135 kDa protoxin (Douville *et al.*, 2005; Mendelsohn *et al.*, 2003). Furthermore, activation by plant proteases within Bt crops was reported by Li *et al.* (2007a), meaning that insects feeding on such plants are exposed to activated toxins. The points mentioned here provide examples of mismatches between what is tested during risk assessment and what is found in the environment, since most safety testing of Bt toxins is performed on Bt protoxins produced by bacteria, typically in *E. coli*, (not by the transgenic crops), and if activation is done, it is usually with mammalian proteases.

2.3.1 Specificity of Bt-toxins and sensitivity of aquatic non-target organisms

The specificity of Bt-toxins – that their effectivity is restricted to a limited range of target organisms (usually restricted to a specific order) - has been lauded as a major advantage for agricultural application because unwanted negative effects on non-target organisms can be minimised (Avisar *et al.*, 2009; Torres & Ruberson, 2008). However, documented negative effects on non-target organisms, such as *Daphnia magna* (Bøhn, Primicero *et al.*, 2008; Bøhn, Rover *et al.*, 2016; Bøhn,Traavik *et al.*, 2010; Holderbaum *et al.*, 2015; Raybould & Vlachos, 2011), which lack the relevant receptors, point to alternative modes of action for Bt toxins. van Frankenhuyzen (2013) reviewed the subject of cross-activity of Bt toxins outside of their primary target orders. The study found that while 64% of the 148 Bt toxins considered were thought to be active within 1 order only, a large portion of these had in fact never been tested on organisms from different orders. Evidence of cross-activity was found in approximately 13% of the Bt toxins investigated (van Frankenhuyzen, 2013).

Given the exposure of Bt toxin in aquatic systems and the uncertainty over alternative mode-of-actions/cross-reactivity of some Bt toxins, more testing of the sensitivity of aquatic organisms seems well justified. In addition, some of the aquatic insects exposed will belong, at some level, in the same taxonomic groups as the terrestrial target pest species. The main insect orders targeted by Bt toxins are Lepidoptera, Diptera and Coleoptera (see Table 2.1). These are biodiversity rich groups with representatives found in aquatic environments. Aquatic stages of e.g.

larval caddisflies, beetles or midges may be vulnerable to Bt toxin exposure, depending on the toxin concentration, the feeding strategy and the sensitivity of each individual species. Depending on the degree of relatedness, these groups may share physiological properties, receptors, etc. which may make them vulnerable to Bt toxins (Rosi-Marshall *et al.*, 2007).

In terms of investigating non-target effects, a few studies set out to characterise the degree to which arthropods were exposed to Bt proteins, regardless of whether they belonged to the target orders, by determining the amount of Bt protein present in arthropod specimens collected from fields of Bt crops (Harwood *et al.*, 2005; Qing-ling *et al.*, 2013; Yu *et al.*, 2014a). What was fascinating about these studies was that, aside from providing baseline data for which species were potentially at risk due to Bt proteins, they were also able to provide data indicating how the concentration of Bt proteins present in the arthropods differed throughout the growing season, especially how these differed before and after anthesis. Furthermore, in some cases it was possible to detect during which life stages the arthropods were most exposed, i.e. contained the highest levels of Bt protein. Exposure pathways of Bt proteins through the food chain were also illuminated, since Bt protein levels were measured in predator species as well (Yu *et al.*, 2014a). Similar investigations of aquatic communities would help determine which species are the most exposed, and which may be good candidates for further investigation.

2.4 Effects of Bt toxins on aquatic organisms

Though effect studies testing Bt-expressing plant material and Bt-toxins, i.e. toxicological testing, feeding trials and field trials are being done with greater frequency with aquatic organisms, large knowledge gaps are still present. Most studies have been done using maize or rice producing Cry1Ab or Cry1Ac. Soy, rapeseed, cotton, and other Bt crops, as well as numerous Bt toxins, are meanwhile underrepresented in terms of investigations of potential effects on aquatic organisms.

2.4.1 Caddisflies

Caddisfly (Trichopteran) larvae have attracted attention as aquatic organisms which may be affected by Bt crops, due to the close relation of the Trichoptera to Lepidoptera, the target order of many Bt toxins. Despite this, only 3 species appear to have been put through feeding trials (Chambers *et al.*, 2010; Jensen *et al.*, 2010; Rosi-Marshall *et al.*, 2007). The results of these studies have been at times contradictory, inconclusive and/or controversial. Methodological issues have been at the root of most of these discrepancies.

The 3 caddisfly species which have been investigated so far are *Lepidostoma lima*, *Pycnopsyche scabripennis* and *Helicopsyche borealis*. *L. lima* experienced reduced daily growth rates of more than 50% (p=0.008) when fed Bt-corn litter as compared to non-Bt litter. Another caddisfly, *Helicopsyche borealis*, was shown to have increased mortality, but this response required exposure to a high concentration of pollen, i.e. 2-3 times higher than maximum pollen density observed in the field (Rosi-Marshall *et al.*, 2007). The Rosi-Marshall *et al.* (2007) study has been criticised for the lack of appropriate controls, for failing to quantify relevant properties of the crop material, especially the Bt-toxin, and also for overstating its conclusions (Beachy *et al.*, 2008; Parrott, 2008). Rosi-Marshall and colleagues (2007) matched the lignin content and the carbon to nitrogen ratio in the maize, rather than using isolines of Bt versus non-Bt treatments. They argued that the nutritional value would not be same with the latter control since Bt-maize contains markedly more lignin compared to its isogenic counterpart (Saxena & Stotzky, 2001).

In a follow-up to the Rosi-Marshall study, Chambers *et al.* (2010) increased the number of Bt and non-Bt varieties sampled, and again found *L. liba* to be negatively affected by Bt maize in terms of growth. In contrast, Jensen *et al.* (2010) did not find any significant differences between *L. liba* larvae fed Bt maize and those fed the non-Bt near isoline. It was noted by these authors that obtaining a true isoline to use as a control can be problematic, making causality difficult to establish. Considering the increased use of stacked events which contain multiple transgenes, the problem of availability of suitable reference material (controls) will become more difficult and complex.

2.4.2 Sedimentary midges

Chironomus dilutus, a filter-feeding aquatic midge frequently used for sediment toxicity testing (OECD, 2010), was known to be sensitive to Bt insecticidal formulations before the advent of GM Bt crops (Dickman, 2000; Yiallouros *et al.*, 1999). *C. dilutus* larvae have since also been found to be sensitive to Cry3Bb1 (Prihoda & Coats, 2008) and Cry1Ac (Li *et al.*, 2013), extracted from Bt maize roots and cotton seeds, respectively. The 2 studies had very different approaches: Prihoda and Coats (2008) exposed *C. dilutus* to Cry3Bb1 by adsorbing the protein to food particles, while Li *et al.* (2013) spiked the sediment or water in which the organisms were placed, with Cry1Ac extract.

A significant reduction in survival was observed in larvae exposed to nominal concentrations of 30 - 48 ng/mL Cry3Bb1 (measured concentrations of Cry3Bb1 adsorbed to food were 19 - 27 ng/mL), though the amount consumed was not determined. The LC50 of Cry1Ac cottonseed

extract was determined to be 155 ng/g DW of spiked sediment and 201 ng/mL in spiked water. Both sets of authors pointed out that the concentrations of Bt protein used in the experiments were far above what these larvae would likely be exposed to in nature. While this is true, it should also be kept in mind that empirical data on aquatic environmental concentrations of Cry3Bb1 and Cry1Ac (or any other Bt protein produced in cotton) are lacking; and the exposure regimes used in these studies do not reflect the "pulsed", chronic exposure which is likely to occur in the field (Prihoda & Coats, 2008; Strain & Lydy, 2015). Furthermore, as filter feeders, *C. dilutus* may feed directly on Bt crop dust (Kratz *et al.*, 2010), which is an avenue of exposure which remains to be investigated.

2.4.3 Waterfleas

The waterflea *Daphnia magna*, a commonly used ecotoxicology model, fed powder from cry1Ab-expressing Bt maize (MON810) kernels over the whole life-cycle showed early reproduction, but reduced survival and fecundity in later life-stages compared to the near-isogenic maize, indicating a weak toxic effect of the Bt-maize (Bøhn *et al.*, 2008). Exposure to the risk of a predator (smell of 3 spined stickleback that had eaten *D. magna*) increased the differences in fitness and population growth rate between animals fed Bt-maize and non-Bt-maize (Bøhn *et al.*, 2010). A follow-up study showed that also Bt-transgenic leaves were capable of producing negative effects in *D. magna* after chronic dietary exposure. Animals fed Bt-maize leaves showed the typical stress-response of producing more resting eggs (Holderbaum *et al.*, 2015).

D. magna have also been used in testing non-target effects of Bt rice. An experiment in which water from Bt and non-Bt rice paddies was used as an environment for culturing *D. magna* indicated that the water from Cry1Ab/1Ac- and Cry2A-expressing rice paddies was less toxic than water taken from the non-Bt control paddy. After 13 d, survival of *D. magna* in water from non-Bt water was 0%, while those cultured in water from Cry1Ab/1Ac-expressing rice was 60% and Cry2A-expressing rice was 52%. Notably, in this study, insecticides were sprayed on both Bt and non-Bt paddies, in a manner which was designed to make the yields of the different rice fields the same, which meant that the non-Bt field received more insecticide. Thus, the mortality in the non-Bt groups likely was caused by the higher concentration of insecticides, potentially masking effects of the Bt toxin. The decreased use of insecticides is a frequently-cited advantage of growing Bt crops, with environmental benefits, as shown in this study (Li *et al.*, 2014). However, the normal amount of pesticide to use on a rice paddy is not given, nor what the yield of an alternative pest management strategy might have been under similar conditions (omitting the insecticide and the Bt). On the other hand, as the authors point out, insecticide use has led to a

drastic reduction in rice paddy biodiversity, a situation which the use of Bt rice has been shown to alleviate (Li *et al.*, 2014).

A study (from Syngenta) showed reduced growth rate for *D. magna* exposed to high doses (0.75 mg/L) of microbially produced Vip3A Bt-toxin over 10 days (Raybould & Vlachos, 2011), indicating weak toxicity or some other physical or chemical effect of the Vip3A test substance. Similarly, exposure to purified Cry1Ab, Cry2Aa toxins (0.75 – 4.5 mg/L) resulted in higher mortality in a life-cycle study by Bøhn *et al.* (2016).

2.4.4 Aquatic vertebrates

In terms of fish, most studies have been done concerning the use of Bt crops in fish food for commercial species, namely salmon and zebrafish, rather than fish likely to live in aquatic environments adjacent to Bt crops. Overall, consumption of Cry1Ab-expressing maize did not have a detrimental effect on the survival, growth and development of either species, though some stress responses were observed in salmon (Gu *et al.*, 2014; Jørgensen, 2012; Sagstad *et al.*, 2007; Sissener *et al.*, 2010). These included decreased activity of certain digestive enzymes (Gu *et al.*, 2014), increased activity of superoxide dismutase and changes in white blood cell population (Sagstad *et al.*, 2007). Conversely, zebrafish growth improved, transcription of superoxide dismutase decreased (Sissener *et al.*, 2010) and white blood cell populations remained unchanged (Jørgensen, 2012).

Frogs are a common feature of rice paddies, and as such have received some attention as non-target aquatic organisms. A feeding study with the model organisms *Xenopus laevis* (Zhu *et al.*, 2015) and field study with *Rana nigromaculata* froglets (Wang *et al.*, 2014a) exposed to Cry1Ab/1Ac-producing rice indicated that neither frog species was significantly physically affected by the Bt protein. *R. nigromaculata* did display some differences in feeding pattern, which is unsurprising considering that there was a substantial decrease in stem borers in the Bt rice paddies. While Cry1Ab/1Ac-expressing rice did not affect tadpole densities, pesticides applied to non-Bt fields did elicit a decrease (Wang *et al.*, 2014a).

2.4.5 Field studies

Field studies have examined the potential influence of Bt crops on communities of organisms, as well as individual species, under realistic conditions. The cost of these realistic conditions is, however, that it is extremely difficult to exclude confounding factors. For example, Douville *et al.* (2009) detected fragments of the Cry1Ab-transgene in tissues of freshwater mussels (*Elliptio complanata*) as well as waterways in proximity to Bt maize fields. At the most exposed site, the

level of recombinant DNA in mussel organs was at its highest, and this coincided with a significantly reduced condition factor (weight/length) and oxidative stress in the mussels. Despite this correlation, the authors could not exclude the influence of chemical pollutants as possible confounding factors for the decreased condition factor (Douville *et al.*, 2009).

In 2 studies of overall invertebrate abundance and diversity, neither Swan *et al.* (2009) nor Chambers *et al.* (2010) reported significant effects on arthropod communities that could be explained by the Cry1Ab Bt maize. However, some taxon-specific effects on certain invertebrates were noticed by Swan *et al.* (2009). *Pycnopsyche* sp. was 11 times more abundant on non-Bt maize compared to single-gene Bt maize litter, while *Caecidotea communis* was statistically less abundant on a stacked Bt event (containing both Cry1Ab and Cry3Bb1) than on the non-Bt and single-gene Bt treatments. Both studies also reported that overall degradation of the aquatic environment made the influence of a single potential stressor, such as a Bt toxin, difficult to discern.

An increase in abundance of aquatic invertebrates was noted on Bt poplar leaves (which expressed Cry3Aa) when leaves were submerged in the Tavelån stream in Sweden. Of the 2 Bt poplar lines tested, 1 displayed an increase in invertebrate abundance of 33%, while the other showed a 25% increase, compared to the isogenic control (Axelsson, *et al*., 2011). The community assemblages of the Bt lines were similar to each other, though both differed from the control. This indicates that Bt crops do have the potential to bring about changes in aquatic invertebrate community composition (Axelsson *et al.*, 2011), though such changes need not necessarily be detrimental.

In the same vein, a much higher abundance and diversity of zooplankton (rotifers, cladocerans and copepods) was found in paddies of Bt rice expressing Cry1Ab/1Ac or Cry2A, compared to non-Bt paddies (Li *et al.*, 2014). As with the studies mentioned above involving frogs, this positive effect was attributed to the difference in pesticide regime between Bt and non-Bt rice paddies, and not to the Bt rice as such. When no pesticides were used, no significant differences in zooplankton density were observed between Bt and non-Bt paddies (Wang *et al.*, 2013b).

There is a need for more field-based studies, and especially for studies which consider food web interactions over time. This includes investigations of transfer of Bt toxins through the food web, as well as possible effects related to prey quality, diversity, and abundance. Interactions of Bt insecticidal sprays (produced using *Bacillus thuringiensis* var. *israelensis* (Bti), and widely used to control mosquitoes) were investigated in a multiyear study conducted in the wetlands of the Camargue area on the French Atlantic coast. Odonata (dragonfly) populations were reported to

be negatively affected due to decreases in Chironomid populations in Bti-sprayed areas compared to control sites. After accounting for factors such as distance from the sea and land cover types between control and Bti-sprayed areas, the authors concluded that Bti use was a potential threat to Odonata (Jakob & Poulin, 2016). Though the results of food web interaction studies of Bti sprays are mixed, and Bt crops and Bti sprays differ in several ways (limiting the usefulness of direct comparisons), further monitoring of aquatic ecosystems to determine the non-target effects of these toxins is warranted (Bøhn *et al.*, 2016; Jakob & Poulin, 2016; Lagadic *et al.*, 2014).

Trophic perturbations reaching vertebrate (bird) populations were also indicated in the Camargue region by Poulin *et al.* (2010). They reported reduced clutch size and fledgling survival of house martins (*Delichon urbicum*), which they explained as an indirect effect of the use of Bti spray, correlating the use of the spray with decreased availability of certain house martin prey (Nematocera and the spiders and dragonflies which feed on them). This study was criticised for having its reference sites too far from its study sites, and a later study by Lagadic in a different area of France did not find any effects due to the application of Bt sprays, in agreement with a number of other studies (Caquet *et al.*, 2011; Davis & Peterson, 2008; Fayolle *et al.*, 2015; Lagadic *et al.*, 2016).

2.4.6 Microbes and horizontal gene transfer

Few studies have been published on the effects that Bt crops may have on aquatic microbial communities, though data from studies on soil microbes suggest that effects will be slight and transient, if they manifest at all (Baumgarte & Tebbe, 2005; Oliveira *et al.*, 2008). Studies on submerged soil communities in rice paddies have followed this pattern. Although changes in microbial phosphatase and dehydrogenase activity have been observed in flooded soils amended with Bt rice straw (Wu *et al.*, 2004a; Wu *et al.*, 2004b), a subsequent study (using the same Cry1Ab-producing rice) did not find this to be the case (Liu *et al.*, 2008). In-stream microbial respiration rates were shown to be the same for both Bt and non-Bt maize residues (Griffiths *et al.*, 2009).

Considering the huge amount of attention which the possibility of horizontal gene transfer of transgenic fragments (especially antibiotic marker genes) to soil and gut microbes attracted when GM crops were first released (Nielsen & Townsend, 2004; de Vries *et al.*, 2003; Gebhard & Smalla, 1998; Mohr & Tebbe, 2007), it is surprising that this topic appears to have been almost entirely passed over in the context of aquatic bacteria. This is possibly due to the manifold challenges encountered when trying to detect such occurrences in soil environments (reviewed

in (Nielsen *et al.*, 2014)). That being said, Douville *et al.* (2009) detected *Cry1Ab* transgene fragments in heterotrophic bacteria present in water samples taken from water sources near Bt corn fields. Freshwater mussels (*Elliptio complanata*) taken from the same sites contained *Cry1Ab* transgene fragments in their gills, digestive glands and gonads. Mussels may however also filter small-sized Bt-maize particles directly, and although the presence of pollen particles was not indicated by the authors, this is difficult to eliminate as a source. Given the filter-feeding habits of mussels, and the detection of transgene fragments in heterotrophic bacteria, the authors speculated that bacteria transformed with transgenic DNA might be responsible for its presence in the mussels, and also that such bacteria acted to maintain and stabilise the presence of the *Cry1Ab* transgene in the aquatic environment (Douville *et al.*, 2009).

2.5 Ecosystem wide effects – might Bt genes have community and ecosystem properties?

When genes have properties that may transform community structure and ecosystem processes, those genes contribute to 'community or ecosystem phenotypes' (Schweitzer *et al.*, 2004; Whitham *et al.*, 2006). For instance, the variation in condensed tannins (which are genetically controlled) in poplar trees, results in multiple effects on higher level properties of the community and even on ecosystem processes (Schweitzer *et al.*, 2004). This case illustrates how a single quantitative trait locus has significant effects in the entire ecosystem: on the community composition, endophyte community, related aquatic community and also nitrogen mineralization and aquatic decomposition (Whitham *et al.*, 2006). Poplar trees thus serve as a model system to study how a single but well characterised gene may affect not only the individual and population, but also organisms that directly and indirectly interact with that gene (Schweitzer *et al.*, 2004).

Whitham *et al.* (2006) argue that genetically modified organisms with traits like insect resistance and herbicide tolerance in genetically modified (GM) plants, or enhanced growth in GM fish, represent a special case of exotic introductions of ecological novelty with several uncertain effects on community and ecosystem phenotypes. Stacked events with more than one cry-gene will strengthen the plant protection, but at the same time may widen the spectrum of potential non-target and cascading effects in the food web. Moreover, adding herbicide co-technology on top of insect resistance in the same plant may increase potential for ecosystem impact.

2.6 Unintended effects

Occasionally, changes are wrought in transgenic crops which do not directly relate to the intended activity of the inserted transgene (Hjältén *et al.*, 2013): increased lignin content of Bt maize, for

instance (Saxena & Stotzky, 2001). Pleiotropy, positional effects and insertional effects may bring about unintended changes in gene expression and regulation (Schnell *et al.*, 2015). In some cases, organisms feeding on Bt crop material have shown negative effects, despite inactivity of the Bt protein (determined through assays with susceptible organisms), leading to questions about unintended effects in these plants.

Jensen *et al.* (2010) observed negative effects on 2 aquatic species: the isopod *Caecidotea communis* and the crane fly larvae *Tippula abdominalis*. This study included single event Cry1Ab maize as well as a stacked event which contained Cry3Bb1 in addition to Cry1Ab. Crane fly larvae experienced 19.6% less growth as well as lower mass when fed Bt maize compared to non-Bt maize. *C. communis* fared worse under Bt feeding regimes (both single-gene and stacked events) compared to non-Bt isolines. Negative effects were noted for body length (49.7% less), final mass (50% less) and survival (43.3% decrease). Considering that a sensitivity assay performed with *Ostrinia nubinalis* on the pre-conditioned maize leaves found no bioactivity of the Bt toxins past 2 weeks of environmental exposure, the authors suggested that the effects seen were due to tissue-mediated factors or micronutrient differences between the near-isolines, rather than due to the Bt toxins themselves.

Li *et al.* (2015) tested the effects of Cry1C and Cry2A-producing rice lines on *Propylea japonica*, and noted that larval development was significantly longer when fed Bt pollen compared to control pollen. However, when the toxins were provided in their pure forms (mixed with rapeseed pollen) at more than 10 times the concentration in the rice pollen, *P. japonica* was not affected. This suggests that unintended changes were induced in the rice genomes when the transgenes were inserted, which caused the negative effects observed on the phenotype. It is interesting, however, that both rice lines should have the same type of effect, despite expressing different Bt toxins and having different insertion events.

2.7 Stacked events and resistance evolution

A challenge for the use of Bt-transgenic crops is to ensure that their solution to insect attacks is sustainable, i.e. that the solution will last over time. The development of resistance to Bt-toxins in pest insects, e.g. in South Africa, India, China, South America and the US (Farias *et al.*, 2014; Tabashnik, 2015; Tabashnik *et al.*, 2009; Van den Berg, 2013) has shown that pest insects may evolve quickly to regain food resources protected by Bt toxins. Resistance development has led to a gradual replacement of the first generation Bt crops that had single cry genes only. A growing majority of Bt crops now contain 2-6 cry-toxin genes at the same time (Niu *et al.*, 2014; Raybould *et al.*, 2012). This means that the total expression of Bt-toxin per plant can be expected to multiply

with a factor of 2-6, which may lead to increased negative effects on a range of non-target organisms. In addition to the higher total expression of Bt-toxins in stacked events, the fact that different Bt toxins aim at target pest species from different taxa, is likely to expand the range of sensitive non-target organisms. Industry data show that stacked Bt-expressing plants may express higher levels of Bt toxins compared to parental lines. Smartstacks maize express on average Cry1A.105 at a 54% and 97% higher level than the parent line MON89034 (Stillwell & Silvanovich, 2007).

High throughput sequencing programs may offer huge potential for the discovery of new Bt-toxin sequences in future (Palma *et al.*, 2014). That may help manage the increasing occurrence of resistance to Bt-based insecticides or transgenic Bt crops (Ferré *et al.*, 2008).

2.8 Added or combinatorial effects of stacked events?

The general trend is that Bt crops with insect protection only seem to be on their way out, replaced by stacked events with herbicide tolerance (HT) as well. By 2012, nearly twice as large area of GM crops with both insect protection and HT traits (mainly for glyphosate) were grown as compared to plants only expressing Bt toxins. In South Africa, 64% of the production of Bt maize in 2014 was also HT (James, 2014). In the US, in 2015, approximately 90% of the maize grown carried HT traits (USDA-ERS., 2015). This means that herbicide co-technology regularly co-occurs with Bt toxins in the environment. Glyphosate-based herbicides are the most common herbicide co-technologies used in conjunction with HT crops.

Herbicide tolerant crops are specifically designed to be used in combination with herbicides and cannot fulfil their function without them. Due to development of glyphosate resistance in weeds, the amount of active ingredient used has increased both per hectare and in total. A 15-fold rise in glyphosate use has been documented globally since the introduction of Roundup Ready HT crops in 1996 (Benbrook, 2016). Surprisingly, co-technology herbicides are usually not tested as a part of the risk-based science and assessment of HT crops. In 13 out of 16 published feeding studies with such crops analysed by Viljoen, 2013, test material had not been sprayed with the relevant herbicide, which represents a serious flaw in the testing procedures. This may lead to underestimated effects of such material on non-target species, not least in aquatic ecosystems (e.g. exposed to run-off crop residues) where it is documented that glyphosate/Roundup is more toxic than previously reported (Cuhra *et al.*, 2013; Relyea & Hoverman, 2006). Furthermore, safety testing (both environmental and health) of the agrochemicals themselves tends to focus on the active ingredient of the herbicides (e.g. glyphosate), and not the commercial formulations

(e.g. Roundup), resulting in unrealistic and misleading safety assessments because the influence of adjuvants is excluded from the testing regime (Mesnage *et al.*, 2014; Surgan *et al.*, 2010).

Stacked events are relatively easy to produce, as conventional breeding is used to combine 2 or more events that are already approved singularly for the market. The maize hybrid MON 89034 x 1507 x MON 88017 x 59122, from Monsanto and Dow AgroSciences, is broadly resistant to insects with 6 different cry genes (*cry1A.105, cry1F, cry2Ab2, cry3Bb1*, *cry34Ab1* and *cry35Ab1*) and tolerant to both glyphosate and glufosinate ammonium herbicides. Multiple-trait stacked crops may be expected to have a wider range of effects (Then, 2010), both on target pest insects (as intended) and on non-target organisms (unintended and unwanted) from their multiple Bt-toxins. However, very few studies have tested plant material from multi-Bt crops.

GM crops recently approved for the market in some countries have tolerance to several herbicides, i.e. different combinations of glyphosate, glufosinate ammonium, 2,4-D and dicamba. One such example is the DAS-44406 soybean, tolerant to glyphosate and glufosinate and 2,4-D (DowAgroSciences., 2011). Tolerance to multiple herbicides is already combined with multiple Bt-toxins, i.e. up to 6 Cry-toxins in the case of 'Smartstacks' maize from Monsanto. Thus, multiple Bt-toxins will co-occur with one or several co-technology herbicides in the same environment. Research on interactions and potential synergies between these toxins is crucial (Then, 2010).

As innovation in agriculture moves towards more complex stacked transgene combinations, with multiple insect-toxins and tolerance to multiple co-technology herbicides/pesticides within the same plant, there is an increasing need to test i) their potential environmental consequences, both as single compounds and as combinations, and ii) assess their risks. One assumption, endorsed by the Food and Agriculture Organization, World Health Organization, Organisation for Economic Co-operation and Development, the International Seed Federation, and Crop Life International is that stacked transgenic events will function as a sum of their separate parts (Pilacinski *et al.*, 2011). Accordingly, safety of stacked transgenic events can be based on previous safety assessments of the single-transgene parental events, i.e. no relevant interactions between the stacked traits are expected. This is controversial and studies of combinatorial effects of multiple stressors are increasingly acknowledged as missing, e.g. from the European Food Safety Authority, EFSA (EFSA, 2008) and elsewhere (Al-Gubory, 2014). Bøhn *et al.* (2016) showed that co-exposure to Cry1Ab and Cry2Aa resulted in higher mortality in the aquatic waterflea *D. magna*, supporting the hypothesis that stacked events may cause stronger effects on non-target organisms.

In addition to the potential additive (or combinatorial) effect of multi-Bt crops, the spraying of agricultural fields throughout the growing season with one or several herbicides on stacked (i.e. Bt/HT) GM crops, will add stress for aquatic organisms that live in nearby aquatic ecosystems. Bio-active herbicides ultimately get into water courses directly or indirectly through processes such as drifting, leaching and surface runoff (Mensah *et al.*, 2012c). Negative effects of herbicides are documented for a number of aquatic species related to the most relevant of the herbicides: glyphosate-based formulations (Roundup), i.e. on amphibians (Relyea & Hoverman, 2006; Relyea, 2005a; Relyea, 2005b; Sih *et al.*, 2004), shrimps (Mensah *et al.*, 2011; Mensah *et al.*, 2012c), and waterfleas (Cuhra *et al.*, 2013).

2.9 Conclusion

Bt and HT crops are in many ways the pioneers of the genetically modified crop movement. The reasons for highlighting the gaps in monitoring and risk assessment are not simply to determine whether these specific crops themselves are problematic, but to consider gaps in the risk assessment of GM crops generally. Almost 10 years of planting Bt crops passed before aquatic ecosystems were seriously considered within risk assessment, and another decade has passed since then with limited improvement. Decline of aquatic biodiversity due to pesticides and agrochemicals is well-documented (Beketov *et al.*, 2013; Fleeger *et al.*, 2003; Relyea, 2005a), but despite this, approximately 90% of major agricultural areas worldwide have not been included in investigations of pesticide concentrations in surface water (Stehle & Schulz, 2015), and the possible contribution of Bt crops to this situation has not been considered. There are significant knowledge gaps for the fate of Bt crops and their potential effects in aquatic systems.

CHAPTER 3

EXPOSURE OF AQUATIC MACROINVERTEBRATES TO BT MAIZE RESIDUES: DNA-BASED METHODS FOR DETECTION AND MONITORING IN A SOUTH AFRICAN REGULATORY CONTEXT

Abstract

The interaction between genetically modified Bt crops and aquatic ecosystems has received very little attention in terms of risk assessment and monitoring compared to terrestrial ecosystems. Our aims were to apply a PCR-based transgene detection method to investigate exposure of aquatic organisms to Bt crop material, in order to gain better understanding of the interaction between Bt crops and aquatic ecosystems. To achieve this, macroinvertebrates were collected from five sites spread across the Vaalharts Irrigation System (North West Province), as well as a control site in the Tshiombo Irrigation System (Limpopo). Following identification according to morphology, total DNA was isolated from the macroinvertebrate specimens. PCR with primers associated with Bt maize MON810 was then done to detect transgene fragments. Seven of the 68 macroinvertebrate individuals yielded positive results, indicating exposure to Bt maize plant residues in water, most likely through diet. In the course of this study, it also became clear that DNA-based monitoring methods such as DNA barcoding and metabarcoding have great potential for the monitoring of aquatic macroinvertrebrates in South Africa. However, this approach is currently underutilised, at least partially due to a lack of supporting data present in public databases.

3.1 Introduction

3.1.1 Monitoring GM crops in South Africa

South Africa is currently the 9th largest grower of genetically modified (GM) crops in the world, and the largest in Africa, with an estimated 2.7 million hectares of land used to grow GM crops in 2016 (James, 2016). The potential effects of Bt crops on non-target organisms has been a point of interest for Bt crops since they were first commercialised. Though a large number of studies have looked at the impact on field invertebrates, soil communities, and pollinators such as bees (Duan *et al.*, 2008; Lövei & Arpaia, 2005; Marvier *et al.*, 2007; Wolfenbarger *et al.*, 2008), aquatic ecosystems have experienced consistent neglect as an environmental context in which Bt (and other genetically modified crops) crops should be studied. A recent review by Pott *et al.* (2018) found only 39 publications which dealt with transgenic crops and aquatic ecosystems. This gap is not only reflected in literature, but in risk assessment of GM crops as well. Carstens *et al.* (2012)

point out that environmental risk assessment of GM crops in aquatic ecosystems is underdeveloped compared to the frameworks in place for terrestrial risk assessment.

In South Africa, the framework for considering risks to the environment can roughly be divided into activities which assess potential impacts at the pre-release phase, and those which follow general release (McGeoch & Rhodes, 2006). The Genetically Modified Organisms Act (GMO Act 15 of 1997 and its amendment Act 23 of 2006) requires that an environmental risk assessment (ERA) must be submitted as part of the application for a permit to conduct activities involving genetically modified organisms. ERAs may be paper-based (essentially literature studies of GM crop research), or field trial-based (crops are tested under local conditions). Field trial-based ERAs are relatively uncommon. Hence, most permits are granted or refused on the basis of paper-based ERAs which are reviewed by the Advisory Committee which falls under the framework of the Department of Agriculture, Forestry and Fishing (DAFF) (McGeoch & Rhodes, 2006). The paper-based approach of most South African GMO ERAs suggests that there is a large gap in both empirical knowledge and GMO risk assessment as far as aquatic ecosystems are concerned. This fits well within the context of the general lack of research into the interaction between GM crops and aquatic ecosystems noted in chapters 1 and 2.

After commercial release, DAFF continues to have a role in ensuring that permit conditions are upheld, and compliance monitoring of this kind involves both the GM crop producers (permit holders) as well as inspectors appointed by DAFF. For the permit holders, this has historically largely been limited to resistance detection and management, refuge compliance, and weed resistance monitoring, depending on the conditions stipulated in the permit (Bothma, 2006; SANBI, 2011). In terms of monitoring the effects of GM crops on the environment and biodiversity, however, there is a shift in oversight from DAFF to the Department of Environmental Affairs (DEA). The South African National Biodiversity Institute (SANBI), under the auspices of the National Environmental Management: Biodiversity Act (NEMBA, Act 10 of 2004), was mandated to monitor and research potential impacts and risks associated with the release of GMOs into the environment and report on these to the minister (Bothma *et al.*, 2010; DEAT, 2009; McGeoch & Rhodes, 2006).

3.1.2 Case-specific monitoring and general surveillance

SANBI's approach to this mandate is broadly similar to the post-market monitoring strategy adopted by the European Commission, and recommends both case-specific monitoring and general surveillance (DEAT, 2009; Züghart *et al.*, 2013; Züghart *et al.*, 2007). Case-specific monitoring examines the effects associated with a particular GM crop. Up to now, SANBI has

conducted case-specific research on one event, cry1Ab-producing MON810 maize (SANBI (2011); Bothma *et al.* (2010)), though other institutions have conducted studies on some non-target organisms (Erasmus & Van Den Berg, 2014; Erasmus *et al.*, 2010). This type of monitoring addresses specific concerns and determine whether aspects of ERAs and EIAs hold true. General surveillance, conversely, detects impacts which were not, or possibly could not be, predicted during earlier assessments (SANBI, 2011).

In the case of the European framework, it was suggested that GM crop monitoring could be integrated with other general surveillance monitoring strategies which are already in place. These strategies are not designed to focus on GM crops specifically, but on environmental health in general. However, though useful in some aspects, such as collection of baseline data, it has been found that such programmes (at least in their current state) are not able to provide data on all the aspects of GM crop monitoring (Smets *et al.*, 2014; Züghart *et al.*, 2007). Lack of standardization of monitoring techniques and methods specifically required for GM crops are some of the obstacles noted (Züghart *et al.*, 2013).

Challenges also exist for general surveillance of GM crops like Bt maize in South Africa. The issues experienced are common to monitoring overall (not just of GM crops) – it is expensive, requires expertise, and may take a long time before trends are clear (ACB, 2010; Birkhead, 2014; Bothma *et al.*, 2010). And yet long-term general surveillance data is vital for monitoring ecosystem health. Several authors have pointed out that conservation of natural resources and ecosystems hinges on the provision of data regarding the presence and distribution of species within an environment (Baird & Hajibabaei, 2012; Dallas & Rivers-Moore, 2014; Thomsen & Willerslev, 2014) – data which are usually supplied through biomonitoring programmes, including SASS, the South African Scoring System. SASS is one of the programmes that forms part of the current National Aquatic Ecosystem Health Monitoring Programme (NAEHMP): River Health Programme (RHP), and analyses macroinvertebrate communities as a measure of stream ecosystem health (DWAF, 2008). SASS uses morphology to identify macroinvertebrates (usually to family level), and records these on a scoresheet. Aquatic taxa have been assigned weighted values which are a measure of each taxon's tolerance of pollution and perturbation. The sum of these values gives the SASS score, which, together with other parameters such as habitat quality and average score per taxon, is used to gauge the relative well-being of the system in comparison with a reference site (Dallas, 2007; Dickens & Graham, 2002).

Although the identification part may be useful to the general surveillance of aquatic ecosystems associated with GM crops, there are several factors which impede the use of the SASS matrix for determining the state of such ecosystems. Firstly, SASS was developed for studying flowing

perennial rivers; its matrix is not applicable for lentic aquatic systems such as dams, which are a fixture of the South African farming landscape. It is estimated that there are 500 000 dams in South Africa, most of them farm dams (Mantel *et al.*, 2010). Furthermore, selection of reference sites which both reflect a non-GM agricultural landscape and meet the 'least-impacted' criteria for SASS reference sites will be difficult for two reasons: the widespread use of Bt and HT crops (see Figure 1.1); and the fact that many streams associated with agricultural systems are under pressure from numerous stressors and understanding the influence of a single potential stressor (such as Bt crops) under such conditions has proven difficult (Chambers *et al.*, 2010; Swan *et al.*, 2009).

Given that research into the interaction of GM crops (such as Bt crops) with aquatic ecosystems is lacking both in South Africa and internationally (see sections 2.2, 2.4 and 2.5 of Chapter 2), surveys based on collection and identification methodologies of programmes such as SASS could provide valuable baseline data regarding the macroinvertebrate communities inhabiting aquatic systems associated with Bt GM crop cultivation.

Combining surveys of arthropods inhabiting fields of Bt soybean and maize with the detection of cry proteins (the toxins produced by Bt crops), provided evidence regarding which species were in contact with Bt proteins (and during which life stage), as well as data regarding how the amount of Bt proteins present differed throughout the growing season and how they were transferred through the food web (Harwood *et al.*, 2005; Yu *et al.*, 2014b). In aquatic ecosystems surrounding Bt crops, macroinvertebrates may be exposed to transgenic proteins and plant material in their water, as well as through their diet. Exposure through diet may occur through direct consumption of transgenic plant material, or via consumption of prey which fed on Bt material, though research in this area for aquatic organisms is scant (see section 2.5.5, Chapter 2).

Transgenic DNA sequences are likewise useful in tracing movement of transgenic material in the environment. Event-specific primers have been used to detect gene flow of transgenes (Iversen *et al.*, 2014; Quist & Chapela, 2001), as well as track the movement of transgenic material through aquatic ecosystems (Douville *et al.*, 2007) and detect their presence in the tissue of freshwaters mussels inhabiting these ecosystems (Douville *et al.*, 2009). This indicates that molecular techniques involving the proteins and DNA of transgenic crops could be useful in monitoring strategies.

DNA can also be used to augment general ecosystem monitoring, by using specific DNA sequences to identify the organisms present (Baird & Hajibabaei, 2012). The standardisation of DNA-based identification methods, known as DNA barcoding, gave rise to the Consortium for the

Barcode of Life. This initiative was set up with the goal of promoting the use of specific regions of DNA (the COI gene for animals, *rbcl* and *matK* for plants) to determine the sequence of those regions which was particular to each species, for use in identification. DNA-based identification methods are particularly useful for identification of cryptic species (Janion *et al.*, 2011; Nadel *et al.*, 2010), juveniles and females (Ekrem *et al.*, 2010), as well as rare species, all of which are commonly among the macroinvertebrates collected during sampling of aquatic ecosystems, but are very challenging to identify morphologically. The Barcode of Life Database (BOLD) is a centralised database in which such barcode sequences, as well as specimen collection and species distribution details, are available to any interested person. The African Centre for DNA Barcoding reports that 12 548 plant species and 1493 animal species had been barcoded in South Africa by July 2013, while the Barcode of Life Database (BOLD) contains 62 926 records for South Africa, 11 392 of which had species names, representing 4541 species (BOLD Systems, 2014). Entries for many aquatic invertebrate taxa of South Africa are currently lacking however, which hinders identification using DNA barcodes – as was discovered by the authors during this study (Venter & Bezuidenhout, 2016).

3.1.3 Study site description: the Vaalharts Irrigation System, an aquatic ecosystem in association with Bt Maize (MON810)

The Vaalharts Irrigation Scheme is the largest irrigation scheme in South Africa, providing water for 36 000 ha of farmland. Vaalharts is located in the Northern Cape at the confluence of the Vaal and Harts rivers, 1050-1150 m above sea level. It is situated in semi-arid conditions where the average annual rainfall is approximately 450 mm, with most of this rainfall occurring in the summer months (Herold *et al.*, 1996; Kruger *et al.*, 2009). Crops grown in this region include wheat, lucerne, groundnuts, pecan nuts, grapes, olives, and barley, as well as maize (Herold *et al.*, 1996; Otieno & Adeyemo, 2011). Bt maize has been grown in this area since its release in South Africa in 1998. MON 810 was the preferred Bt event in the region at the time of sampling, with 96% of farmers choosing to grow this event at the time of sampling (Kruger *et al.*, 2009).

The irrigation scheme comprises of a network of canals which carry water from the Vaalharts Weir (between Christiana and Warrenton) to the boundaries of the farms in the scheme (Herold *et al.*, 1996). The Main Canal (18.4 km) splits into the West Canal (22 km) and the North Canal (82 km), and together they make up the three primary canals. From these canals, a number of feeder and tertiary canals extend to the farms. There are 5 balancing dams, as well as numerous small overnight dams within the scheme. Natural drainage of farmland in Vaalharts is insufficient to prevent waterlogging, so a network of subsurface drains (also known as tile drains) was installed to assist with drainage. This assists in keeping the water level in the fields below root level. Water

from the drains is collected in drainage canals, which merge to form collective canals. Water from drainage canals may be used to irrigate pasture land, while some is collected in buffer dams (Herold *et al.*, 1996). The drainage canals flow towards the Harts River, which is where most drainage water ends up (Ellington, 2003).

The Tshiombo Irrigation Scheme is a smallholder irrigation scheme situated in the Vhembe region of Limpopo Province. It is 847 ha in size, with average plot size of about 1.28 ha (Thagwana, 2010; Van Averbeke, 2012). Water is drawn from the Mulale river and weirs and canals are used to transport the water to the plots. Flood irrigation was used in the area, but has been replaced by floppy irrigation in some parts. Crops which are grown in this area include maize, sweet potato, groundnuts and tea (Fourie, 2014; Thagwana, 2010). Although maize is grown in this area, to the best of our knowledge, no GM maize was planted at the time of sampling. The three sampling points of the study were the Maraxwe Canal, the Mutale River and the Mbahela Dam, though macroinvertebrates were only collected from the river.

The dams and rivers of the Vaalharts Irrigation Scheme are aquatic ecosystems which are exposed to Bt plant material, specifically Bt maize. As discussed more fully in Chapter 2, Bt plant material may end up in the aquatic environment due to the effects of wind and rain (Rosi-Marshall *et al.*, 2007), as well as via the subsurface drains which drain water from the maize fields (Tank *et al.*, 2010). In this chapter, molecular techniques are used in combination with general surveillance techniques to investigate the interaction of Bt maize with adjacent aquatic ecosystems. To this end, primers developed for the detection of MON810 Bt maize (which expresses cry1Ab protein) were used to detect the presence of MON810 transgenes in DNA isolated from aquatic macroinvertebrates collected from the Vaalharts Irrigation Scheme, and thus provide a snapshot of the aquatic invertebrates in this system which are exposed to Bt transgenic maize.

3.2 Materials and Methods

3.2.1 Study Area: Vaalharts Irrigation Scheme

The Vaalharts Irrigation Scheme (see Figure 3.1 below) is situated on the border of the Northern Cape and North West Provinces. Bt maize has been grown in this area since its commercial release in South Africa in 1998 (Kruger *et al*., 2009). Five sampling sites were selected at water bodies across the irrigation scheme. These were: Site A (Spitskop Dam); Site B (Ganspan Dam); Site C (Espagsdrif); Site D (Dam 8) and Site E (North Canal).

3.2.2 Control Site: Tshiombo Irrigation Scheme

The Tshiombo Irrigation Scheme is situated in Limpopo in the Venda region (see Figure 3.2 below). Farming practice in this region differs significantly from that in Vaalharts in that most farms in this region are small subsistence farms. Although maize is grown in this area, to the best of our knowledge, no GM maize was planted at the time of sampling. Even though three locations (a dam, a river and a canal) were selected within the Tshiombo scheme, the scheme as a whole is viewed as one control site. The three sampling points were the Maraxwe Canal; the Mohale River and the Mbahela Dam.

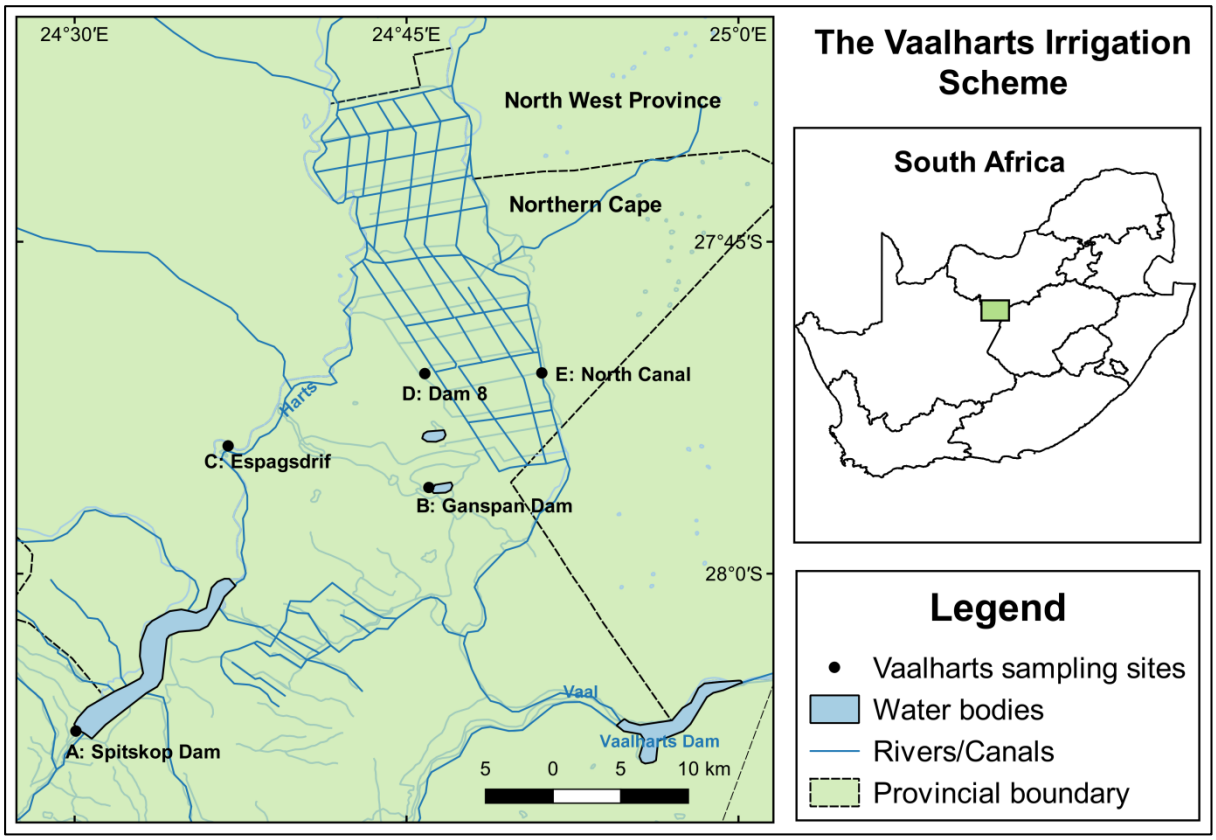


Figure 3.1: A map of the Vaalharts Irrigation Scheme (QGIS version 2.14)

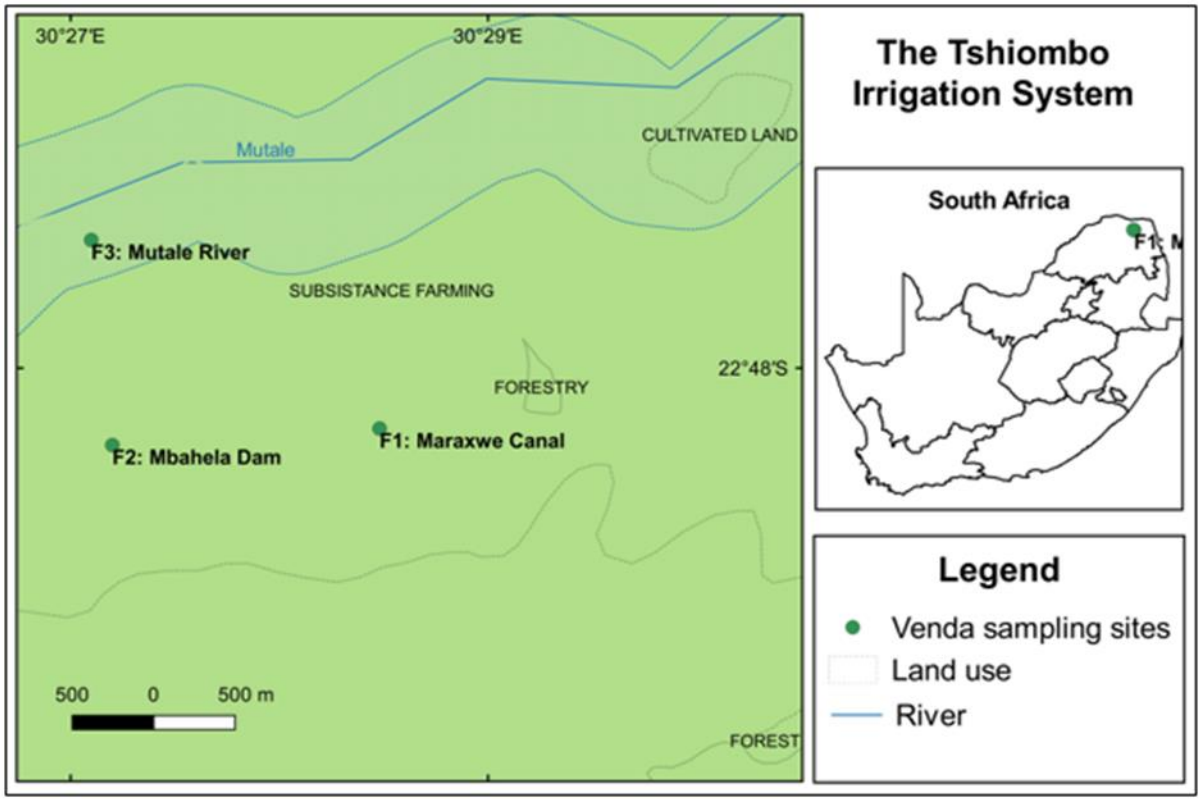


Figure 3.2: A map of the Tshiombo Irrigation Scheme (QGIS version 2.14)

3.2.3 Macroinvertebrate Specimens: collection and identification

Macroinvertebrates were collected in April 2011 according to the procedures documented in the River Health Programme Implementation Manual (DWAF, 2008), using a variety of nets for surface invertebrates as well as those found in the vegetation and gravel/sand/mud biotypes. Macroinvertebrates were collected from the surface, as well as by passing the nets through the eater and disturbing the rocks, gravel and sand underfoot by kicking. These specimens were stored in plastic containers and placed in cooler boxes. They were then stored individually in alcohol until they could be identified, and their DNA extracted. Macroinvertebrate samples were identified by Hilde Kemp (Subject group Zoology; NWU Potchefstroom Campus). Representatives of each morphotype were then photographed before DNA isolation.

3.2.4 DNA Isolation

To isolate DNA from macroinvertebrates, the Zymo Research Micro, Mini and Midi Insect kits (Zymo Research, US) were used. The specimens were weighed, as the weight determined which kit would be used. Very small macroinvertebrates (< 10 mg) of the same type and from the same site were occasionally pooled together in one tube for DNA isolation. Addition of Proteinase K

(Thermo Scientific, US) to samples during the initial lysis step was found to improve yields. Following this step, the protocols were followed according to the manufacturer's instructions.

MON810 maize DNA was required as a positive control. Maize plants were grown at EcoRehab (NWU). Leaves were used for DNA isolation, using the NucleoSpin Plant II kit (Macherey-Nagel, Germany) according to the protocol supplied by the manufacturer.

3.2.5 DNA Amplification

3.2.5.1 COI amplification

"Universal" primers refer to primers which can be used to amplify specific regions of DNA from a broad range of representatives of certain groups of organisms (Folmer *et al.*, 1994). For animals, primers targeting the mitochondrial cytochrome oxidase subunit I gene (COI) are frequently used for DNA barcoding and identification, due to their broad scope (Hebert *et al.*, 2003). In the present study, primers targeting the COI gene were used to demonstrate that the isolated macroinvertebrate DNA was amplifiable; and to provide a template for sequencing at a later stage (see Table 3.1 below).

Table 3.1: Primer sequences used to amplify COI gene of macroinvertebrate DNA isolates

Primer	Sequence	Amplicon size	Reference
LCO1490	5'-GGTCAACAAATCATAAAGATATTGG-3'	710 bp	Folmer *et al.* (1994)
HCO2198	5'-TAAACTTCAGGGTGACCAAAAAATCA-3'		

The PCR reaction mixture contained 12.5 µl double strength PCR master mix [(0.05 U/µl Taq DNA Polymerase in reaction buffer, 0.4 mM of each dNTP; 4 mM $MgCl_2$ (Fermentas Life Sciences, US)]; 10 pmol of each primer (Applied Biosystems, UK); 50 ng DNA template; and PCR-grade nuclease-free water (Fermentas Life Science, US) to fill up the volume to 25µl. The reaction took place in an ICycler thermal cycler (BioRad, UK) under the following conditions: 94°C for 300 s; 35 cycles of 94°C for 30 s; 50°C for 30 s; 72°C for 45 s; and final extension at 72°C for 300 s (Otomo *et al.*, 2009).

3.2.5.2 PCR with MON810 Primers

MON810 primers were used to detect the presence of MON810 transgene sequences in the DNA of the aquatic organisms (see Table 3.2). Three primer set combinations were chosen, using the PCR conditions described in Table 3.3 below.

Table 3.2: The sequences of the MON810 primers used in this study

Primer	Sequence	Reference
CM01 F	5'- CTCTACAAATGCCATCATTGCGATA-3'	Matsuoka *et al.* (2000)
CM02 R	5'-CTTATATAGAGGAAGGGTCTTGCGA-3'	Matsuoka *et al.* (2000)
HSP70 F	5'-GATGCCTTCTCCCTAGTGTTGA-3'	Kuribara *et al.* (2002)
Cry1A(b) R	5'-GGATGCACTCGTTGATGTTTG-3'	Kuribara *et al.* (2002)
1213 F	5'-GGCACGGTGGATTCCCTGGACGAGAT-3'	Hernández *et al.* (2003)
2388 R	5'-GGAGAAGTGGTGGCTGTGGTGGGC-3'	Hernández *et al.* (2003)

The reaction mix for all the MON810 primer sets comprised of 12.5 µl double strength PCR master mix [(0.05 U/µl Taq DNA Polymerase in reaction buffer, 0.4 mM of each dNTP; 4 mM $MgCl_2$ (Fermentas Life Sciences, US)]; 20 pmol of each primer (Fermentas Life Sciences, US); 50 ng DNA template; and PCR-grade nuclease-free water (Fermentas Life Science, US) to fill up the volume to 25 µl. Each primer set had its own cycling conditions which are listed in Table below. For each PCR reaction using MON810 primers, a positive control (isolated MON810 maize DNA), a negative control (DNA from a sample from the control site), and a no-template control were included.

Table 3.3: The three primer sets used for amplification of transgene fragments, with expected amplicon sizes and cycling conditions

Primer Set	Primers used	Cycling conditions	Expected Amplicon Size
1	CM01F CM02R	95°C for 300 s; 40 cycles of 95°C for 30 s; 66°C for 30 s; 72°C for 30 s; and final extension at 72°C for 420 s.	220 bp
5	HSP70F Cry1A(b)R	95°C for 300s; 35 cycles of 95°C for 30 s; 59°C for 30 s; 72°C for 30 s; and final extension at 72°C for 420 s.	113 bp
8	1213F 2388R	95°C for 300 s; 40 cycles of 95°C for 30 s; 68°C for 30 s; 72°C for 60 s; and final extension at 72°C for 420 s	1197 bp

3.2.5.3 Electrophoresis and visualization of PCR products

After amplification, 5 µl of PCR product was mixed with 3 µl of 6 x O'range Loading Dye (Fermentas Life Sciences, US) containing Gel Red, and loaded into a 1.5% (w/v) agarose gel. Electrophoresis proceeded at 80V for 50 minutes in 1 x TAE buffer (20mM acetic acid (Merck, US); 40mM Tris (Sigma Aldrich, US); 1mM EDTA (Merck, US) at pH 8). Amplicons were visualized using the BioRad Gel Doc Imaging system (BioRad, UK).

3.2.6 Sequencing

Sequencing was done in-house. The NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel, Germany) was used to purify the amplicon DNA, followed by sequencing PCR using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, UK). The products of this reaction were purified again using the Zymo Research DNA Sequencing Clean-up Kit™ (Zymo Research, US) according to the manufacturer's instructions, and loaded into the 3130 Applied Biosystems Genetic Analyser, where sequencing electrophoresis took place using a 36 cm capillary array and POP-7™polymer (all Applied Biosystems, UK). Sequence analysis was done by the 3130 Genetic Analyser Data Collection software (Applied Biosystems, UK). The resulting chromatograms were analysed in BioEdit v7.1.3 software.

3.3 Results

3.3.1 Collection and identification of macroinvertebrate specimens

Macroinvertebrate specimens collected in the field were identified at least to family level, genus or species level where possible by Hilde Kemp of the Zoology subject group (NWU), see Table 3.4 below. The specimens were stored in 70% ethanol until they were photographed and their DNA isolated (though in some specimens were too damaged to be recognisably photographed). In some cases when the specimens were very small, they were pooled for DNA isolation. In such cases a single isolate code was assigned to the pool as a whole, bringing the number of macroinvertebrate isolates in Table 3.4 to 68. No macroinvertebrates could be collected from Site E (North Canal). The macroinvertebrates belonged to 18 families, from 9 orders. Of the 21 families identified, 15 were collected from Vaalharts (sites A-D), while 9 were collected from the control site (Venda, site F). Between the study site and control site, only 3 families were held in common: Corixidae, Baetidae and Coenagrionidae. Photographs of collected specimens can be found in the appendix.

Table 3.4: Identification of collected macroinvertebrate specimens, and the number of DNA isolates per family

Order	Family	Genus	Number of DNA isolates	Site
Araneae	Tetragnathidae		2	Vaalharts
Coleoptera	Elmidae		1	Venda
	Gyrinidae		1	Vaalharts
	Hydrophilidae		1	Vaalharts
Decapoda	Atyidae	*Caridina*	9	Vaalharts
Diptera	Ceratopogonidae		1	Vaalharts
	Chironomidae	*Chironminae*	1	Vaalharts
	Culcidae		1	Vaalharts
Ephemeroptera	Baetidae		2	Vaalharts
			1	Venda
Hemiptera	Belastomatidae	*Appasus*	14	Vaalharts
	Corixidae	*Micronecta*	2	Vaalharts
			1	Venda
	Gerridae	*Eurymetra*	4	Venda
	Mesoveliidae		4	Vaalharts
	Notonectidae	*Anipsops*	7	Vaalharts
	Pleidae		1	Vaalharts
Lepidoptera	Pyralidae		1	Venda
Odonata	Aeshnidae	*Anax*	1	Vaalharts
	Coenagrionidae	*Pseudagrion*	3	Vaalharts
			5	Venda
	Gomphidae	*Ceratogomphus*	1	Venda
	Libellulidae		1	Venda
Trichoptera	Hydropsychidae	*Cheumatopsyche*	1	Venda

3.3.2 Amplification with Cytochrome Oxidase (COI) primers

After morphological identification, DNA was isolated from the macroinvertebrates collected at the various sites. PCR with cytochrome oxidase (COI) primers was done to determine whether the DNA was of sufficient quality to act as a template for further PCR reactions (see Figure 3.3 below). DNA isolates which failed to generate amplicons of the expected size at this stage were rejected from PCR with MON810 primers.

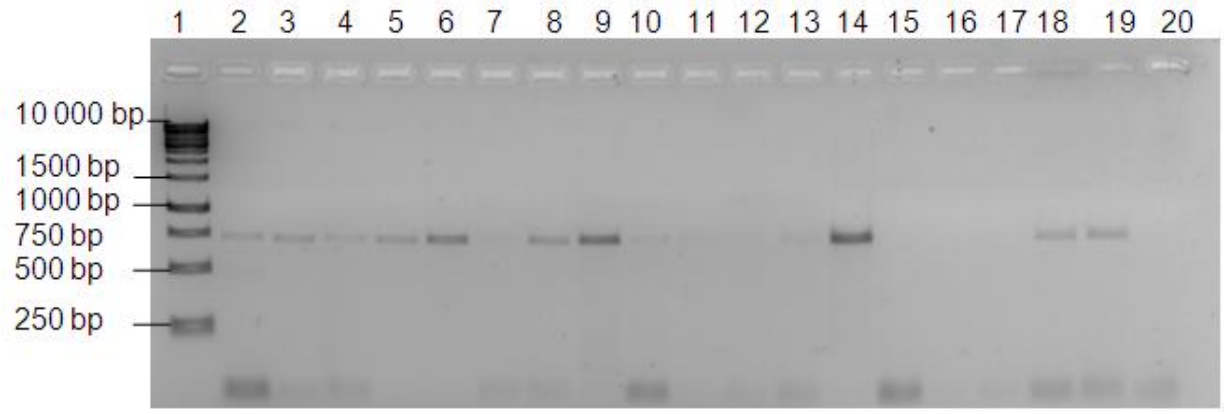


Figure 3.3: **A 1.5% (w/v) agarose gel after electrophoresis of the products of amplification of macroinvertebrate DNA using the COI primer set. The expected 710 bp product can be seen in lanes 2-11, 13-14, and 17-18. Lane 20 contains the no-template control, while lane 1 contains the molecular weight marker (O'GeneRuler™, Fermentas Life Science, US).**

Although in some cases (lanes 7, 10, 11, and 13 of Figure 3.3, for example) the result was weak, the expected 710 bp amplicon was produced for 68 macroinvertebrates (12 isolates did not amplify successfully). PCR with CO1 primers served to demonstrate that the isolated DNA was amplifiable, such that negative results obtained with MON810 primer sets would not be attributed to DNA quality. Although only DNA isolates which generated the expected 710 bp amplicon were used for further PCR with MON810 primers, those which did not were not removed from the list of macroinvertebrates collected and identified morphologically during this study. In total, 68 macroinvertebrate DNA isolates were used, of which 15 were from site A, 7 from site B, 7 from site C, 21 from site D and 18 from site F. No macroinvertebrates were collected from site E.

3.3.3 Amplification with MON810 primers

Figure 3.4 is an example of electrophoresis results following PCR of macroinvertebrate DNA isolates with primer set 1.

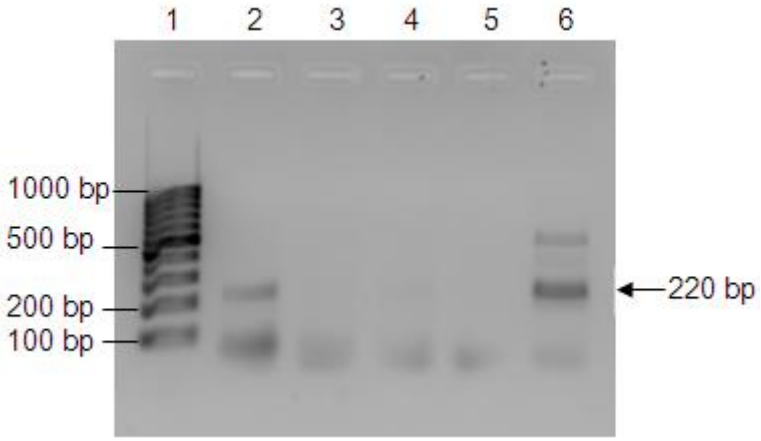


Figure 3.4: **An inverted image of a 1.5% agarose gel after electrophoresis of PCR products following amplification of macroinvertebrate DNA with primer set 1. The 100 bp molecular weight marker (O'GeneRuler™, Fermentas Life Science, US) was loaded in lane 1, the negative control in lane 4, the no-template control in lane 5 and the positive control in lane 6.**

The size of the amplicon in lane 2 of Figure 3.4 corresponds to the expected size of primer set 1 (220 bp). Lanes 4 and 5 contain the negative and no-template controls, respectively, while lane 6 contains the positive control. The isolate in lane 2 of Figure 3.4 is B87-11/04/08mi. Four of macroinvertebrate DNA isolates had amplicons of this size after amplification with set 1 primers. These correspond to isolates A91-11/04/08mi, D71-11/04/08mi, D80-11/04/08mi and B87-11/04/08mi.

Primer set 5 generated more positive results than primer set 1. Figure 3.5 below is an example of the results obtained after PCR of macroinvertebrate DNA with primer set 5. In Figure 3.5, positive results can be seen in lanes 6-7 (B87-11/04/08mi and D71-11/04/08mi), and 10-12 (D73-11/04/08mi, D77-11/04/08mi, and D80-11/04/08mi). Lane 16 contains an amplicon which is larger than the expected 113 bp fragment produced by the positive control (see lane 20). Lane 18 is the negative control, while lane 19 contains the no-template control. Six macroinvertebrate DNA isolates produced amplicons of the expected size (113 bp) after amplification with primer set 5. These were spread among site A (A95-11/04/08mi), site B (B87-11/04/08mi) and site D (D71-11/04/08mi, D80-11/04/08mi, D73-11/04/08mi and D77-11/04/08mi). No positive results were obtained for site C or the control site, and no macroinvertebrates were collected from site E.

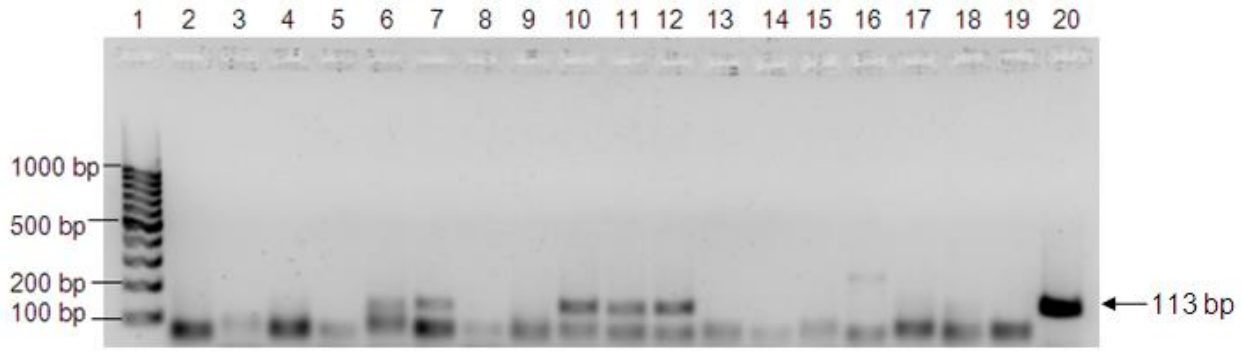


Figure 3.5: **Inverted image showing the PCR results after amplification with primer set 5 following electrophoresis. Lane 20 contains the positive control, while lanes 18 and 19 represent the negative control and no-template control respectively.**

Figure 3.6 below represents an example of the results obtained when macroinvertebrate DNA underwent PCR with primer set 8.

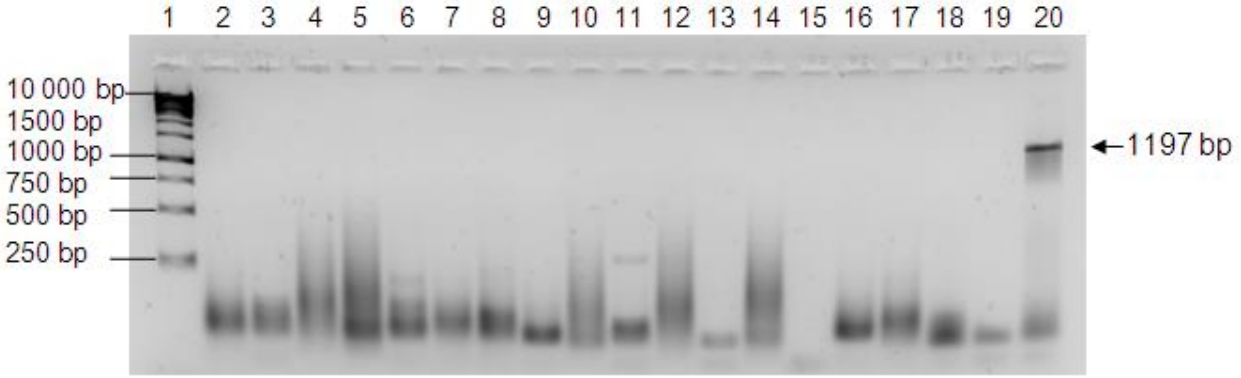


Figure 3.6: **An inverted image of a 1.5% gel containing GelRed-stained PCR products of macroinvertebrate DNA amplified with primer set 8. Lane 1 contains the 1 kb molecular weight marker (O'GeneRuler™, Fermentas Life Science, US), lane 18 the negative control, lane 19 the no-template control and lane 20 the positive control.**

Primer set 8 did not generate any amplicons of the expected size among the macroinvertebrate DNA isolates. The macroinvertebrate DNA often formed smears when amplified with primer set 8, as can be seen in Figure 3.6 above. When bands were produced, they tended to be in the region of 250 bp, such as the one seen in lane 11. Lane 11 contains the only single-band amplicon obtained in this set of results, other than the positive control amplicon in lane 20 (Figure 3.6 above).

The positive results after PCR with the MON810 primers is summarised in Table 3.5 below, which also provides the morphological identification of the organisms for which positive results were obtained.

Table 3.5: Summary of the isolates for which positive results were recorded for the MON810 primer sets.

Isolate code	Primer set with positive results	Morphological identification
A91-11/04/08mi	1	Hemiptera: Mesoveliidae
A95-11/04/08mi	5	Hemiptera: Notonectidae Genus: *Anisops*
B87-11/04/08mi	1, 5	Decapoda: Atyidae Genus: *Caridina*
D71-11/04/08mi	1, 5	Hemiptera: Belastomatidae Genus: *Appasus*
D73-11/04/08mi	5	Odonata: Coenagrionidae Genus: *Pseudogarion*
D77-11/04/08mi	5	Decapoda: Atyidae Genus: *Caridina*
D80-11/04/08mi	1,5	Decapoda: Atyidae Genus: *Caridina*

3.3.4 Sequencing of MON810 amplicons

The PCR products were sequenced to determine their similarity to the MON810 transgene. Successfully sequenced amplicons were aligned with the transgene in BioEdit (v7.1.3). Figure 3.7 below, an example of such an alignment is given for primer set 1. The forward and reverse sequences obtained after amplification with primer set 1 align with the transgene sequence from position 55 (start of forward sequence) until position 276 (end of reverse sequence). The total is 220 bp, the entire expected length of the amplicon produced by primer set 1. Although a few differences are observed, for example at positions 55 – 57 and position 276, overall the sequence of the D71-11/04/08mi amplicons closely matches that of the transgene.

3.3.5 Sequencing and attempted identification of the COI amplicons

An attempt was made to use the sequences of the COI amplicons to identify the collected organisms down to species level. In some cases, the entire length of the amplicon could not be sequenced, meaning that the sequences were too short to make species-level identifications. In other cases, the sequences were long enough, but identification was still unreliable. Further investigation indicated that there is a dearth of South African aquatic macroinvertebrate sequences in public databases, which prevents their identification using DNA sequences – this became the subject of a publication (Venter & Bezuidenhout, 2016).

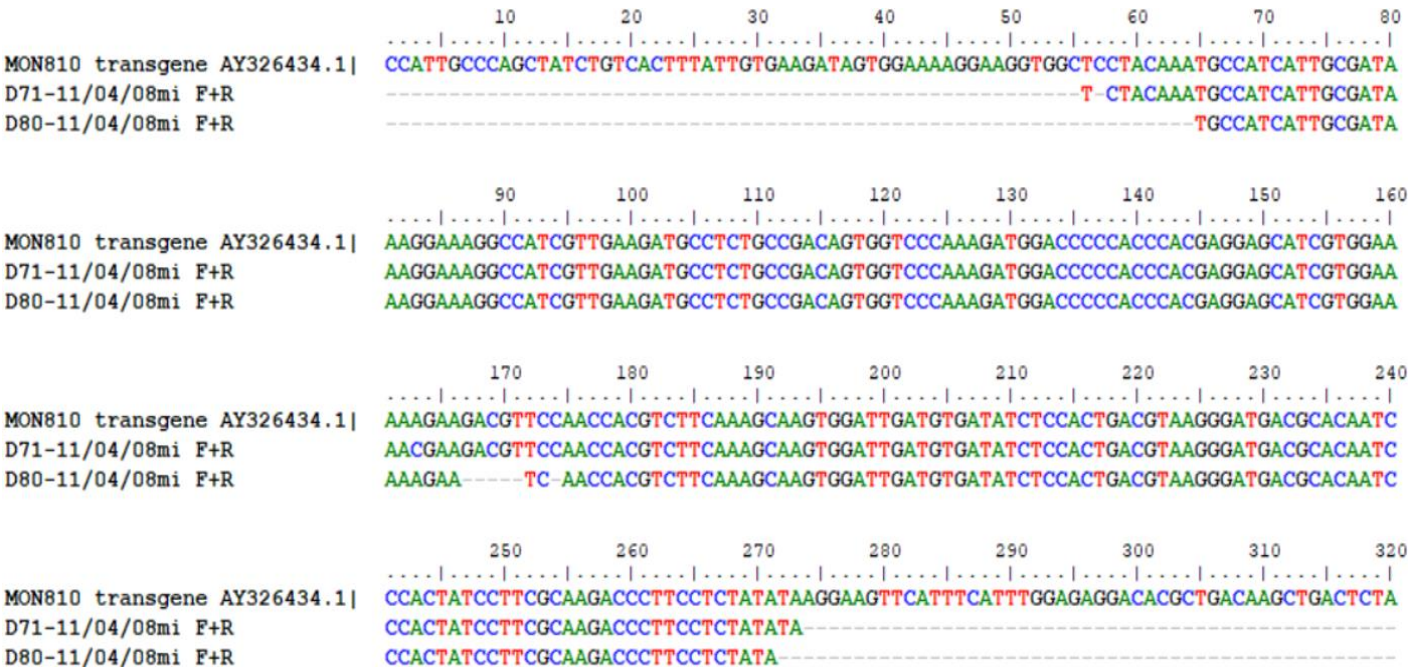


Figure 3.7: Alignment of the sequences obtained from the amplicons produced by PCR of D71-11/04/08mi and D80-11/04/08mi, with that of the MON810 transgene.

3.4 Discussion

3.4.1 Identification of macroinvertebrates

While this was not an exhaustive study into the make-up of the aquatic macroinvertebrate communities of the Vaalharts and Tshiombo irrigation systems, this study does give a snapshot of the organisms which were found in these environments at the time of sampling. The macroinvertebrates collected were identified morphologically, and found to belong to 21 families, from 9 orders (see Table 3.4 and appendix). DNA isolation yields from the macroinvertebrate specimens were generally low. Reasons for low yields may be due to incomplete lysis of the macroinvertebrate tissue. However, PCR with COI primers indicated that the isolated DNA was amplifiable in most cases.

3.4.2 Detection of MON810 sequences

The site with the most positive results was site D, which is unsurprising since site D (Dam 8) is situated between maize fields, and was observed to have received a substantial amount of maize plant debris. Sites A (Spitskop Dam) and B (Ganspan) are further away from the maize fields and represent possible downstream movement and spread of Bt maize plant material throughout the irrigation system. However, no transgene sequences were detected in DNA isolated from

macroinvertebrates collected from site C (Espagsdrif). The flowing water at this site, combined with the fact that it was in flood at the time of sampling, as well as the rather small sampling size may explain why no transgenic constructs were detected among macroinvertebrates from this site. Reasons for this may include a dilution effect (reducing contact with Bt material), Bt material being washed away, as well as other nutrient sources being available.

The primer sets which successfully amplified transgene DNA were those which had short amplicons (113 – 220 bp). Primer set 8, which targeted a larger region (1197 bp) was not successful, although it did successfully amplify the target region in the positive control DNA. These results are consistent with DNA damage and fragmentation which is expected once released into the environment (Epp *et al.*, 2012; Thomsen & Willerslev, 2014). Primer set 1 was developed to detect the presence of transgenic DNA in food (Matsuoka *et al.*, 2000). It was also used by Quist and Chapela (2001) to detect transgenic DNA in native Mexican maize landraces, indicating that gene flow had taken place. This primer set amplifies the CaMV 35S promoter region, which is used in a number of GM crops including several events of Bt maize, cotton and soybean (Kuribara *et al.*, 2002; Matsuoka *et al.*, 2000; Randhawa *et al.*, 2010) and may also be present in CaMV-infected members of the Brassica family. Positive results for this primer set have been used to indicate that transgenic DNA is present, though it does not specify which transgenic event is present. Primer set 5 is the same primer pair which Douville *et al.* (2009) used to detect the presence of the MON810 transgene in freshwater mussel tissue and aquatic bacteria. This primer set is specific to MON810 (Kuribara *et al.*, 2002). Primer set 8, which targeted a larger region (1197 bp) was not successful, though it did successfully amplify the target region in the positive control DNA (a portion of the MON 810 transgene). The inability to amplify larger regions of DNA is consistent with DNA damage and fragmentation which is expected once released into the environment (Epp *et al.*, 2012; Thomsen & Willerslev, 2014).

Given the nature of the DNA isolation procedure used, sources of positive signals detected in the DNA isolated from macroinvertebrates may include consumed Bt maize plant material, ingested prey which consumed Bt plant material, the presence of Bt DNA in the water (although the lack of positive signals among many of the specimens makes this unlikely), Bt material taken in by filter feeders, or the presence of microbes transformed with fragments of transgene constructs (as put forward by Douville *et al.* (2009)). Surveys of terrestrial arthropods collected from fields of Bt crops and tested for Bt proteins using ELISA have indicated that Bt protein can be detected among organisms of higher trophic levels, suggesting transfer from prey to predator (Harwood *et al.*, 2005; Yu *et al.*, 2014b). Feeding studies conducted on the predator *Poecilus cupreus* (Coloptera: Carabidae) fed *Spodoptera littoralis* (Lepidoptera: Noctuidae) larvae reared on Bt maize indicated that Cry1Ab protein was transferred from prey to predator, albeit with a tenfold

reduction in concentration (Alvarez-Alfageme *et al.*, 2009). This indicates that transgenic DNA may survive the transfer across trophic levels. There is a need for similar studies examining multi-trophic transfer of transgenic DNA and proteins within aquatic organisms and ecosystems.

Relatively few herbivorous macroinvertebrates were collected in this study. All of the macroinvertebrates which tested positive after amplification with MON810 primers – with the exception of the Atyidae – belonged to families with predatory feeding habits (Picker *et al*., 2004). Atyidae (genus *Caridina*, freshwater shrimps) feed on algae, detritus of plant and animal origin, and bacteria (Hart *et al.*, 2001). The Atyidae family contained the most positive results, and also contained the only two individual isolates which gave positive results for both primer set 1 and primer set 5 (B87-11/04/08mi and D80-11/04/08mi). The diverse feeding strategy of the freshwater shrimps provides a number of possible exposure pathways, including direct consumption of Bt plant material, consumption of detritus of organisms which had consumed Bt material, ingestion of bacteria transformed with transgenic constructs, and feeding on algae to which Bt proteins had adsorbed (Douville *et al.*, 2009; Kratz *et al.*, 2010; Wang *et al.*, 2014b). If only Bt protein was ingested (if adsorbed to algae, for instance) a positive result might not have produced with PCR. The diets of Mesoveliidae, Coenagrionidae and Notonectidae include aquatic insects and their larvae, crustaceans, small fish, and insects trapped on the water surface, and Belastomatidae will essentially feed on any organism they can subdue with their forelegs (Picker *et al*., 2004). The predatory feeding strategies of most of the positive macroinvertebrates suggests that the source of the positive results is not direct consumption of Bt maize plant material, but transfer via prey. More comprehensive sampling efforts are needed to better understand multi-trophic transfer of transgenic DNA and proteins within aquatic ecosystems. An alternative explanation is that the macroinvertebrates may have harboured bacteria transformed with the transgenic DNA, as suggested by Douville *et al*. (2009).

Sequences associated with the Bt MON810 transgene using DNA-based techniques were detected. Among the macroinvertebrates collected in Vaalharts were aquatic representatives of the orders Diptera and Coleoptera (see Table 3.4), orders which are sensitive to some Bt toxins (Bravo *et al.*, 2007). Although no transgene DNA was detected in the individuals collected in this study, the sampling size was small and only one sampling run was done. Follow-up studies detecting Bt proteins and plant materials are needed. Surveys such as those used by Yu *et al.* (2014b) and Harwood *et al.* (2005) to detect Bt proteins in terrestrial invertebrates, are recommended. Combining such collection efforts with DNA barcoding and metabarcoding may make future aquatic monitoring endeavours simpler, by building up the sequence databases needed to identify organisms by their DNA. Though long-term monitoring of aquatic ecosystems generally falls within the province of the National Aquatic Ecosystem Health Monitoring

Programme (which currently employs a suite of other methods), the approach has some overlap with the general surveillance aspect of SANBI's monitoring policy and may be of use to both.

3.4.3 Recommendations

From this study follows two recommendations are made for further study of the interaction of genetically modified crops and aquatic ecosystems:

- Use of *Caridina* as a candidate organism for further study
- Promote the use of DNA techniques in aquatic ecosystem monitoring

3.4.3.1 *Caridina* as a candidate for further study

When selecting surrogate species to be used for risk assessment of GM crops, Romeis *et al.* (2013) list relevance, potential sensitivity, and availability and reliability as the three criteria which the selection should be based upon. From the results of this study, *Caridina* emerged as a contender for further study based on these criteria. In terms of relevance, it has been demonstrated that atyid shrimp were exposed to Bt transgenic material at levels which could be detected with standard PCR. Atyid shrimp were found at three sampling sites throughout the Vaalharts Irrigation Scheme, and are widespread throughout most of northern and eastern South Africa (Hart *et al.*, 2001), where Bt crops are also abundant. These freshwater shrimps accounted for almost half the positive results for the Bt maize transgene in the present study. The detrivorous feeding strategy of these organisms, which includes feeding on plant detritus as well as bodies of dead macroinvertebrates etc. (Hart, 2001; Hart *et al.*, 2001), makes it seem likely that exposure occurred through their diet, though it is possible that the shrimp could have come into contact with the transgene in their surrounding water or due to bacteria transformed with the transgene.

In this study, exposure to Bt crop residues was inferred from the presence of transgene DNA. Sensitivity to Bt proteins was not investigated. However, potential sensitivity to Bt proteins may be suggested by results of experiments with other crustaceans and cry1Ab toxin. Rusty crayfish (*Oronectes rusticus* (Decapoda: Cambaridae)), also a decapod, experienced 31% reduction in survival when fed cry1Ab-expressing Bt maize, compared to an isogenic control variety (Linn & Moore, 2014). Experiments on more distantly related crustaceans (*Daphnia magna* (Cladocera: Daphniidae) also suggest potential non-target effects, though a mechanism for this has not yet been proposed (Bøhn *et al.*, 2008; Bøhn *et al.*, 2016; Bøhn *et al.*, 2010).

Availability and reliability of *Caridina nilotica* as a test organism has been demonstrated by its use in several studies (Mensah *et al.*, 2012a-c; Mensah *et al.*, 2011; Mensah *et al.*, 2013). The fact that these studies document the effects of glyphosate on this organism mean that the foundation is laid for possible combinatorial studies of Bt crops and glyphosate, which is an extremely understudied aspect of the interaction of genetically modified crops and their environments (see Chapter 2). Glyphosate is the active ingredient in Roundup, the herbicide which most herbicide tolerant (HT) crops are genetically modified to be tolerant to, and are designed to be used in conjunction with. Stacked crops bearing both Bt and HT traits have outstripped single gene Bt crops in popularity (see Figure 1.1). Glyphosate is a systemic herbicide, and accumulates in metabolic sinks of treated plants, including seeds (Duke *et al.*, 2003), nodules (Reddy & Zablotowicz, 2003) and roots (from which it may also be released as part of the root exudate (Kremer *et al.*, 2005)). The conversion of glyphosate to its metabolite aminomethylphosphonic acid (AMPA) in plants sprayed with glyphosate-based herbicides can result in the presence of AMPA alongside glyphosate residues in the plant tissues (Bøhn *et al.*, 2014; Duke *et al.*, 2003; Komoßa *et al.*, 1992; Reddy *et al.*, 2004). Aside from drainage and runoff of herbicide residues, the stacked Bt/HT crop material which enters aquatic ecosystems represents a potential entry route for glyphosate, AMPA and Bt toxins simultaneously which has not been investigated.

3.4.3.2 The potential of DNA identification techniques in aquatic ecosystem monitoring

The usefulness of molecular methods in monitoring scenarios (detecting transgenic DNA) has been shown in this study. However, when attempts were made to use macroinvertebrate COI sequences (DNA barcodes) to provide further resolution to morphological identifications of macroinvertebrates, only weak matches could be found with sequences in the GenBank database. These matches often did not agree with the morphological identifications, or corresponded to species found in Northern Hemisphere countries or Australia. In addition, the BOLD Systems Databases' records for South Africa indicates a noticeable dearth of aquatic macroinvertebrates. For instance, at the time of writing, there were zero entries referring to families such as Baetidae and Ephemeridae, and only 43 records for the whole order Ephemeroptera. Similarly, Plecoptera only had 13 entries, while Odonata had 66 and Trichoptera 138. When compared to the 3621 records for Coleoptera and 3150 for Hemiptera (BOLD, 2014) it is clear that there is considerable room for improvement for DNA barcoding of aquatic organisms.

The possibilities of using DNA in general surveillance techniques are not limited to barcoding, which is a very useful identification aide but is still reliant on relatively invasive and destructive

sampling techniques. The approach of collecting specimens and isolating their DNA individually in order to identify them (as in the case of DNA barcoding), has been shifting towards making use of the DNA which organisms naturally shed into their environments (eDNA). Being able to work with eDNA precipitated from samples from aquatic ecosystems opens up a range of monitoring and research opportunities (Bohmann *et al.*, 2014; Mächler *et al.*, 2014). DNA metabarcoding, for example, is an approach which uses bulk DNA collections, such as faeces (De Barba *et al.*, 2014), sediment meiofaunal communities (Lallias *et al.*, 2015), and eDNA (Thomsen & Willerslev, 2014), coupled with next generation sequencing to obtain an overview of the organisms which are present in an environment as a whole. Instead of targeting one or a few species, this approach aims to give a more holistic view of ecosystem composition (Cristescu, 2014). The reads which result from such high throughput sequencing are clustered into operational taxonomic units (OTUs). Depending on the genetic loci used, OTUs can be used to match to sequences in databases such as BOLD (if CO1 was used) or GenBank, in order to identify the organisms whose DNA was in the sample (Cristescu, 2014; Ji *et al.*, 2013; Leray & Knowlton, 2015). In other cases, in which primer sets other than COI are used or when no reference sequences are available for the OTUs produced, taxonomic assignation is not possible. However, that does not mean the data are not useful. Molecular taxonomic units (MOTUs) refer to representative sequence clusters which have been grouped together using particular algorithms. MOTU data can be utilised *in lieu* of 'true' species data, by comparing MOTU profiles of different environments or time period (Jones *et al.*, 2011). Although the metabarcoding approach still is being refined, it potentially allows monitoring of community-level responses to change, including responses to remediation strategies and climate change (Epp *et al.*, 2012; Ji *et al.*, 2013).

The fact that DNA has a relatively short turnover time in aquatic systems means that aqueous eDNA is likely to represent a 'real-time' view of species present within a relatively small window of time (Turner *et al.*, 2014). Strickler *et al.* (2015) investigated the effects of temperature, pH and UVB radiation on eDNA in water and found that it degraded faster in warmer water, with a neutral pH and a moderate UVB level. Because these conditions are also amenable to microbial growth, the authors speculated that eDNA breakdown was at least partially facilitated by microbial action. During investigations into DNA persistence in both laboratory and field conditions (ponds), species could be detected using eDNA for 25 days and 21 days after removal of the organisms, respectively (Dejean *et al.*, 2011). However, the fact that DNA may be concentrated and survive much longer in sediments may be a complicating factor (Turner *et al.*, 2014).

DNA dispersal in flowing streams and rivers is also a concern, as it may give false positive results downstream where the organism in question is not found. DNA dispersal was investigated by Laramie *et al.* (2015) who traced eDNA of Chinook salmon (*Oncorhynchus tshawytscha*), and

Deiner and Altermatt (2014) who studied eDNA of a daphnid (*Daphnia longispina*) and swollen river mussel (*Unio tumidus*). Both studies found that eDNA signals tend to decrease as distance from the source increases (Deiner & Altermatt, 2014; Laramie *et al.*, 2015). In Deiner & Altermatt's study, DNA from lake-dwelling invertebrates was detected 12 km downstream from the lake inhabited by the target organisms. The authors suggest that when using eDNA to estimate biodiversity in such ecosystems, sample sites should be 5–10 km apart, and follow the stream hierarchy (Deiner & Altermatt, 2014).

Primers and markers have been developed for eDNA, which are often species-specific (in contrast to the broad range of COI), and which have been applied to the detection of indicator species as well as rare, invasive or pathogenic species. Mächler *et al.* (2014) demonstrated the potential of using eDNA and specific primers to detect macroinvertebrate species in both river and lake systems. Using a standard PCR-based method, they were able to detect both indicator and non-native species using their own primer design. Specially designed primers and probes used in conjunction with qPCR were successfully used to survey the population of European weather loach in Denmark (Sigsgaard *et al*., 2014). During this study, this near-extinct fish was detected at sites where its presence had been observed recently, as well as one location where it had not been observed since 1995. In addition to successful detection, the authors report that this approach is less costly, both economically and in terms of effort (person-hours).

However, to harness the usefulness of these techniques, an effort must be made to collect the necessary data to identify DNA sequences. We advocate for the establishment of regional collections which link identified aquatic species with their DNA sequences, which can be used to develop primer sets and standard methods for the use of eDNA in biomonitoring. Expertise in bioinformatics, particularly working with high-throughput sequencing data, will also be required (Venter & Bezuidenhout, 2016).

3.4.4 Conclusion

In this chapter, the aims were to gain understanding of which organisms were exposed to the Bt maize, and to identify possible candidate organisms for further study. Though not exhaustive, a snapshot of the aquatic macroinvertebrates present in aquatic ecosystems in close association with GM crops has been compiled. Using PCR-based techniques, transgene sequences were detected in a number of these organisms, which has led to the recommendation that atyid shrimps be put forward as candidates for further testing for GM crops and their associated herbicides. The advantages of using DNA-based techniques, such as DNA barcoding and metabarcoding, in aquatic ecosystem monitoring have been highlighted. The molecular techniques used in this

chapter may be of use regarding monitoring of GM crops, and in understanding the interaction between these crops and aquatic ecosystems.

CHAPTER 4

DETECTING TRANSGENE FRAGMENTS IN AQUATIC MICROORGANISMS

Abstract

Horizontal gene transfer (HGT) of transgenic DNA to environmental microbiota has been a concern connected to GM crops since their development. This concern about HGT was particularly connected to the antibiotic resistance marker genes borne by many GM crops, and the potential that soil bacteria might be able to take these up and become resistant to the antibiotics in question. In this chapter, we have investigated the potential of HGT of transgene DNA to microorganisms sampled from aquatic ecosystems adjacent to fields growing transgenic Bt maize. Instead of focusing on antibiotic resistance marker genes, we have aimed to detect the transgene of Bt maize MON810. Water samples were collected from the Vaalharts Irrigation System, where MON810 was widely grown, as well as from the Tshiombo Irrigation system where Bt maize was not grown. After cultivation on various media, isolates were selected for further study and DNA was extracted from each. A PCR-based approach using a suite of primers targeting the MON810 transgene was used to detect transgene sequences in the isolates' DNA. Positive results were detected in 56 bacteria, and 20 yeast and fungi isolates. Sequencing allowed identification of all but 2 of these microorganisms. In order to confirm whether the transgene sequenced were integrated into the genomes of these isolates, ten bacterial isolates were selected for whole genome sequencing.

4.1 Introduction

Agriculture and aquatic ecosystems are frequently found in close association with each other. There are estimated to be 500 000 small dams in South Africa, most of which are farm dams (Mantel *et al.* 2010 and references therein). The number of studies which have focused on the effects of Bt crops on aquatic organisms is small, a subject which is discussed in Chapter 2 and 3. Studies which consider the interaction of transgenes in aquatic settings are in even shorter supply than those which consider the effect of Bt proteins. In this chapter, the presence and longevity of transgenic DNA in aquatic ecosystems will be discussed, as well as the mechanisms of uptake and integration which may lead to the maintenance and proliferation of transgenic constructs in communities of aquatic organisms.

4.1.1 Entry and fate of DNA/transgenes in aquatic ecosystems

The routes by which Bt plant material (and transgenes) may enter aquatic ecosystems are discussed more fully in Chapter 2. Briefly, these include:

- Deposition of plant material due to wind, rain etc. (Rosi-Marshall *et al.*, 2007)
- Runoff water and sediment from the field (Strain & Lydy, 2015)
- Drainage water (Tank *et al.*, 2010)

Douville *et al.* (2007) investigated the presence and persistence of the *cry1Ab* transgene in aquatic environments, and found that transgenic DNA constructs from Bt maize could be detected several kilometres downstream from the fields where the maize was grown. This indicates that rivers and streams could be implicated in spreading transgenic DNA away from the immediate surroundings of crop fields. Factors such as UV-B radiation, temperature, pH, and the activity of microbes damage DNA and affect the amount of time it persists in aquatic environments (Barnes *et al.*, 2014; Douville *et al.*, 2007; Strickler *et al.*, 2015). Transgenic constructs were detectable after a 21-day incubation in surface water, and after 40 days in sediment (Douville *et al.*, 2007). These results are consistent with those of Dejean *et al.* (2011), who reported the persistence of environmental DNA in surface water to be 21-25 days, though another study by Strickler *et al.* (2015) has reported it to be over 58 days. Longer persistence of DNA in sediment compared to surface water has also been reported by Turner *et al.* (2014). Adsorption to minerals in sediment and soil particles protects DNA from degradation by microbes and extracellular nucleases (Nielsen, Calamai *et al.*, 2006; Nielsen, Johnsen *et al.*, 2007).

Apart from degradation and adsorption, another potential fate of DNA in such environments is uptake by organisms such as bacteria. Douville *et al.* (2009) were able to isolate *cry1Ab* transgenic DNA from heterotrophic bacteria in two out of six samples taken from surface water within the Bt maize-growing study area, as well as from tissues of freshwater mussels (*Elliptio complanata*). Mussels from a pristine site contained no transgenic DNA, while mussels transferred from the pristine site to the agricultural stream became contaminated with *cry1Ab* transgenic DNA. The authors hypothesized that the bacteria had taken up the transgenic DNA constructs, and were responsible for their sustained presence in the mussel tissue and the aquatic ecosystem overall.

The process by which bacteria take up exogenous (free) DNA from their surroundings and incorporate it into their own genomes is known as natural transformation, and is one of the mechanisms by which horizontal gene transfer (HGT) may take place (Chen & Dubnau, 2004). Any gene transfer event outside of parent-to-offspring transfer is a form of horizontal gene transfer

(Keese, 2008). In addition to transformation, the other main mechanisms of HGT among prokaryotes include conjugation and transduction (Davison, 1999). Conjugation refers to specialised processes which result in the direct transfer of DNA between donor and recipient cells by means of a conjugation apparatus (Grohmann *et al.*, 2003)). Although conjugation occurs primarily between bacteria, transfer from bacteria to yeast (Heinemann & Sprague, 1989), plants (Zupan & Zambryski, 1995), and filamentous fungi (De Groot *et al.*, 1998) is not unheard of. Transduction describes the process by which DNA is transferred between prokaryotes due to the action of bacteriophages (Paul, 1999). Transformation is considered to be the most likely route by which transgenic DNA might be taken up by environmental organisms such as bacteria (Bertolla & Simonet, 1999). Transformed bacteria may then have the ability to sustain the presence of the transgene in the environment, and possibly transfer it to other organisms (de Vries & Wackernagel, 2005; Douville *et al.*, 2009).

4.1.2 Natural transformation and integration of exogenous DNA

For natural transformation to occur, a cell has to become competent. Competency is a transient, highly regulated physiological condition involving between 20 and 50 proteins, during which bacteria are able to take up DNA from their environments (Thomas & Nielsen, 2005). Approximately ninety bacterial species are currently known to be naturally transformable under specific conditions (de Vries & Wackernagel, 2005), though there may be many more whose particular competency-inducing conditions are not yet known (Chen & Dubnau, 2004; Mell & Redfield, 2014).

Once a cell is in a state of competency, transformation proceeds according to the following steps. Firstly, the exogenous double-stranded DNA binds to receptors on the surface of the cell. Some proteobacteria (such as species of the *Neisseriaceae* and *Pasteurellaceae*) are only capable of taking up DNA which contains specific DNA uptake sequences that match with sequences scattered throughout their genomes (Ambur *et al.*, 2007; Elkins *et al.*, 1991; Mell & Redfield, 2014). Other species are less particular, and may potentially take up virtually any piece of DNA from their environment (Lorenz & Wackernagel, 1994).

After binding to the cell surface, the DNA may be transported into the cell. For this to happen in Gram positive bacteria, the DNA must first be nicked (so as to introduce a free end) and made linear for the DNA to be transported into the cytoplasm. In Gram negative bacteria, the DNA is transported across the outer membrane into the periplasmic space via protein channels such as secretins (Chen & Dubnau, 2004). Proteins which are related to those involved in type IV pilus (T4P) and type II secretion system (T2SS) assemblage form part of a DNA translocation complex

which assists in transporting the exogenous DNA into the cell. In both Gram positive and Gram negative organisms, only one of the DNA strands is taken up into the cell. The strand which is not taken up is degraded (Chen *et al.*, 2005).

DNA which has been taken up by a cell may be broken down to be used as nucleotides or as a source of carbon, nitrogen or phosphate. Alternatively, the donor DNA may be integrated into the genome, which completes the transformation process (Dell'Anno & Danovaro, 2005; Mell & Redfield, 2014).

Homologous recombination is the most common way by which donor DNA can be integrated into the recipient cell's genome. Insert sizes of 1 – 10 kb are common (Kung *et al.*, 2013), though the size of integrated DNA fragments can be as large as 90 kb (Coupat *et al.*, 2008), and it is not uncommon for more than one DNA molecule to be transformed when cells are in a state of competency (Mell *et al.*, 2011). The recombination insertion process is mediated by a suite of proteins, including the virtually ubiquitous RecA (or a homolog thereof), which is also involved in housekeeping DNA repair efforts (Claverys *et al.*, 2009). RecA is responsible for finding the areas of homology in the cell's genome, as well as facilitating strand-invasion and strand-exchange which leads to the incorporation of the donor DNA into the genome where a strand break has occurred (Kowalczykowski, 2000; Rocha *et al.*, 2005). As the name suggests, homologous recombination relies upon sequence homology between the incoming DNA and the recipient DNA for integration to be successful. Regions of non-homologous sequence can be sandwiched between flanks of homologous regions, however, leading to the insertion of 'foreign' non-homologous DNA sequences. Both the size of the flanking areas of homology and the size of the non-homologous sequence influence recombination efficiency. Kung *et al.* (2013) found that recombination efficiency rose exponentially as the size of the homologous flanking regions increased from 96 bp to 1 kb (but plateaued after that). Efficiency decreased, however, with increasing size of the non-homologous region. The maximum non-homologous insert region was 6 kb, with 1 kb stretches of homologous sequence on each flank (Kung *et al.*, 2013).

When there is insufficient homology between donor and recipient sequences, integration may still occur via illegitimate recombination (IR). The likely success of IR depends largely on how 'illegitimate' the potential insert sequence is: homology-facilitated illegitimate recombination (HFIR) can increase the frequency of successful integration of foreign DNA 10^5-fold compared to IR with no homology, for example. HFIR occurs when a short region of homology located on one of the insert's flanks recombines with the recipient DNA, and is then able to act as an 'anchor' which allows the rest of the exogenous DNA sequence to become integrated (de Vries & Wackernagel, 2002). Two modes of double IR, which involves two instances of IR on a single

donor DNA fragment, have been demonstrated in *Acinetobacter baylyi*. One requires no homology, and is estimated to occur 10^{10}-fold less frequently than homologous recombination. The other, homology-facilitated double IR, involved small homologous regions, flanked on either side by non-homologous sequences, which facilitated integration by a RecA-dependent mechanism (Hülter & Wackernagel, 2008). Micro-homology, meanwhile, requires only very small regions (approximately 9 base pairs) of weak homology between the exogenous DNA and recipient genome during double strand break repair for the foreign DNA to be inserted (Kohli *et al.*, 1999).

The above mechanisms typically operate with inserts which are several kilobases in length. Environmental DNA, however, seldom persists as such large fragments for long: damage and degradation rapidly reduce eDNA to short fragments, often less than 100 bp in length. Fragments this length are too short to contain whole genes or encode lengthy sequences for novel proteins, and have long been considered to be too short to be of much use in terms of HGT (Nielsen *et al.*, 2007). This view has been challenged by the work of several groups (Chan *et al.*, 2009; Croucher *et al.*, 2012; Croucher *et al.*, 2016; Domingues *et al.*, 2012b; Fani *et al.*, 2014; Harms *et al.*, 2016; Maddamsetti & Lenski, 2018; Overballe-Petersen *et al.*, 2013). Chan *et al.* (2009), found evidence that gene fragments had been inserted within other genes, and Overballe-Petersen *et al.* (2013) showed that very short (as few as 20 bp), damaged, even ancient, DNA fragments can undergo transformation. Insertion of these short fragments is RecA independent and instead relies upon DNA replication; incoming fragments are incorporated during the discontinuous synthesis of the lagging strand (Overballe-Petersen *et al.*, 2013). Though it is unknown how prevalent this process is under natural conditions, RecA independent oligonucleotide recombination has been adopted as a tool of “recombineering” (genetic engineering of bacteria), also known as oligonucleotide mediated mutagenesis. This tool relies upon sequence homology between target regions and single strand oligonucleotides for recombination to occur. This allows variations in the form of polymorphisms, insertions, deletions, and mutations to be introduced into target genomes and bring about a desired phenotype. The mechanism responsible is believed to be conserved in both prokaryotes and eukaryotes, making it a powerful tool in genetic engineering (Bryan & Swanson, 2011; Swingle *et al.*, 2010; Swingle, 2013; van Pijkeren & Britton, 2012).

While recombineering requires substantial sequence homology to function optimally, Harms *et al.* (2016) have recently described integration of short heterologous DNA fragments by means of double illegitimate recombination. They have dubbed this process Short-Patch Double Illegitimate Recombination (SPDIR). Regions of microhomology as short as 12 bp were sufficient for integration to occur, even allowing for gaps and mismatches. DNA sources for SPDIR are

predominantly intragenomic, though HGT can also contribute. The result of SPDIR events are clusters of polymorphisms, which ranged from 3-77 bp in this study.

4.1.3 HGT in Fungi

HGT is acknowledged to be an important factor driving prokaryote evolution, but its role in eukaryote evolution is less clear cut (Keeling & Palmer, 2008; Marcet-Houben & Gabaldon, 2010). Although cases of HGT amongst eukaryotic organisms have been documented, they appear to have occurred much less frequently than in prokaryotes (Richards *et al.*, 2011). This has prompted some scientists to speculate on the relative importance of HGT as a mechanism of eukaryote evolution (Choi & Kim, 2007). Others argue that HGT does occur in most lineages to a certain extent, and that as more genomes are sequenced, more evidence of HGT among eukaryotes may be found and its role in eukaryote evolution elucidated (Fitzpatrick, 2012; Loreto *et al.*, 2008; Marcet-Houben & Gabaldon, 2010). Mechanisms of HGT in eukaryotes include endosymbiotic gene transfer from plastids (Nikoh *et al.*, 2008) conjugation (Heinemann & Sprague, 1989), viral transfer (Liu *et al.*, 2010), phagotrophy (Keeling & Palmer, 2008), transposon transfer, and anastomosis (Richards *et al.*, 2011).

There are several factors which could prevent or limit HGT in eukaryotes. These include the presence of a nuclear membrane; packaging of DNA in chromatin; separation of somatic and germline cells; discordance between the promoter of the in-coming gene and the recipient transcription system; and incompatibility of the transferred DNA with the host intron-splicing mechanisms (Richards *et al*., 2011). Additionally, incompatible codon usage patterns (Fitzpatrick *et al*., 2008); limited exposure to foreign DNA for uptake and lack of selective pressure to maintain the presence of the transferred DNA (Andersson, 2005) can also play a role.

Despite the barriers to HGT, there are numerous cases where they have been overcome, especially in fungal genomes. Richards *et al.* (2011) re-evaluated putative HGT events relating to HGT into fungal genomes and were able to confirm 323 cases. Another study of 60 fungal genomes found evidence of 713 genes acquired from prokaryotes through HGT, distributed among 88% of the genomes analysed (Marcet-Houben & Gabaldon, 2010). The acquired genes were mainly of bacterial origin, with bacterial racemase, catalase and arsenate reductase genes being among those transferred. Heinemann and Sprague (1989) demonstrated that plasmids could be transferred from *Escherichia coli* (bacteria) to *Saccharomyces cerevisiae* when conjugative plasmids were present in the bacteria. *S. cerevisiae* is also capable of receiving genetic material via *Agrobacterium tumefaciens*-mediated transformation (Bundock *et al.*, 1995). Several species of filamentous fungi, including *Aspergillus awamori*, *Aspergillus niger*, *Agaricus*

bisporus, *Fusarium venenatum*, *Trichoderma reesei* and *Neurospora crassa*, are likewise able to undergo *A. tumefaciens*-mediated transformation (De Groot *et al.*, 1998). Transformation in fungi is mostly a laboratory-bound affair requiring measures such as protoplast formation, or physical methods such as electroporation, biolistics, vacuum infiltration, or shock waves in order to transport DNA into the cells (Rivera *et al.*, 2014).

4.1.4 Detection of HGT from transgenic crops

HGT from transgenic crops is recognised as a potential risk associated with their release into the environment. While some authors doubt whether HGT from transgenic crops will occur at a frequency which is cause for concern (Mendelsohn *et al.*, 2003), others regard HGT as inevitable (Heinemann & Traavik, 2004). Previous studies which have investigated HGT from transgenic plants to surrounding organisms have mostly centred round the antibiotic resistance genes which served as marker genes during the development of the crop (de Vries *et al.*, 2003; Gebhard & Smalla, 1998; Mohr & Tebbe, 2007). This is partly due to the need to investigate the risk of spreading antibiotic resistance by means of HGT of these marker genes to bacteria (Hug, 2007; Keese, 2008); and partly because it is thought that sequence similarity between the antibiotic resistance genes and the genomic sequences of certain bacteria may increase the likelihood of HGT (de Vries *et al.*, 2003; Monier *et al.*, 2007). One such marker gene is *nptII* (neomycin phosphotransferase type II), which is considered to be relatively safe in terms of potential spread of antibiotic resistance since resistance to the antibiotics against which *nptII* works (kanamycin and neomycin) is already prevalent among bacteria in the environment (EFSA, 2004).

NptII was the focus of a study which tracked the spread of transgenic crop DNA in the environment (de Vries *et al.*, 2003). During this study, the homology between the *nptII* marker gene from transgenic crops and a truncated version of the gene in an *Acinetobacter* sp. strain was exploited during transformation of the bacterial strain. This method was successful in detecting the presence of DNA from transgenic potato, although other crops tested failed to produce any transformants. The study demonstrated that transgenic crops spread biologically active DNA into their environments, and that this DNA is capable of transforming bacteria – particularly if the bacteria in question contain homologous sequences to facilitate homologous recombination (de Vries *et al.*, 2003).

Evidence of HGT in the field instead of under laboratory conditions is more difficult to detect. Gebhard and Smalla (1998) detected transgenic construct-specific sequences in the bacterial fraction of soil samples taken from the soil surrounding transgenic sugar beets. They were unable to confirm whether HGT had taken place due to the possibility that free DNA persisted by adhering

to cells or soil particles, or was transiently taken up by the bacteria. During the culture-dependant portion of the same experiment, 4000 bacteria grown on selective agar were screened, although none of them was found to contain the transgenic construct. A hurdle in this methodology is that naturally occurring resistance genes make it difficult to screen potential transformants from naturally resistant species (Nielsen & Townsend, 2004).

Other factors which make detection of HGT events to microorganisms from complex natural environments extremely challenging are the size and diversity of the microbial community, and the limitations imposed by the methodologies used. Use of culture-dependent techniques excludes non-culturable organisms which make up a large portion of environmental microbes (Rizzo *et al.*, 2013). However, the culturable organisms present a few grams of soil can be an unmanageable amount to test, even if a screening technique is employed. In the Gebhard and Smalla (1998) study mentioned above, 10 000 isolates were part of the initial screening before 4000 of those were screened for the transgene). In spite of the effort involved, such techniques may not be sensitive enough to detect HGT events (Nielsen *et al.*, 2013; Nielsen & Townsend, 2004). HGT events of very short fragments, such as those described by Overballe-Petersen *et al.* (2013) and Harms *et al.* (2016), would be even more difficult to detect, since differentiating between the endogenous DNA of uncharacterised environmental bacteria, short regions of exogenous inserts, and variations arising due to mutations and polymorphisms, may be virtually impossible (Hanage, 2016; Harms *et al.*, 2016; van Pijkeren & Britton, 2012).

This study used a method similar to that employed by Douville *et al.* (2009): instead of screening for antibiotic resistance genes which might be conferred from marker genes contained in transgenic crops, the aim was to detect *cry1Ab* transgene constructs of the MON810 event. Although a culture-dependent method was employed, no screening of microorganisms for the presence of the transgenic constructs was done at the culturing stage. For the detection of MON810 sequences, a collection of primers was selected from previous studies, which used PCR-based methods in a number of different contexts. These include detecting food and feed contamination (Hernández *et al.*, 2003; Kuribara *et al.*, 2002; Matsuoka *et al.*, 2000), detecting gene flow (Pineyro-Nelson *et al.*, 2009; Quist & Chapela, 2001), measuring longevity and transport of transgenes under field conditions (Douville *et al.*, 2009; Douville *et al.*, 2007), and characterizing the transgene construct and its junction regions where it was integrated into the maize genome (Rosati *et al.*, 2008).

The transgene detection methods and primers were designed for unambiguous detection of transgenic DNA, to ensure traceability of GM crops in food and feed products. However, while positive results for some primer sets indicate that transgenic DNA is present, they do not

necessarily specify which transgenic event is the source, since they target transgenic regions such as promoters which may be present in several events and crops (Hernández *et al.*, 2003; Kuribara *et al.*, 2002). For example, the promoter which is present in MON810, is also present in lines of cotton and soybean. Other primer sets target transgenic regions which are specific to certain events, such as the junction between introns and coding sequences, or junctions between transgene and genome where integration occurred (Hernández *et al.*, 2003; Kuribara *et al.*, 2002; Rosati *et al.*, 2008). The suite of primers selected for this study include both types. By combining different forward and reverse primers, it was possible to target regions across the length of the transgene.

In this chapter, the aim was to determine whether transgene fragments associated with genetically modified maize event MON810 could be detected in DNA isolated from aquatic microorganisms, as an indicator of possible horizontal gene transfer of transgenic DNA. Water samples were collected from the Vaalharts Irrigation Scheme (study site) and Tshiombo Irrigation scheme (negative control site). Microorganisms (bacteria, yeast and fungi) were cultured from the water samples, and DNA isolated from selected colonies. A PCR-based approach was used to detect whether transgenic constructs were present in the DNA isolated from the microorganisms.

4.2 Materials and Methods

4.2.1 Sampling Procedure

The sampling locations were described in the methods section of the previous chapter (see sections 3.1.3 and 3.2.1-3.2.2). The study area was the Vaalharts Irrigation Scheme, where Bt maize has been grown in this area since its commercial release in South Africa in 1998 (Kruger *et al*., 2009). Five sampling sites were selected at water bodies across the irrigation scheme in 2011. These were: Site A (Spitskop Dam); Site B (Ganspan Dam); Site C (Espagsdrif); Site D (Dam 8) and Site E (North Canal). An earlier sampling trip in 2010 collected samples from Site A, B and C only. The control site was the Tshiombo Irrigation Scheme is situated in Limpopo in the Venda region, where no GM maize was planted at the time of sampling. Even though three locations (a dam, a river and a canal) were selected within the Tshiombo scheme, the scheme as a whole is viewed as one control site. The three sampling points were the Maraxwe Canal; the Mohale River and the Mbahela Dam.

Water samples (2 L per site) were collected from each site and contained in sterilised bottles. The samples were then stored in cooler boxes until they could be processed in the laboratory (± 24 hours). For each site, two dilution series (10^{-1} – 10^{-6}) were set up using the sample water. The

dilution series were then used to inoculate various types of agar using the spread plate technique. R2A agar (Merck, Germany) and nutrient agar (Merk, Germany) were used for the cultivation of bacteria. For the cultivation of yeast and filamentous fungi PDA (Potato Dextrose Agar), SDA (Sabouraud Dextrose Agar), YEA (Yeast Extract Agar) and MEA (Malt Extract Agar) were used (all Merck, Germany). A variety of media were selected for fungi and yeast cultivation in order to obtain a wider diversity of isolates (Pereira *et al.*, 2010). To suppress bacterial growth, the PDA, SDA, YEA and MEA were supplemented with 0.03 mg/ml chloramphenicol (Sigma-Aldrich, US). In addition to the chloramphenicol, the MEA contained 0.1 mg/ml streptomycin (Sigma-Aldrich, US).

The nutrient agar plates were incubated at 37ºC for 1-2 days. The other agar plates were incubated at room temperature for 5 days. The plates were then examined for growth, and colonies were selected for further purification based on morphology. Selected colonies were then transferred to fresh agar plates (of the same type as that on which the colony was originally cultivated) using the streak plate method. This was repeated at least twice to generate pure cultures. In the case of filamentous fungi, blocks of mycelium mats were cut and transferred to fresh plates. Gram staining was then done on the purified colonies (with the exception of filamentous fungi) to ensure that they were single cultures.

4.2.2 DNA Isolation

For each type of organism, a suitable DNA isolation method was selected, using a commercially available kit where applicable. Following isolation, the concentration of the DNA was determined using a NanoDrop™ 1000 Spectrophotometer (Thermo Fischer Scientific, US). Aside from concentration, the NanoDrop™ reading also determined the 260:280 and 230:260 ratios. DNA with a 260:280 ratio of 1.7-1.9 is considered to be of good quality, and relatively free of RNA and protein contamination, while a 230:260 of below 2 is indicative of organic contamination (NanoDrop technical brief).

Bacteria

Selected colonies were inoculated into nutrient broth (Merck, Germany), which was then incubated overnight before DNA isolation. In cases where the bacteria did not grow in the broth, culture was scraped directly from the agar plate. A pre-incubation with 20 mg/ml lysozyme was found to improve the overall yield of DNA extractions (particularly for Gram positive bacteria), so the cells were incubated in lysozyme at 37ºC for up to 1 hour before DNA isolation. DNA isolation

then continued using the NucleoSpin® Tissue Kit (Macherey-Nagel, Germany), following the manufacturer's support protocol for bacterial DNA isolations.

Fungi (yeast and filamentous fungi)

DNA was isolated from yeast isolates using the protocol described by Tarr (2004), with a few modifications. The washing steps of this protocol were replaced. Instead of pelleting the DNA after precipitation with isopropanol, the iso-propanol DNA mixture was passed through a NucleoSpin® column. The wash and elution buffers from the NucleoSpin® Tissue kit (Macherey-Nagel, Germany) were then used to wash and elute the DNA. In the case of filamentous fungi, the mycelium mats were scraped from the agar plates and placed in 2 ml microcentrifuge tubes, after which the NucleoSpin® Plant II DNA isolation kit (Macherey-Nagel, Germany) was used according to the manufacturer's support protocol.

Maize

MON810 maize DNA was required as a positive control. Maize plants were grown at EcoRehab (NWU). Leaves were used for DNA isolation, using the NucleoSpin Plant II kit (Macherey-Nagel, Germany) according to the protocol supplied by the manufacturer.

4.2.3 DNA Amplification

Universal Primers

"Universal" primers refer to primers which are theoretically able to amplify any representative of a certain group of organisms. They are frequently used to identify organisms. The purpose of the universal primers was firstly, to demonstrate that the genomic DNA of the isolate was amplifiable; and secondly, providing a template for sequencing to identify the isolate at a later stage. The universal primers used in this chapter are included in Table 4.1 below.

The PCR reaction mixture contained 12.5 µl double strength PCR master mix [(0.05 U/µl Taq DNA Polymerase in reaction buffer, 0.4 mM of each dNTP; 4 mM $MgCl_2$ (Fermentas Life Sciences, US)]; 10 pmol of each primer (Applied Biosystems, UK); 50 ng DNA template; and PCR-grade nuclease-free water (Fermentas Life Science, US) to fill up the volume to 25µl. The reaction took place in an ICycler thermal cycler (BioRad, UK) under the following conditions: 95°C for 300 s; 35 cycles of 95°C for 30 s; 52-55°C for 30 s; 72°C for 60 s; and final extension at 72°C for 420 s.

Table 4.1: The sequences of the universal primers used in this study, as well as their amplicon sizes and target regions

Region targeted	Primer Name	Sequence	Amplicon Length	Reference
16S rRNA: bacteria	27F 1492R	5‘-AGAGTTTGATCMTGGCTCAG-3’ 5‘-TACGGYTACCTTGTTACGACTT-3’	1465	Lane (1991)
18S rRNA: fungi	EF3 EF4	5’-TCCTCTAAATGACCAAGTTTG-3’ 5’-GGAAGGGRTGTATTTATTAG-3’	1500	Smit *et al.* (1999)
26S rRNA: yeast	NL1 NL4	5’-GCATATCAATAAGCGGAGGAAAAG-3’ 5’-GGTCCGTGTTTCAAGACGG-3’	400 bp	Kurtzman and Robnett (1998)

PCR with MON810 Primers

Primers were selected from previous studies to detect the presence of MON810 transgene sequences in the DNA of the aquatic organisms. Eight primer set combinations were chosen, (see Tables 4.2 and 4.3). By combining different forward and reverse primers, it would be possible to amplify regions across the length of the transgene. The lengths of the amplicons which are expected after such combinations, are given in Table 4.3. A graphical representation of the placement of the primers in relation to the transgene is given in Figure 4.1.

Table 4.2: The sequences of the primers used for detection of MON810 transgenic DNA in this study

Primer	Sequence	Reference
CM01 F	5’- CTCTACAAATGCCATCATTGCGATA-3’	Matsuoka *et al*., 1999
CM02 R	5’-CTTATATAGAGGAAGGGTCTTGCGA-3’	Matsuoka *et al*., 1999
HSP70 F	5’-GATGCCTTCTCCCTAGTGTTGA-3’	Kuribara *et al*. 2002
Cry1A(b) R	5’-GGATGCACTCGTTGATGTTTG-3’	Kuribara *et al*. 2002
1213 F	5’-GGCACGGTGGATTCCCTGGACGAGAT- 3’	Hernández *et al*., 2003
1213 R	5’-ATCTCGTCCAGGGAATCCACCGTGCC-3’	Hernández *et al*., 2003
2388 R	5’-GGAGAAGTGGTGGCTGTGGTGGGC-3’	Hernández *et al*., 2003

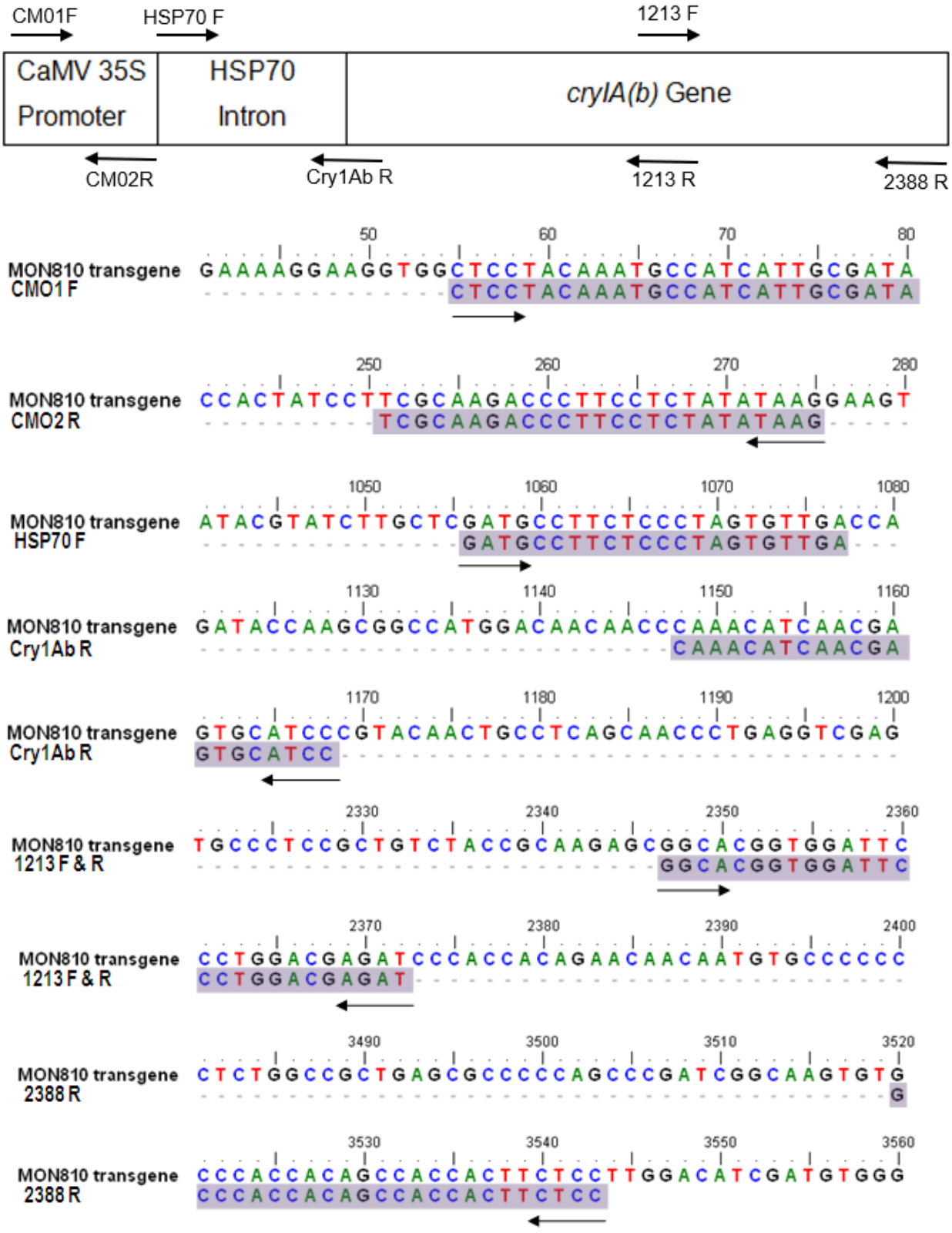


Figure 4.1: A graphical representation of the organisation of the MON 810 transgene. Below, an alignment of primer sequences with the MON810 transgene to indicate their positions. Regions which did not contain a primer sequence were omitted, though the numbers above the transgene sequence still reflect their positions relative to the transgene sequence.

With the exception of primer sets 4 and 7 (see Table 4.3 below for primer combinations), the amplification with MON810 primers took place in the following reaction mix: 12.5 µl double strength PCR master mix [(0.05 U/µl Taq DNA Polymerase in reaction buffer, 0.4 mM of each dNTP; 4 mM $MgCl_2$ (Fermentas Life Sciences, US)]; 20 pmol of each primer (Fermentas Life Sciences, US); 50 ng DNA template; and PCR-grade nuclease-free water (Fermentas Life Science, US) to fill up the volume to 25 µl. Each primer set had its own cycling conditions which are listed in Table 4.3 below.

For primer set 4 and 7, the reaction mix differed from the other sets and contained the following: 12.5 µl double strength PCR master mix [(0.05 U/µl Taq DNA Polymerase in reaction buffer, 0.4 mM of each dNTP; 4 mM $MgCl_2$ (Fermentas Life Sciences, US)]; 30 pmol of each primer (Fermentas Life Sciences, US); 100 ng DNA template; an additional 0.5 µl Taq polymerase (Kapa Biosystems, US), an additional 1 µl $MgCl_2$ (Fermentas Life Science, US), and PCR-grade nuclease-free water (Fermentas Life Science, US) to fill up the volume to 25 µl.

For each PCR reaction using MON810 primers, a positive control (isolated MON810 maize DNA), a negative control (DNA from a reference site isolate), and a no-template control (water) were used for quality control purposes and to indicate whether contamination had taken place. Agarose gel electrophoresis was done with the products following PCR.

4.2.4 Electrophoresis and visualization of PCR products

During the course of this project, the Subject Group Microbiology transitioned from using ethidium bromide to visualise DNA, to using Gel Red (Biotium, US). At the same time, the GeneGenius Bio Imaging System (Syngene Synoptics, UK), and accompanying GeneSnap software (version6.00.22) was replaced by the BioRad GelDoc Imaging System (BioRad, UK). Thus, reactions which were completed earlier in the project were analysed using ethidium bromide and the GeneGenius and GeneSnap imaging system.

When ethidium bromide was used, 5 µl of PCR product was mixed with 5 µl of 6 x O'range Loading Dye (Fermentas Life Sciences, US) and loaded into a 1.5% (w/v) agarose gel containing 0.001 mg/ml ethidium bromide (BioRad, UK). When Gel Red was used, 3 µl of loading dye containing Gel Red was mixed with 3 µl of PCR product, and loaded into a 1.5% agarose gel which did not contain any ethidium bromide. Before loading the samples, the agarose gel was placed into 1 x TAE buffer (20mM acetic acid (Merck, US); 40mM Tris (Sigma Aldrich, US); 1mM EDTA (Merck, US) at pH 8). Electrophoresis then proceeded at 80V for 50 minutes. Amplicons were visualised

using either the GeneGenius Bio Imaging System (Syngene Synoptics, UK), and accompanying GeneSnap software (version6.00.22); or the BioRad Gel Doc imaging system (BioRad, UK).

Table 4.3: The eight primer sets used for amplification of transgene fragments, with expected amplicon sizes and cycling conditions

Primer Set	Primers Used	Cycling Conditions	Expected Amplicon Size
1	CM01 F CM02 R	95°C for 300 s; 40 cycles of 95°C for 30 s; 66°C for 30 s; 72°C for 30 s; and final extension at 72°C for 420 s.	220 bp
2	CM01 F CryIAb R	95°C for 300 s; 40 cycles of 95°C for 30 s; 61°C for 30 s; 72°C for 60 s; and final extension at 72°C for 420 s.	1113 bp
3	CM01 F 1213 R	95°C for 300 s; 40 cycles of 95°C for 30 s; 62°C for 30 s; 72°C for 120 s; and final extension at 72°C for 420 s.	1957 bp
4	CM01 F 2388 R	95°C for 360 s; 33 cycles of 95°C for 30 s, annealing at 70°C (decreasing by 0.5°C every cycle until 55°C was reached), extension at 72°C for 240 s; followed by an additional 10 cycles of 95°C for 30s, 60°C for 30 s, 72°C for 180 s and a final extension at 72°C for 420 s.	3488 bp
5	HSP70 F CryIAb R	95°C for 300s; 35 cycles of 95°C for 30 s; 59°C for 30 s; 72°C for 30 s; and final extension at 72°C for 420 s.	113 bp
6	HSP70 F 1213 R	95°C for 300 s; 40 cycles of 95°C for 30 s; 65°C for 30 s; 72°C for 90 s; and final extension at 72°C for 420 s.	1317 bp
7	HSP70 F 2388 R	95°C for 6 min; 33 cycles of 95°C for 30 s, annealing at 70°C (decreasing by 0.5°C every cycle until 55°C was reached), extension at 72°C for 180 s; followed by an additional 10 cycles of 95°C for 30s, 60°C for 30 s, 72°C for 180 s and a final extension at 72°C for 420 s.	2488 bp
8	1213 F	95°C for 300 s; 40 cycles of 95°C for 30 s; 68°C for 30 s; 72°C for 60 s; and final extension at 72°C for 420 s.	1197 bp

4.2.5 Sanger Sequencing

Positive PCR results (amplicons) were selected for sequencing. The NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel, Germany) was used to purify the amplicon DNA and to remove traces of PCR reagents and primer-dimers. The concentration and 260:280 ratio of the purified amplicon DNA were determined using the NanoDrop™ 1000 Spectrophotometer (Thermo Fischer Scientific, US). This DNA then served as the template for sequencing PCR.

The BigDye® Terminator v3.1Cycle Sequencing Kit (Applied Biosystems, UK) was used for the sequencing PCR. The reaction mixture contained 2.5 X Ready Reaction Premix (Applied Biosystems, UK), 5 X BigDye Sequencing Buffer (Applied Biosystems, UK), 3.2 pmol of either the forward or reverse primer (Applied Biosystems, UK) – the same as was used in the original PCR reaction - and PCR-grade nuclease-free water (Fermentas Life Science, US) to fill up the volume to 20 µl. The reaction took place in an ICycler thermal cycler (BioRad, UK) under the following conditions: 96°C for 1 min; 25 cycles of 96°C for 10 s; 50°C for 5 s; and extension at 60°C for 4 min followed by holding at 4°C until purification as per the manufacturer's instructions.

The products of this reaction were purified again using the Zymo Research DNA Sequencing Clean-up Kit™ (Zymo Research, US) as per the manufacturer's instructions. The eluent was loaded into a 96-well plate and placed in the 3130 Applied Biosystems Genetic Analyser, where sequencing electrophoresis took place using a 36 cm capillary array and POP-7™polymer (all Applied Biosystems, UK). Analysis of the data was done by the 3130 Genetic Analyser Data Collection software (Applied Biosystems, UK). The resulting chromatograms were analysed in BioEdit v7.1.3 software, and were entered into BLAST (Basic Local Alignment Search Tool (NCBI, 2014)) to confirm their identities.

4.3 Results

4.3.1 Sample collection and isolation of microorganisms

The number of isolates included in the present study comprised of 298 bacteria isolates, and 123 yeast and fungi isolates. Sites A, B and C have more bacteria isolates than the other sites due to a carry-over of isolates from a previous sampling trip in 2010. The date of the sampling trip, as well as the site and type of organism, can be gleaned from the code which was ascribed to each isolate or specimen. Thus, A10-11/04/08bac represents a bacterial (bac) isolate A10, which was cultivated from a water sample collected from site A on 2011/04/08. Yeast isolates end in the suffix "ye", filamentous fungi in "fn".

4.3.2 DNA Isolation

After DNA isolation, the concentration and purity of the DNA was determined spectrophotometrically using the NanoDrop spectrophotometer. Concentrations of isolated DNA ranged from 3 ng/µl to 841 ng/µl, while the 260:280 ratios of the DNA samples were between 1.5 and 2.4. The wide ranges in these values can be attributed to the large variety of different organisms. All isolates underwent PCR with appropriate "universal" primers to ensure that the DNA was amplifiable.

4.3.3 DNA Amplification

Universal primers

Genomic DNA isolated from each isolate was used as a template for PCR using appropriate universal primers. Isolates which failed to generate amplicons of the expected size at this stage were rejected from PCR with MON810 primers. After each PCR amplification, the PCR products were subject to agarose gel electrophoresis and visualised to determine whether amplification had taken place.

Amplification with MON810 primers

In order to determine that all 8 primer combinations were capable of successfully amplifying a product of the expected size, PCR was done with each set using MON810 maize (positive control) DNA. A compilation of these results can be seen in Figure 4.1 below.

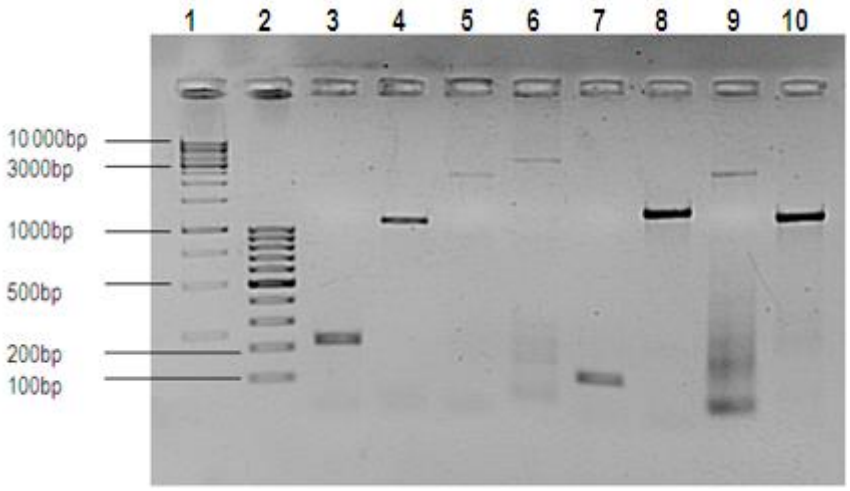


Figure 4.2: **A negative image of a 1.5% (w/v) agarose gel containing 0.001mg/ml ethidium bromide, showing the amplicons of all 8 primer sets (lanes 3-10 represent primer sets 1-8 respectively). Lanes 1 and 2 contain a 1 kb and 100 bp molecular weight markers (O'GeneRuler™, Fermentas Life Science, US), respectively.**

Figure 4.2 indicates the successful amplification of MON810 maize DNA using the primer sets and conditions listed in Table 4.3. In each case, an amplicon of the expected size was generated. However, when applied to the non-maize DNA, only primer sets 1, 5 and 8 were able to produce amplicons (although not always of the expected size).

MON810 primers: Bacteria isolates

PCR with primer set 1 indicated 16 positive results. Below (Figure 4.3) are inverted images of two gels which contain the products of such reactions.

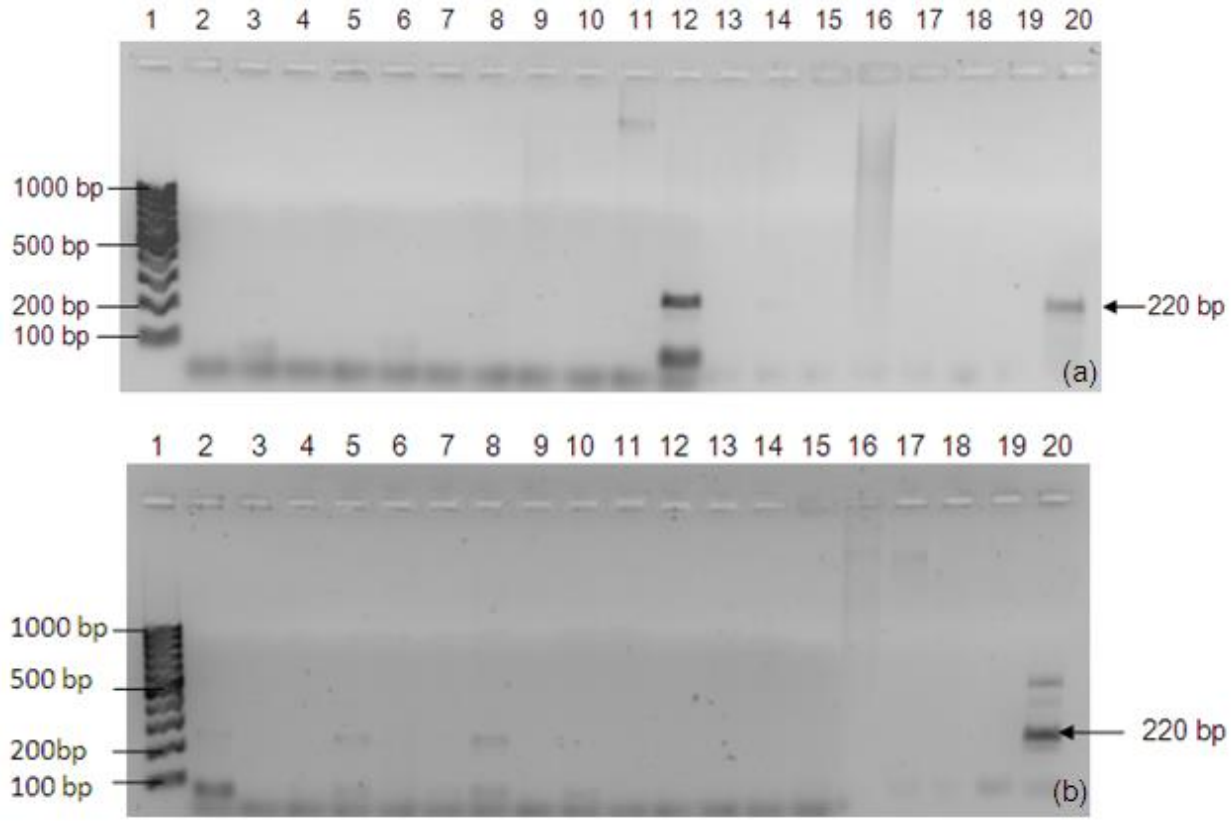


Figure 4.3: **Examples of results obtained after amplification of bacteria DNA with set 1 primers. The PCR products underwent electrophoresis in a 1.5% agarose gel and were stained using GelRed. Lane 1 contains the 100 bp molecular weight marker (O'GeneRuler™, Fermentas Life Science, US). In both gel (a) and gel (b), lanes 17 and 18 contain the negative control and no-template controls respectively. The positive control (lane 20 in both Figures) indicates the expected 220 bp amplicon**

In Figure 4.3(a) above, positive results can be seen in lane 12 (B8-11/04/08bac), as well as in lanes 1 (D15-11/04/08bac), 5 (D32-11/04/08bac) and 8 (D35-11/04/08bac) of 4.3(b). In both cases, lane 18 was the negative control, lane 19 the no-template control, and lane 20 the positive control. A total of 16 bacteria isolates indicated similar positive results with this primer set. Each site had representatives amongst these positive results, except the reference site. The division between the sites was as follows: site A, 1 isolate; site B, 5 isolates; site C, 3 isolates; site D, 6 isolates and site E, 1 isolate.

More positive results were detected with primer set 5 than any other primer set. This can be seen in the agarose gel electrophoresis of the primer set 5 PCR products which are shown in Figure 4.4 below.

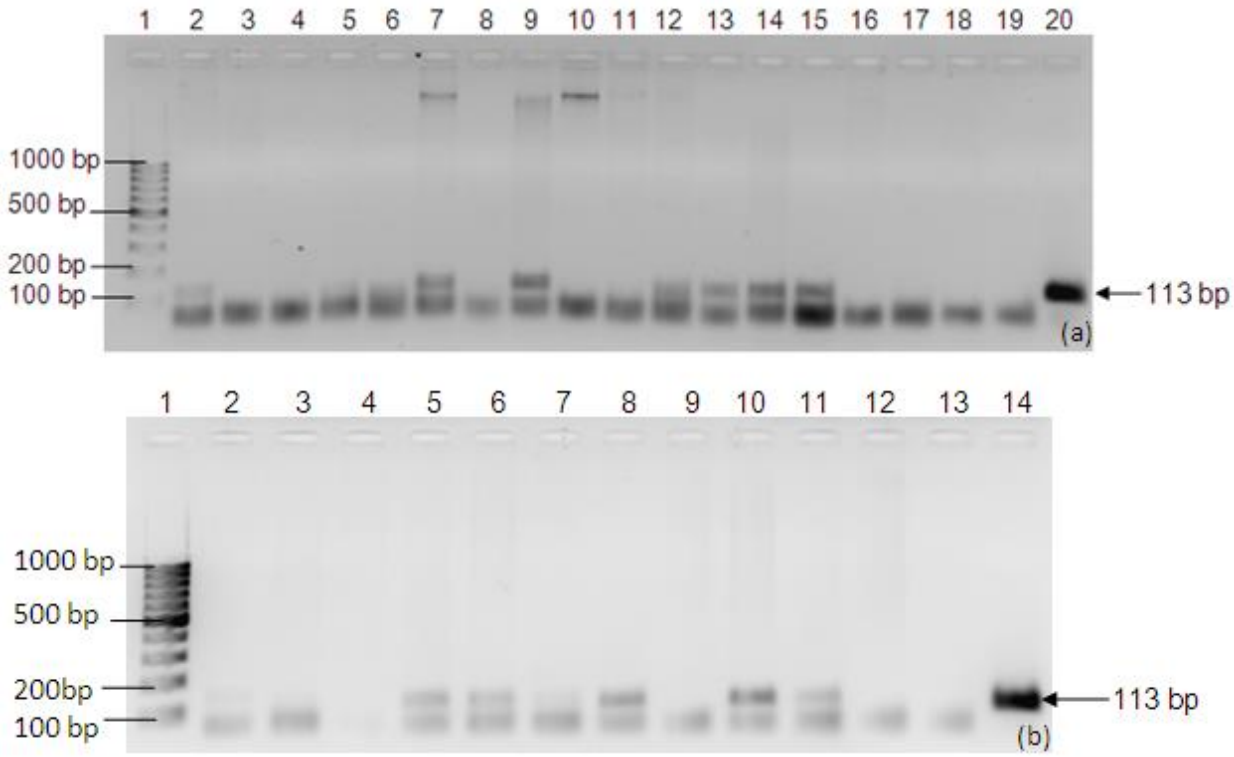


Figure 4.4: **An example of the results obtained after performing PCR on bacteria genomic DNA using primer set 5. Following PCR, the amplicons underwent electrophoresis in a 1.5% agarose gel and were stained with GelRed. Lanes 18-20 contain the negative control, no-template control and positive control, respectively. The molecular weight marker (O'GeneRuler™, Fermentas Life Science, US) can be seen in lane 1.**

Set 5 generated the most positive results of all the primer sets tested, 45 in total among the bacterial isolates. Examples of positive results can be seen in Figure 4.4 (a) and 4.4(b), which show representatives of sites D and B respectively. Lanes 7, 9 and 12-15 of Figure 4.4(a) contain bands which correspond to the expected 113 bp fragment size (isolates D22-11/04/08bac, D24-11/04/08bac, D7-11/04/08bac, D10-11/04/08bac, D29-11/04/08bac and D15-11/04/08bac respectively). In Figure 4.4 (b), the amplicons in lanes 5, 6, 8, 10 and 11 belong to B51-10/06/11bac, B59-10/06/11bac, B19-11/04/08bac, B52-10/06/11bac and B9-11/04/08bac, as well as lane 14 which contains the positive control. Positive results (see Table 4.4) included 13 isolates for site A, 5 for site B, 14 for site C, 6 for site D and 7 for site E and none for the reference site.

Primer set 8, although successful in amplifying MON810 maize DNA, was unable to generate amplicons of the correct size when used with DNA from bacterial isolates, as can be seen in Figure 4.5 below.

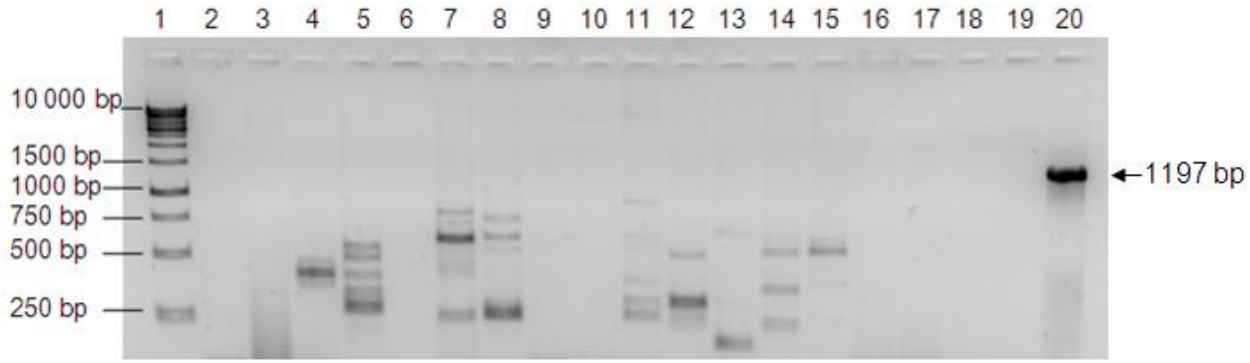


Figure 4.5: Inverted image of a 1.5% gel containing the PCR products of amplification of bacterial DNA with primer set 8. The 1 kb molecular weight marker (O'GeneRuler™, Fermentas Life Science, US) was run in lane 1, while lanes 18-20 contain the controls (negative, no-template and positive, respectively).

In the case of primer set 8, no bacterial isolates produced fragments of the expected 1197 bp size (see Figure 4.5 above), although for the positive control amplified a fragment of the expected size. A plethora of smaller fragments were generated, all smaller than 800 bp. In lanes 4 and 15 in Figure 3.4, for example, single bands of approximately 500 bp were produced. In lanes such as 5, 7-8, and 11-14 of Figure 4.5, multiple bands can be observed. In some cases, even the negative control produced fragments. Efforts to sequence these fragments to determine whether they bore

any similarity to the MON810 transgene sequence were largely unsuccessful. The few short sequences obtained after sequencing did not align with the MON810 transgene.

MON810 primers: Yeast and Fungi isolates

Figure 4.6 is an inverted image of an agarose gel following electrophoresis of yeast and fungi amplicons after PCR with primer set 1.

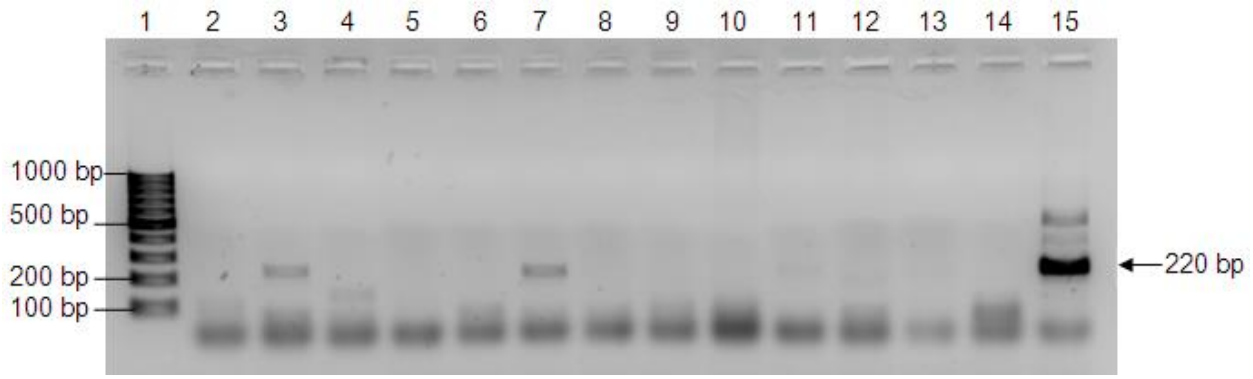


Figure 4.6: **Above is an example the results PCR with primer set 1 of yeast and fungi DNA. Amplicons were electrophoresed in a 1.5% agarose gel and stained with GelRed. The 100 bp molecular weight marker (O'GeneRuler™, Fermentas Life Science, US) was loaded into lane 1, while lanes 13 and 14 contain the negative control and no-template control respectively. The positive control can be seen in lane 15**

Positive results can be seen in lanes 3 and 7 of Figure 4.6 above. For the yeast and fungi, these were the only two positive results (B69-11/04/08ye and D53-11/04/08fn) obtained with primer set 1.

Primer set 5 was more successful than primer set 1, 19 positive results were obtained with this primer set. In Figure 4.7 (a) below, lanes 2 (D47-11/04/08fn) and 3 (D53-11/04/08fn) contain amplicons which match the size of that of the positive control. Lanes 9 (E30-11/04/08ye), 11-13 (C76-11/04/08ye, B62-11/04/08ye, and E35-11/04/08ye) and 15-16 (E37-11/04/08ye and B69-11/04/08ye) of Figure 4.7 (b) also contain amplicons. A total of 19 positive results were obtained after amplification with primer set 5. There is a bias of positive results in favour of site A, which represents 11 of the potentially positive samples. The remaining 8 positive results are more evenly spread across the sites: 1 from site B, 2 from site C, 2 from site D, and 3 from site E.

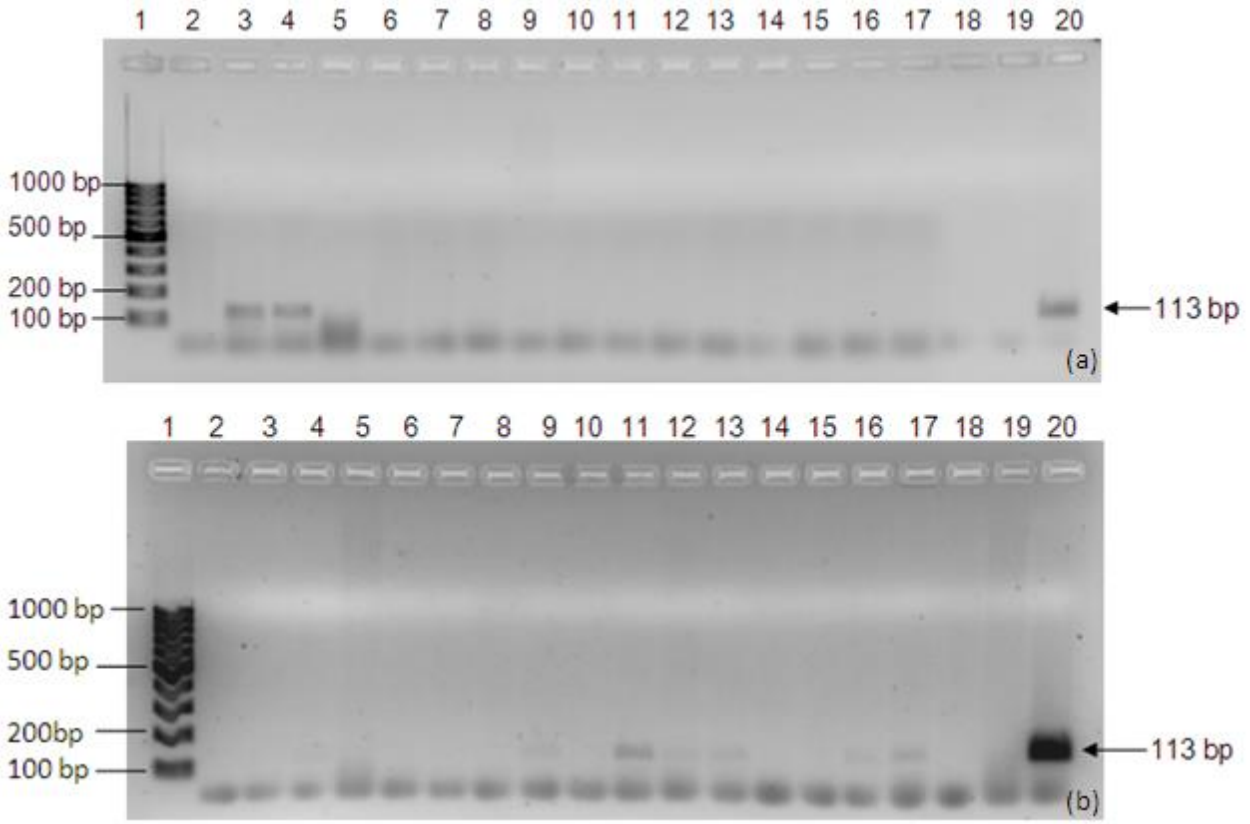


Figure 4.7: **Inverted images of electrophoresis results of fungi and yeast DNA amplified with primer set 5. Lane 20 contains the positive control (in both gel (a) and (b)), which is the expected 113 bp amplicon size. In both Figures (a) and (b), lane 18 contains the negative control, while lane 19 contains the no-template control and lane 1 the molecular weight marker (O'GeneRuler™, Fermentas Life Science, US).**

Figure 4.8 is an inverted image of an agarose gel following electrophoresis of yeast and fungi amplicons after PCR with primer set 8, stained with GelRed.

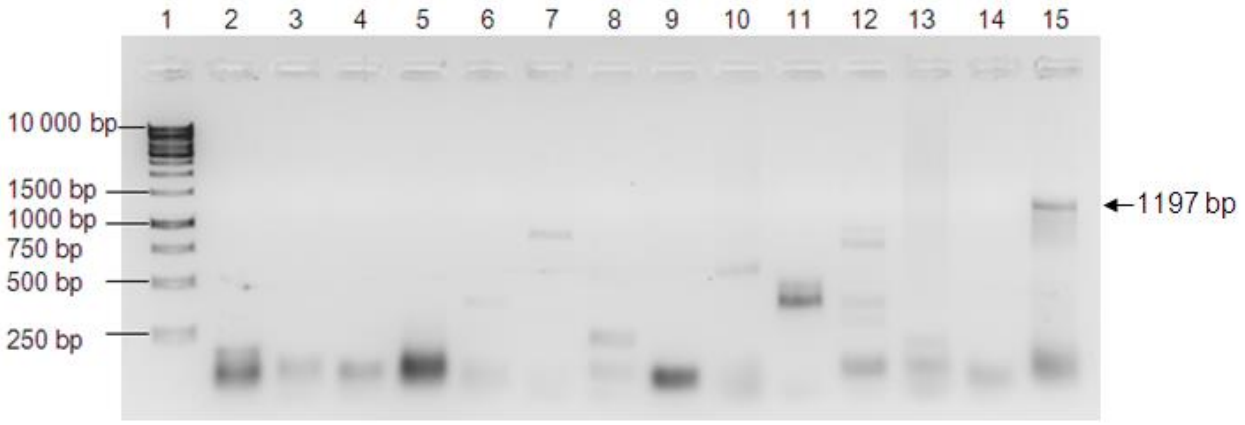


Figure 4.8: **An inverted image of a 1.5% agarose gel after electrophoresis of PCR products stained generated by the amplification of yeast and fungi DNA with primer set 8. The molecular weight marker was loaded into lane 1 (O'GeneRuler™, Fermentas Life Science, US), while lanes 13-15 contain the controls (negative, no-template and positive, respectively).**

As seen with the bacteria isolates, no yeast or fungi isolates generated an amplicon of the correct size for primer set 8 (see Figure .7 above). Lane 7 and 12 of Figure 4.8 indicate fragments which are approximately 750-800 bp in size, although most other isolates which produced amplicons, produced fragments smaller than 800 bp. Isolates frequently did not produce any fragments, as can be seen lanes 2-5 and 8-9 of Figure 4.8

4.3.4 Sequencing

The PCR products of samples which produced a positive result were sequenced to determine their similarity to the MON810 transgene. Sample sequences were aligned with the transgene (accession number AY326434.1) in BioEdit (v7.1.3). Below, an example of such an alignment is given in Figure 4.9.

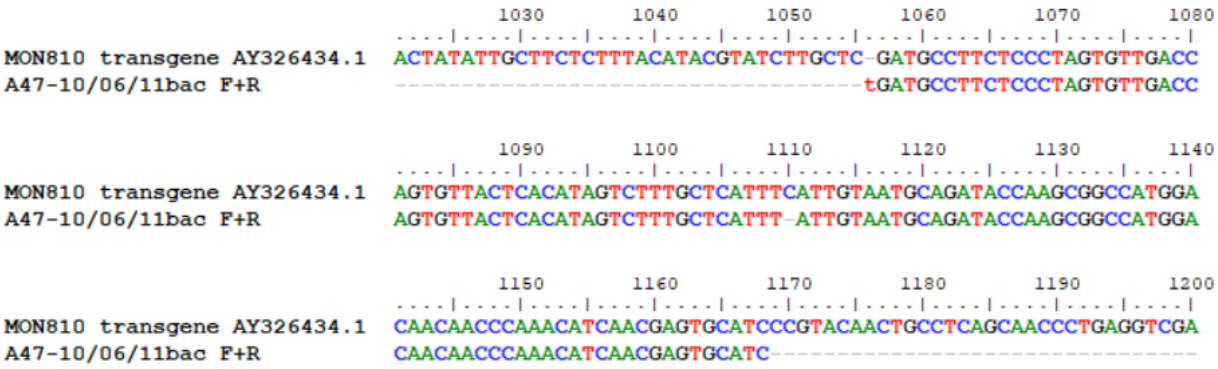


Figure 4.9: **An example of an alignment of sequenced PCR amplicons from one of the isolates (A47-10/06/11bac), with the MON810 transgene.**

While a few differences between the sample sequence and the transgene sequence can be seen, the sample sequence does follow that of the transgene closely.

The 16S (bacteria), 18S (fungi) or 23S (yeast) rRNA sequences of isolates which gave positive results were also sequenced, and the resulting sequences were used to identify the isolates by using BLAST.

Table 4.4: **List of bacteria isolates which generated positive results with primer sets 1 and 5, as identified after sequencing 16S amplicons and using BLAST to identify the sequences.**

Isolate Code	Primer set(s)	Organism GenBank ID	e value	% similarity	Accession Number
A22-11/04/08bac	5	*Bacillus marisflavi*	0.00	99	MK182898
A29-11/04/08bac	1	*Aeromonas veronii*	0.00	100	MK182872
A34-11/04/08bac	5	*Bacillus* sp.	0.00	99	MK182894
A40-10/06/11bac	5	*Aeromonas* sp.	0.00	99	MK182890
A42-10/06/11bac	5	*Bacillus safensis*	0.00	99	MK182900
A43-10/06/11bac	5	*Aeromonas veronii*	0.00	99	MK182892
A46-10/06/11bac	5	*Bacillus safensis*	0.00	100	MK182901
A47-10/06/11bac	5	*Fictibacillus* sp.	0.00	99	MK182903
A52-10/06/11bac	5	*Bhargavaea ginsengi*	0.00	99	MK182895
A53-10/06/11bac	5	*Bacillus* sp.	0.00	99	MK182899
A54-10/06/11bac	5	*Bacillus subtilis*	0.00	100	MK182905
A55-10/06/11bac	5	*Aeromonas veronii*	0.00	100	MK182893
A59-11/04/08bac	5	*Pseudomonas* sp.	0.00	99	MK182915
A61-11/04/08bac	5	*Pseudomonas* sp.	0.00	100	MK182918
B1-11/04/08bac	1	*Staphylococcus cohnii*	0.00	100	MK182887
B19-11/04/08bac	5	*Rheinheimera* sp.	0.00	99	MK182920
B20-11/04/08bac	1	*Klebsiella aerogenes*	0.00	99	MK182881

B22-11/04/08bac	1	*Staphylococcus cohnii*	0.00	99	MK182882
B36-11/04/08bac	1	*Bacillus* sp.	0.00	99	MK182874
B51-10/06/11bac	5	*Bacillus* sp.	0.00	99	MK182896
B52-10/06/11bac	5	*Aeromonas salmonicida*	0.00	100	MK182891
B59-10/06/11bac	5	*Bacillus* sp.	0.00	99	MK182897
B8-11/04/08bac	1	*Staphylococcus cohnii*	0.00	100	MK182886
B9-11/04/08bac	5	*Staphylococcus cohnii*	0.00	99	MK182923
C15-11/04/08bac	1	*Arthrobacter* sp.	0.00	99	MK182873
C20-11/04/08bac	5	*Micrococcus luteus*	0.00	99	MK182913
C26-11/04/08bac	5	*Devosia* sp.	0.00	99	MK182907
C32-11/04/08bac	5	*Pseudomonas protegens*	0.00	99	MK182917
C34-11/04/08bac	5	*Microbacterium testaceum*	0.00	99	MK182912
C38-11/04/08bac	5	*Pseudomonas* sp.	0.00	99	MK182916
C39-11/04/08bac	5	*Aeromonas* sp.	0.00	99	MK182888
C4-11/04/08bac	1 & 5	*Bacillus* sp.	0.00	99	MK182876
C41-11/04/08bac	1 & 5	*Pseudomonas mendocina*	0.00	99	MK182883
C56-10/06/11bac	5	*Pseudomonas* sp.	0.00	99	MK182919
C62-10/06/11bac	5	*Bacillus* sp.	0.00	99	MK182906
C64-10/06/11bac	5	*Micrococcus luteus*	0.00	99	MK182914
C65-10/06/11bac	5	*Bacillus safensis*	0.00	99	MK182902
C71-10/06/11bac	5	*Rheinheimera tangshanensis*	0.00	99	MK182921
C8-11/04/08bac	5	*Staphylococcus cohnii*	0.00	99	MK182926
D10-11/04/08bac	5	*Microbacterium imperiale*	0.00	99	MK182911
D15-11/04/08bac	1 & 5	*Pseudomonas protegens*	0.00	100	MK182884
D2-11/04/08bac	1	*Bacillus* sp.	0.00	99	MK182877
D22-11/04/08bac	5	*Staphylococcus cohnii*	0.00	99	MK182927
D23-11/04/08bac	1	*Pseudomonas* sp.	0.00	100	MK182885
D24-11/04/08bac	5	*Hydrogenophaga pseudoflava*	0.00	99	MK182909
D29-11/04/08bac	5	*Exiguobacterium acetylicum*	0.00	99	MK182908
D32-11/04/08bac	1	*Bacillus* sp.	0.00	99	MK182875
D35-11/04/08bac	1	*Bacillus safensis*	0.00	100	MK182879
D7-11/04/08bac	1 & 5	*Microbacterium* sp.	0.00	99	MK182880
E10-11/04/08bac	5	*Serratia fonticola*	0.00	100	MK182922
E15-11/04/08bac	5	*Janibacter melonis*	0.00	100	MK182910
E17-11/04/08bac	5	*Staphylococcus arlettae*	0.00	99	MK182924
E2-11/04/08bac	1 & 5	*Massilia* sp.	0.00	99	MK182878
E22-11/04/08bac	5	*Staphylococcus arlettae*	0.00	99	MK182925
E23-11/04/08bac	5	*Aeromonas hydrophila*	0.00	99	MK182889
E25-11/04/08bac	5	*Bacillus subtilis*	0.00	97	MK182904

Selected isolates from Table 4.4 were chosen for further analysis in Chapter 5.

Table 4.5: List of eukaryote isolates which generated positive results, as identified after sequencing 18S or 23S amplicons and using BLAST to identify the sequences.

Isolate Code	Primer set(s)	Organism GenBank ID	e value	% similarity	Accession Number
A70-11/04/08fn	5	*Alternaria alternata*	0.00	99	MK177499
A71-11/04/08fn	5	*Pulcherricium caeruleum*	0.00	99	MK177500
A72-11/04/08fn	5	*Pleosporales* sp.	0.00	99	MK177501
A73-11/04/08fn	5	*Entyloma gaillardianum*	0.00	100	MK177502
A74-11/04/08fn	5	*Pleosporales* sp.	0.00	100	MK177503
A76-11/04/08fn	5	*Cheimonophyllum candidissimum*	0.00	99	MK177504
A77-11/04/08fn	5	*Marchandiomyces buckii*	0.00	99	MK177505
A78-11/04/08fn	5	*Massarina phragmiticola*	0.00	99	MK177506
A83-11/04/08fn	5	*Phoma* sp.	8.00E-94	99	MK177506
B62-11/04/08ye	5	*Rhodotorula* sp.	0.00	99	MK182281
B69-11/04/08ye	1	*Cryptococcus cistialbidi*	0.00	99	MK182280
C76-11/04/08ye	5	*Tremellales* sp.	0.00	99	MK182282
C90-11/04/08fn	5	*Pseudoclitocybe cyathiformis*	0.00	98	MK177508
D47-11/04/08fn	5	*Phomopsis* sp.	0.00	100	MK177509
D53-11/04/08fn	1 & 5	*Rhizochaete filamentosa*	0.00	99	MK177498
E30-11/04/08ye	5	*Tremellales* sp.	0.00	99	MK182283
E35-11/04/08ye	5	*Tremellales* sp.	0.00	100	MK182284
E37-11/04/08ye	5	*Bullera unica*	7.00E-169	100	MK182285
A69-11/04/08fn	5	Sequencing unsuccessful	-	-	-
A75-11/04/08fn	5	Sequencing unsuccessful	-	-	-

4.4 Discussion

4.4.1 General

The first objective of the current study was to collect samples of aquatic organisms and isolate their DNA. In total, 421 isolates were collected or cultured, and underwent DNA isolation. This number reflects the 298 bacterial isolates and 123 yeast and fungi isolates. The yield of the isolated DNA differed widely between the samples. PCR with 16S, 23S and 18S primers was done to determine whether the DNA was amplifiable during PCR. These amplicons also provided templates for sequencing and identification of isolates.

To use PCR-based methods to determine if *cry1Ab* genes and/or their promoters were present in the DNA isolated from aquatic organisms was the third objective of this study. The isolates' DNA underwent PCR with selected MON810 primer sets (see Table 4.3). Unfortunately, only primer sets 1 and 5 were successful in amplifying fragments of the correct size. Primer set 8 amplified non-specific fragments (see Figures 4.5 and 4.8), whereas the other primer sets were unsuccessful in amplifying fragments from any DNA save that of the positive control. Sequencing

of the fragments obtained during the standard PCR aligned well with the MON810 transgene. Considering the results with primer set 1 and primer set 5 together, positive results were detected among 56 of the bacterial isolates and 20 among the yeast and fungi.

4.4.2 Detection of MON810 sequences

Primer set 1 was developed to detect the presence of transgenic DNA in food (Matsuoka *et al.*, 2000). It was also successfully used by Quist and Chapela (2001) to detect transgenic DNA in native Mexican maize landraces, indicating that gene flow had taken place. This primer set amplifies the CaMV 35S promoter region, which is used in a number of GM crops including several events of Bt maize, cotton and soybean (Kuribara *et al.*, 2002; Matsuoka *et al.*, 2000; Randhawa *et al.*, 2010). Thus, while positive results for this primer set indicate that the CaMV sequence is present, it does not specify which transgenic event is present. Positive results with primer set 1 were obtained with 23 isolates. The detection of CaMV 35S promoter sequences raises concerns as to the effects that its insertion may have on the host organism. CaMV was suggested to be a "universal promoter" due to its ability to function in a wide variety of organisms, including bacteria (Assaad & Signer, 1990), yeasts (Pobjecky *et al.*, 1990) and basidomycete fungi (Sun *et al.*, 2002). Insertion of a promoter sequence may interfere with the regulation and transcription patterns of the organism's own genes (Traavik *et al.*, 2009).

Primer set 5 is the same primer pair which Douville *et al.* (2009) used to detect the presence of the MON810 transgene in freshwater mussel tissue and aquatic bacteria. This primer set is event-specific to MON810, and spans the junction between the *hsp70* intron of the construct and the *cry1Ab* gene (Kuribara *et al.*, 2002). Primer set 5 generated the most positive results of all the primer sets.

The amplicons of both these primer sets are very short (210 bp and 113 bp respectively). Primer sets which targeted longer regions were unsuccessful in amplifying them when using the isolates' DNA as a template, though all the primer sets successfully amplified their target regions when MON810 maize DNA was used (see Figure 4.1). Assuming that the amplified transgene regions detected by primer sets 1 and 5 are indicative of HGT and integration of transgene sequences, a number of explanations for the lack of success of the other primer sets exist. The shortness of the fragments detected suggests that the whole transgene was not transferred. The transgene may have become fragmented and damaged while exposed to the aquatic environment, resulting in it being broken up into shorter regions which were subsequently take up by bacteria. Damage to environmental DNA is expected and well-documented in other studies (Deagle *et al.*, 2006; Dejean *et al.*, 2011; Hajibabaei *et al.*, 2006). It is not uncommon for DNA to be reduced to

fragments in the region of 100 bp long (Overballe-Petersen *et al.*, 2013). However, although uptake of short DNA molecules has been documented in the lab, the extent to which it happens under natural conditions is currently unknown. The difficulty in detecting such events, which may be mistaken for polymorphisms and mutations, may contribute to the lack of detection. Several authors have expressed the view that this type of transformation may be an underestimated form of HGT (Bryan & Swanson, 2011; de Aldecoa *et al.*, 2017; Dutra *et al.*, 2007; Hanage, 2016; Harms *et al.*, 2016; Overballe-Petersen *et al.*, 2013).

The second possibility is that during insertion into the recipient genome, portions of the transgene containing the primer binding sites were lost or rearranged (Holst-Jensen *et al.*, 2012). If either primer binding sites are absent or altered, amplification of the target region may not be achieved. Transgene recombination upon insertion may also be the reason that primer set 8 generated amplicons which were not of the expected size, although non-specific binding may be a more likely explanation.

The number of positive results in this study may appear high when compared to studies such as that of Gebhard and Smalla (1998), in which no transformants were detected among the 4000 (cultured) isolates screened. However, during the same study, transgenic DNA was detected in the bacterial fraction when DNA was isolated directly from the soil sample. Those authors were not able to determine whether this was due to HGT, or other factors such as soil particle-bound DNA or transiently taken up DNA. In a study by Douville *et al.* (2009), Bt maize transgenic DNA was detected in heterotrophic bacteria in the water and sediment samples downstream of Bt maize fields. Although the authors do not specify the number of bacteria samples tested, they did detect the transgene in bacteria from two of the six sites studied. Douville *et al*. (2009) hypothesised that transformed bacteria may have been the source of the Bt transgene detected in the tissue of freshwater mussels, and that the transgene was detected in mussels from all sites except their reference site, the number of positive results in that study was also quite high.

The positive results among the yeasts and fungi isolates were unexpected, especially the number seen with primer set 5. HGT from plants to fungi has been recorded, but such events are considered extremely rare and ancient (Nikolaidis *et al.*, 2013; Richards *et al.*, 2009). Fungi make a major contribution to the breakdown of leaves in streams, accounting for approximately 15-18% of leaf mass loss (Hieber & Gessner, 2002), and comprising roughly 5-10% of the total carbon of the leaf detritus (Baldy *et al.*, 1995). Fungi are therefore present in large numbers in aquatic systems and would likely be involved in the decomposition of Bt maize detritus in such systems (Hieber & Gessner, 2002, Baldy *et al*., 1995).

4.4.3 Limitations and recommendations for future studies

Although precautions were taken to prevent contamination of samples, contamination cannot be written off as a possible reason for the number of positive results. Contamination from positive control DNA is unlikely, since amplicons of the correct size were generated for primer sets such as set 8 when positive control DNA was used, while this was not the case for the microbial DNA. This does not eliminate cross-contamination between samples, however.

As a culture-based study, there are limitations inherent to this approach which are present. Firstly, very many species of bacteria and fungi are not culturable in the laboratory, and were therefore precluded from this study (Amann *et al.*, 1995; Pereira *et al.*, 2010; Yarza *et al.*, 2014). Likewise, though an effort was made to use a variety of media to isolate a wide range of microorganisms, species which were not culturable under the conditions selected (media, temperature, incubation time etc.) were excluded. Finally, for practical reasons, only a selection of the colonies present after the initial dilution series were selected for further testing.

Future studies should consider a new technique which has become available which could overcome this limitation, namely epicPCR. Emulsion, Paired Isolation and Concatenation PCR (epicPCR) (Spencer *et al.*, 2016) is a high-throughput technique which allows coupling of different genes in uncultured cells, followed by next generation sequencing, allowing detection of both genes at once. The technique is based on similar principles to those used in 454 sequencing library preparation, and uses an emulsion technique in which single cells are isolated within droplets. The droplets are able to contain the genome after cell lysis, but also allow the diffusion of PCR reagents which amplify the target regions and link them by fusion PCR. Following this, high-throughput sequencing is done (these authors used Illumina MiSeq), and the sequences can be processed and identified using the same methods as other metagenomics studies. epicPCR has been applied to link the V4 region of the 16S gene with sulphate reduction gene *dsrB* to characterise which bacteria were involved in sulphate reduction in a lake community (Spencer *et al.*, 2016), to track the spread of a plasmid between members of a biofilm (Cairns *et al.*, 2018) and for studying the distribution of antibiotic resistance genes in a community (Hultman *et al.*, 2018). Applied to a project such as this one, a similar approach would be useful: in-tandem detection of 16S regions and transgene sequences. This would allow screening for transgene sequences in bacterial cells, without the need to culture them first.

From the results of this chapter, it is not possible to confirm the insertion of the fragments into the isolates' genomic DNA, or to comment on the location of the insertion sites and the implications thereof. Detail regarding the genomic context of insertions can be obtained using whole genome sequencing (WGS), which has emerged as a powerful tool to study the composition of genomes,

and has shed light on processes which lead to genomic variation, including polymorphisms, horizontal gene transfer and recombination events (Chewapreecha *et al.*, 2014; Guan & Sung, 2016; Tattini *et al.*, 2015; Trappe *et al.*, 2014; Trappe *et al.*, 2016). Next generation sequencing has seen huge improvements in read lengths (longer than 10kb), which allows better assemblies, more complete genomes and improved understanding of areas of the genome dominated by long repeat sequences (Schmid *et al.*, 2018). As more genomes have been sequenced and studied, HGT, particularly in prokaryotes, has been revealed to be a major evolutionary force (Bansal *et al.*, 2013; Darmon & Leach, 2014; García-Aljaro *et al.*, 2017; Syvanen, 2012). While many of the techniques for detecting and characterizing HGT regions within genomes pre-date the use of high-throughput WGS technologies, methods are rapidly being developed which make use of WGS data directly for this purpose, allowing HGT regions to be studied in the genomic context which they occur (Trappe *et al.*, 2016).

CHAPTER 5

WHOLE GENOME SEQUENCING AND ITS USE IN THE DETECTION OF HORIZONTAL GENE TRANSFER

Abstract

Following the detection of MON810 transgene-associated sequences in the DNA of bacterial isolates in the previous chapter, 10 isolates were selected for whole genome sequencing (WGS). WGS was done using Illumina MiSeq, and the genomes were assembled using the CLC Genomics Workbench. It was hoped that detection of transgene fragments in the assembled genomes of the selected organisms would provide information regarding the genomic context of the insertion sites and any genes which had been interrupted due to recombination with transgene fragments. However, after scrutinising the genomes and the sequencing reads using a mapping based approach (Daisy and a BWA-MEM), traces of transgene DNA could not be detected in the isolates' draft assemblies. This may be due to a lack of sequencing coverage in some areas of the genome, or possibly due to loss of the sequences over time. Though HGT was not detected in this study, a set of methods has been outlined which could be utilised by others.

5.1 Introduction

Whole genome sequencing and comparative analysis has upgraded our understanding of horizontal gene transfer from biological curiosity to evolutionary driving force. As more genomes are sequenced and analysed, the extent of HGT and its importance for evolution have been revealed. This trend has accelerated as whole genome sequencing (WGS) by high-throughput 'next generation' sequencing (NGS) became more widespread, as the technology became more affordable (Koonin & Wolf, 2008; Syvanen, 2012).

The two main approaches for detecting HGT in genome data rely on incongruences in either sequence composition or phylogenetics introduced by HGT. Both are used to detect traces of HGT which have occurred at some point in the organism's evolutionary lineage. Sequence composition methods scrutinise the genome of the recipient organism, and search for regions of sequence which deviate from the average 'signature' composition of that genome in terms of nucleotide composition: GC content, codon usage, or k-mer frequency etc., indicating that they may be influenced by HGT (Cortez *et al.*, 2009; Langille & Brinkman, 2009; Ravenhall *et al.*, 2015b). Phylogenetic methods examine the phylogenies of genes compared to the species' phylogeny, and identify conflicts which may be due to HGT (Langille & Brinkman, 2009; Ravenhall *et al.*, 2015b; Vernikos & Parkhill, 2006). Another approach is to treat horizontally transferred

sequences as structural variants, and use NGS data directly to detect instances of HGT (Trappe *et al.*, 2016).

Structural variants (SVs) are genomic variations which are longer than 50 bp in length, while indels account for variations shorter than 50 bp (Alkan *et al.*, 2011; Tattini *et al.*, 2015). Unlike HGT where variation is introduced due to the DNA of another organism, the variation represented by SVs originates due to rearrangements within the genome itself. SVs encompass deletions, insertions, inversions, duplications and copy number variations, and translocations (including inter-chromosomal translocations in multi-chromosomal organisms). SVs can be detected in mapping inconsistencies between reference genomes and the available reads of the query genome (Tattini *et al* 2015).

There are 4 main approaches to using NGS data to detect SVs:

- Read count: the distribution of reads mapped to a suitable reference genome is assumed to be proportional to the number of copies present in the genome. If the read depth (coverage) of a region is significantly divergent from that of the rest of the background levels of the rest of the genome, this may denote a copy number variation (Alkan *et al.*, 2011; Tattini *et al.*, 2015). A dearth of reads in an area reflects a loss, while increased coverage reflects a gain (see Figure 5.1, column RC).
- Read pair: Sequencing reads which are part of read pairs or mate pairs are expected to map to certain positions relative to each other when mapped to a reference genome. Discordances in the mapping positions and orientation of paired reads (see Figure 5.1, column RP) are indicative of possible SVs (Alkan *et al.*, 2011; Trappe *et al.*, 2016).
- Split-reads: Also taking advantage of mate and read pair relative positioning, split-read algorithms use a one-end anchored mapping approach: when one read maps to the reference genome but the other maps partially or imprecisely because it spans a breakpoint where an SV is present (see Figure 5.1, column SR). These methods can provide single-base resolution (Karakoc *et al.*, 2012; Tattini *et al.*, 2015; Trappe *et al.*, 2016).

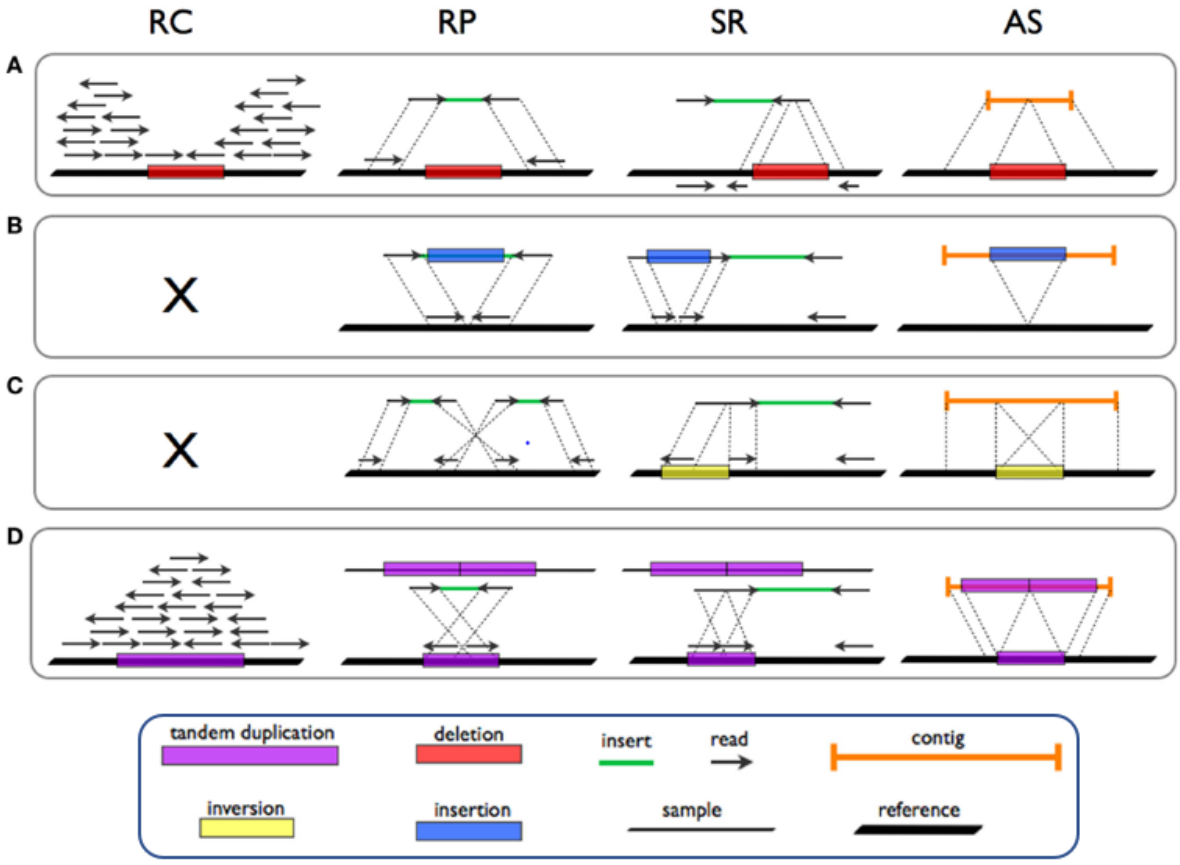


Figure 5.1: **Classes of structural variants and their detection with NGS data. A: deletion, B: insertion, C: inversion, D: duplication. RC: read count, RP: read pair, SR: split-reads, AS: assembly methods. Originally published in Tattini *et al*. (2015) in Frontiers in Bioengineering and Biotechnology (creative commons attribution licence), used with permission of the authors.**

- Assembly-based tools: Instead of mapping reads to a reference genome, assembly-based methods assemble the genomes de novo, and collate and compare information by analysing multiple genomes simultaneously, to detect SVs (Iqbal *et al.*, 2012; Nijkamp *et al.*, 2012).

Similar disparities between reads and reference genomes are created by insertions generated through HGT. Trappe *et al.* (2016) have exploited this fact to develop Daisy, a pipeline of tools which takes advantage of methods of analysing NGS data developed for SV detection, and applies them to HGT detection. Unlike phenotypic and compositional methods, which analyse the content of the genome, Daisy utilises the characteristics of the NGS data directly. Daisy maps reads to both a reference genome and a donor genome, then performs a combination of split-read mapping followed by analysis of coverage and read pair information to identify regions where

insertion due to HGT has taken place. Daisy is, to the best of our knowledge, the first HGT detection tool to utilise a mapping-based approach.

Aligning (mapping) reads to a genome is often used in genome assembly, where a reference genome is used to guide the positioning of NGS reads, based on sequence similarity between the reads and the reference genome. It has also been used to detect recombinant sequence segments in transformation experiments. Transformants were sequenced and the reads aligned with a reference genome, after which single nucleotide polymorphisms (SNPs) were identified. The presence of donor polymorphisms (particularly several them clustered close together) in the transformant sequences indicated that those regions had undergone homologous recombination with segments of donor DNA (Croucher *et al.*, 2012; Fani *et al.*, 2014; Power *et al.*, 2012).

Alignment of reads can also be used to locate particular sequences in a genome, by matching a known sequence to the reference genome in question. Allowances for non-matching bases within the query sequence can be made by adjusting the gap penalties, which include penalties to the alignment score for opening a gap and extending a gap (Li, 2013; Li & Durbin, 2009). This means that a query sequence need not match the reference genome exactly in order to find a region to align to. A minimum number of identical bases can be set to ensure that a certain degree of similarity is required for a hit to be reported. The BWA collection of algorithms is frequently used for mapping low divergent reads to a larger reference sequence or genome. In this study, the BWA MEM algorithm was used. This algorithm was designed to cope with reads of 70 bp and much longer, compared to other BWA algorithms which were designed when sequencing reads were only about 36 bp long, meaning it could handle the Illumina reads which were mostly in the region of 100 – 300 bp. It makes use of a seed-and-extend method, and is capable of automatically selecting between local and end-to-end alignment (Li, 2013). Instead of assembling the reads to a reference genome, since assembly was done de novo, the BWA MEM algorithm was used to match sequences in the isolate genomes and the MON810 transgene. The inverse was also done: MON810 sequences (chopped up to resemble sequencing reads) were also aligned to the draft assemblies of the selected isolates.

In the previous chapter, the results suggested that transgene sequences may be present in the genomes of some of the aquatic microorganisms sampled. The regions detected were quite short, suggesting that the whole transgene was not transferred. Whole genome sequencing of a set of bacterial genomes was done to confirm whether transgene sequences were present, and if so, to characterise the site of insertion.

5.2 Materials and methods

5.2.1 Whole genome sequencing

The culturing of the bacteria and DNA isolation methods are described in more detail in chapter 4. Briefly, cells from overnight cultures grown in nutrient broth or onto nutrient/R2A agar were harvested by either centrifugation or scraping from the plate. DNA isolation was done according to the manufacturer's instructions, using the NucleoSpin® Tissue Kit (Macherey-Nagel, Germany). Quality assessment by spectrophotometric analysis (NanoDrop™ 1000 Spectrophotometer (Thermo Fischer Scientific, US)) and electrophoresis followed to ensure the isolated DNA met the quality requirements for next generation sequencing. These entailed a 260:280 ratio of 1.7 - 2.0, and high molecular weight, without RNA contamination. Final assessment was done with a Qubit 2.0 (Thermo Fischer Scientific, US).

Library preparation was done using the Nextera Library Prep Kit (Illumina, city, USA), according to the manufacturer's instructions. Paired-end sequencing took place on a MiSeq sequencer (Illumina), using V3 Illumina chemistry. Both library preparation and sequencing were done in-house by Dr. Charlotte Mienie.

5.2.2 Assembly

The CLC Genomics Workbench version 9.0 (CLC bio, USA) was used to process and assemble the reads. Quality assessment was done, followed by quality trimming according to the following parameters. After trimming, a variety of settings were tested to find the combination which yielded the best assembly. These included iterations where pairs were merged before assembly (including a trim step after merging), assembly without pair merging, and assemblies in which word size, bubble size and minimum contig size were adjusted.

To assess the quality of the assemblies, FASTA files of the assemblies were uploaded onto the QUAST online tool (Gurevich *et al.*, 2013). This allowed direct comparison of draft assemblies across a range of parameters. Choosing the 'best' assembly is a matter of subjectivity, and in this case the parameters which carried the most weight for the selection of the assembly for each isolate were the total length, the number of contigs, and the number of ambiguous bases (#Ns) present. An alternative to comparing these parameters is to assess genome completeness based on the presence or absence of single copy genes which are expected to be present. BUSCO 3.0.1 (Benchmarking Universal Single Copy Orthologs) was used for this purpose (Simão *et al.*, 2015; Waterhouse *et al.*, 2018), assessing the assemblies using different lineage sets. Each lineage set is compiled of a list of single-copy orthologs which are expected to be found in

representatives of that lineage. In this case, two generalised Bacteria lineage sets, and a more specific lineage set (Gammaproteobacteria, Firmicutes, Betaproteobacteria and Actinobacteria), were used according to the identity of the isolates.

The isolates were identified using both BLAST (NCBI, 2014) and specI (Mende *et al.*, 2013). specI uses a set of 40 universal single copy phylogenetic marker genes (which also form the basis of one of the generalised bacteria lineages used in the BUSCO assessment), and a threshold of 96.5% sequence identity to make species-level identifications. Contig ordering was done using the Mauve Contig Mover, part of the Mauve Aligner (Darling *et al.*, 2004, 2010). The draft assemblies were annotated using the RAST (Rapid Annotation using Subsystem Technology) prokaryote genome annotation service (Aziz *et al.*, 2008; Overbeek *et al.*, 2014). This provided information on the composition of the genome and the functions of various genes.

5.2.3 Transgene construct detection strategies

5.2.3.1 Simulation

To ensure that the detection strategies described below would be capable of detecting the presence of horizontally transferred transgene fragments, should they occur in the genomes, a simulation of such a genome was created using HgtSim (Song *et al.*, 2017). HgtSim was developed to test out pipelines that detect incidences of HGT in microbial genomes, particularly metagenomes. Given the input of a recipient genome and the desired insert sequences, the HgtSim algorithm simulates HGT of the inserts, and includes the ability to add flanking sequences and manipulate the mutation level. Here, the A29 isolate assembly was used as the recipient genome (it was judged to be one of the most complete assemblies, see parameters in results below), and portions of the MON810 transgene (specifically the cry1Ab and hsp70 regions) were used as inserts, creating a genome in which transgene sequences were guaranteed to be found. HgtSim provides details on where in the genome the sequences have been inserted, and these were checked in BioEdit to ensure the specified transgene sequences were indeed present.

5.2.3.2 Daisy

Daisy requires a reference genome, a donor genome, and the paired WGS reads of the query genome. A reference genome was selected for each isolate after the draft assemblies were constructed and identified. The MON810 transgene construct (accession number AY326434.1, Hernández *et al.* (2003)) was used as the ‘donor genome’. The paired reads of each isolate underwent quality control with Trimmomatic (V0.32, Bolger *et al.* (2014)) to remove short and/or

low quality reads and adapter sequences. These trimmed reads (both forward and reverse) were then entered the Daisy pipeline. Wgsim (Li, 2011) was used to create sequence “reads” from the HGT-simulation genome assembly, essentially un-assembling it, resulting in two fastq files mimicking paired end sequence files which were used as input reads for Daisy. Daisy was run for each isolate on both default parameters and with a reduced minimum insert length of 50 bp.

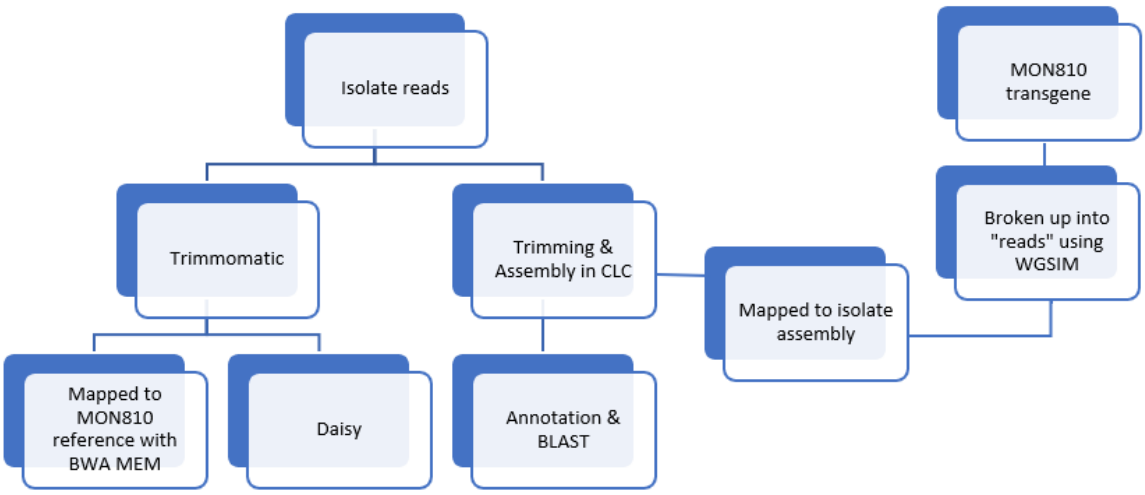


Figure 5.2: Summary of the methods used to assemble the genomes and of the selected isolates and search them for transgene sequences

5.2.3.3 BWA MEM

Using BWA MEM (Li, 2013), the trimmed sequencing reads of the isolates were aligned to the MON810 transgene sequence, and transgenic sequences were aligned to the assembled genomes of the isolates. The raw sequencing reads of the isolates were trimmed as described in the previous section. Wgsim (Li, 2011) was used to simulate sequence reads from the MON810 transgene sequence, creating two fastq files (mimicking paired end sequence files) each containing 10 000 ‘reads’ 150 bp in length. The base error rate was set to zero to prevent Wgsim from simulating many sequencing errors/SNPs etc. When the isolate reads were used as input for BWA MEM, the MON810 transgene sequence acted as a reference genome. When the MON810 ‘reads’ were used as input, the assembled genomes of the isolates were used as reference genomes. The BWA MEM algorithm was run using various parameters, including reducing the minimum seed length from its default of 19 bp and reducing the mismatch penalty and gap open penalties from their respective values of 4 and 6 to 0. The resulting SAM files were inspected for matches, and were converted to BAM files, sorted and indexed using SAMtools (version 0.1.19-44428cd) (Li *et al.* (2009)). Alignments were visualised in the Integrative

Genomics Viewer (IGV, version 2.4.3) (Robinson *et al.* 2011; Thorvaldsdóttir *et al.* 2013) and BioEdit (version 7.0.5.3) (Hall 2011, 1999).

5.3 Results

5.3.1 Assembly results

The parameters of the selected draft assemblies were assessed by Quast, and are presented in Table 5.1 below. The completeness of the genomes, in terms of the presence of expected genes, was assessed by BUSCO, though RAST also provides an estimation of the predicted number of coding sequences, RNA sequences, and missing genes. The different lineages used for the BUSCO assessments provided slightly different results due to the different ortholog sets used by each. In the case of E10 (*Serratia* sp.), for example, for the measurement of completeness to be 100% according to the pMG lineage set, but be missing 11 genes (92% complete) when assessment was done with the Bacteria lineage set. Taken together, these results show that while each of the genomes is missing some genes and there are some gaps in the genomes which are not covered, they are relatively complete (see Table 5.1). For the purpose of comparison, some of the reference genomes were assessed using the same parameters. Figure 5.3 shows the results of this comparison between each draft assembly and its respective reference genome. The completeness of the draft genomes of this study compare favourably with reference genomes (such as the *A. veronii* RefSeq genome) maintained by the NCBI database, and somewhat better than some of the other assemblies in the database (see especially *Arthrobacter* sp. Hiyo4, which is only 58% complete). This serves to demonstrate that the draft assemblies are of relatively good quality.

Table 5.1: Table summarising assembly quality parameters according to different tools

		Isolates									
		A29	**A52**	**A55**	**B19**	**B52**	**C15**	**C41**	**D15**	**E2**	**E10**
Quast	**Size (bp)**	4482053	3111393	4497612	4472126	4650301	3450154	5401332	6632104	587382	5637698
	Contigs	56	50	120	171	89	76	88	453	284	419
	N50	165894	2801210	67035	56190	123944	90246	140955	35063	45670	29601
	GC content	58.84%	53.74%	58.97%	65.96%	58.75%	65.93%	62.52%	62.34%	65.94%	53.75%
	Coverage										
RAST	**Predicted CDS**	4021	3124	4082	4024	4184	3138	4924	6029	5230	4986
	Predicted RNAs	111	81	109	60	111	56	63	65	70	68
	Predicted missing genes	46	7	50	72	45	28	66	81	92	88
BUSCO	**Bacteria Lineage**	138 (93.2%)	138 (93%)	138 (93%)	139 (93.9%)	137 (92.6%)	(88.5%)	139 (93.9%)	147 (99.3%)	138 (93.2%)	136 (91.9%)
	Duplicated	0	1 (0.7%)	0	0	0	0	0	0	0	0
	Missing	10 (6.8%)	8 (5.4%)	10 (6.8%)	9 (6.1%)	10 (6.7%)	17 (11.5%)	9 (6.1%)	1 (0.7%)	9 (6.1%)	11 (7.4%)
	fragmented	0	1 (0.7%)	0	0	1 (0.7%)	0	0	0	1 (0.7%)	1 (0.7%)
	pMG	40 (100%)	39 (97.5%)	40 (100%)	38 (95%)	40 (100%)	40 (100%)	39 (97.5%)	38 (95%)	39 (97.5%)	40 (100%)
	Lineage										
	Missing	0	1 (2.5%)	0	1 (2.5%)	0	0	1 (2.5%)	2 (5%)	0	0
	fragmented	0	0	0	1 (2.5%)	0	0	0	0	1 (2.5%)	0
	More specific Lineage	Gamma-proteo-bacteria (452)	Firmicutes (232)	Gamma-proteo-bacteria (452)	Gamma-proteo-bacteria (452)	Gamma-proteo-bacteria (452)	Actino-bacteria (352)	Gamma-proteo-bacteria (452)	Gamma-proteo-bacteria (452)	Betaproteo-bacteria (582)	Gamma-proteo-bacteria (452)
	Complete-ness	448 (99.1%)	225 (97%)	446 (98.7%)	440 (97.3%)	446 (98.7%)	337 (95.7%)	445 (98.2%)	441 (97.3%)	562 (96.6%)	446 (98.7%)
	Duplicated	0	0	0	0	0	0	1 (0.2%)	1 (0.2%)	1 (0.2%)	0
	Missing	4 (0.9%)	2 (0.9%)	1 (0.2%)	3 (0.7%)	3 (0.7 %)	4 (1.1%)	7 (1.6%)	7 (1.6%)	9 (1.5%)	2 (0.4%)
	fragmented	0	5 (2.1%)	5 (1.1%)	9 (2%)	3 (0.7%)	11 (3.2%)	0	4 (0.9%)	10 (1.7%)	4 (0.9%)

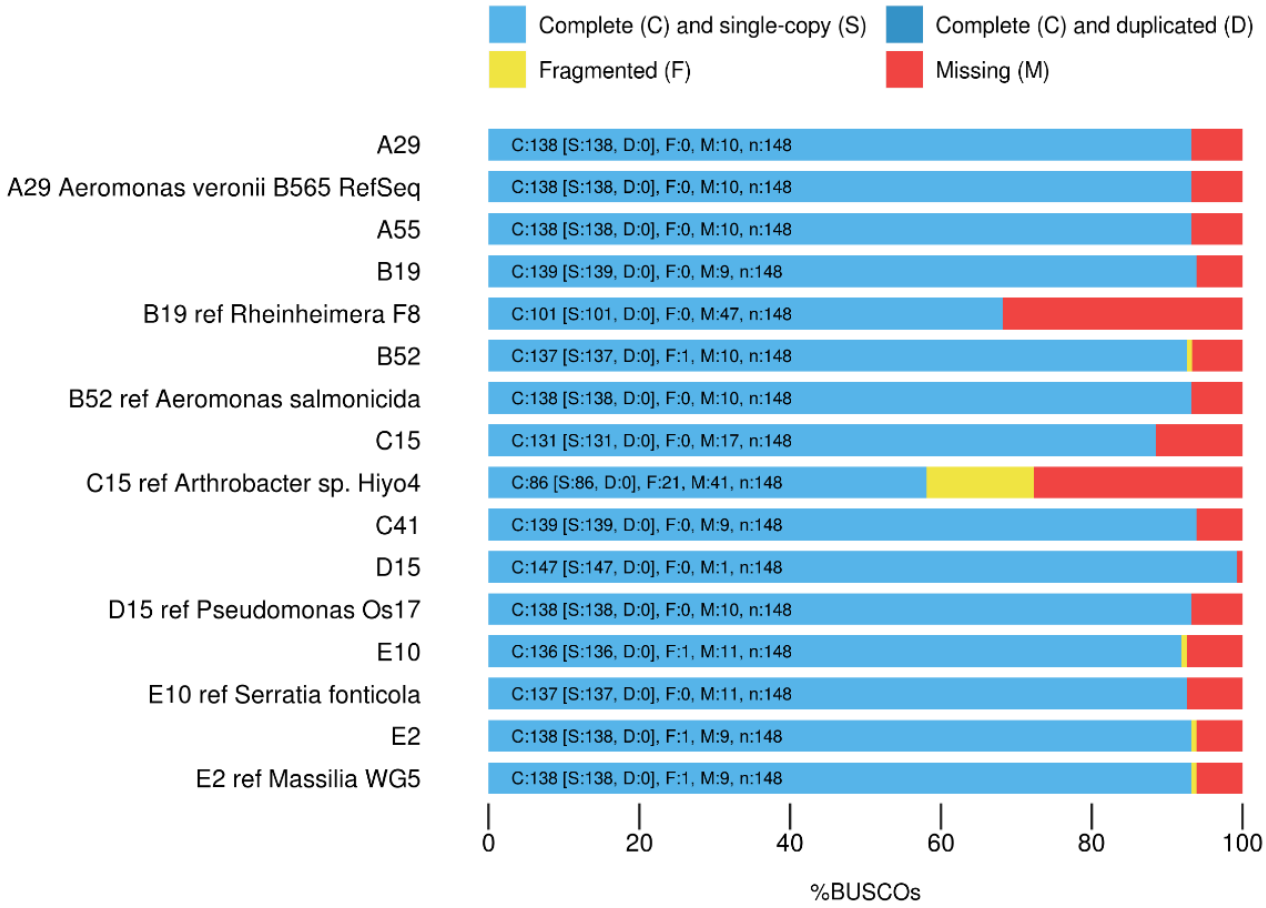


Figure 5.3: A comparison of the completeness of the assemblies from this study with and their closest matches in the NCBI database, as well as a Refseq reference genome.

Identification of the isolates was done using the whole assembly (see Table 5.2 below). For specI, which makes use of a list of 40 phylogenetic marker genes (the same list as the pMG lineage of BUSCO), only 3 assemblies were able to pass the 96.5% threshold required for confident taxonomic classification: A29, A55 (both identified as *Aeromonas veronii*) and B52 (*Aeromonas salmonicida)*. These assemblies also had high query coverage and percentage identity when entered into the BLAST database, which meant that they were easily identified, and suitable reference genomes were selected with no trouble. In the cases of C15, C41, D15, and E10, though the specI identifications did not pass the threshold (they were above 90%), their BLAST scores were acceptable, and matched the specI identification at genus if not at species level. B19 had low scores for both types of identification, though they were in agreement with each other. E2 also had low scores for both identification methods, which placed it in the same taxonomic groups up to family level, but placed it in different genera. Reference genomes for these assemblies were selected according to the BLAST results. A52 presented more of a challenge, since it did not have close matches in either database. In fact, specI suggested that A52 may

represent a new undescribed genus, though it is outside of the scope of this study to investigate this claim further.

5.3.1 Transgene detection

5.3.1.1 Daisy

Simulation genome

Daisy successfully identified the transgene inserts in the HGT-simulation genome. The final HGT evaluation file provided information on the coordinates of the inserts as they corresponded to their location in the acceptor and donor genomes. The coordinates of the donor genome (MON810) matched the location of the inserts which were used to create the simulation genome.

Daisy was unable to detect regions of MON810 transgene sequence in any of the isolates, using either default parameters or when adjusted for shorter insert size. In some cases, Daisy identified candidate regions where split-read SVs were detected in the genomes, though none of these were associated with insertion sequences of the MON810 donor genome and did not pass the remaining steps and quality control measures.

Table 5.2: Table summarising identification of isolates according to different databases

	Isolates									
	A29	**A52**	**A55**	**B19**	**B52**	**C15**	**C41**	**D15**	**E2**	**E10**
specI Identification	*Aeromonas veronii*	Could not be assigned	*Aeromonas veronii*	*Rheinheimera* sp. A13L	*Aeromonas salmonicida*	*Arthrobacter chlorophenolicus*	*Pseudomonas mendocina*	*Pseudomonas protegens*	*Herbaspirillum seropedicae*	*Serratia odorifera*
Average % Identity	98.3%	NA	98.2%	92.7%	97.9%	81.6%	92.9%	96.2%	82.7%	91.4%
Species cut-off	Above cut-off	NA	Above cut-off	Below cut-off	Above cut-off	Below cut-off	Below cut-off	Below cut-off	Below cut-off	Below cut-off
BLAST ID	*Aeromonas veronii* strain CB51	*Bacillus sp. OxB-1*	*Aeromonas veronii* strain CB51	*Rheinheimera* sp. F8	*Aeromonas salmonicida* strain A527	*Arthrobacter* sp. Hiyo4	*Pseudomonas mendocina* S5.2	*Pseudomonas* sp. Os17	*Massilia* sp. WG5	*Serratia fonticola* strain GS2
Accession	CP015448.1	AP013294.1	CP015448.1	CP013656.1	CP022550.1	AP014718.1	CP013124.1	AP014627.1	CP012640.1	CP013913.1
Query coverage	89%	4%	89%	12%	90%	16%	77%	88%	52%	87%
% Identity	97%	78%	98%	85%	98%	91%	89%	92%	87%	99%
E value	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

5.3.1.2 BWA MEM

No exact matches along the entire length of a read were found with either the alignments of isolate reads to MON810 transgene reference, or MON810 transgene sequences to isolate assemblies. Shorter matches were found, which concentrated around stretches of ±19 bp on some contigs (see Figure 5.4 below).

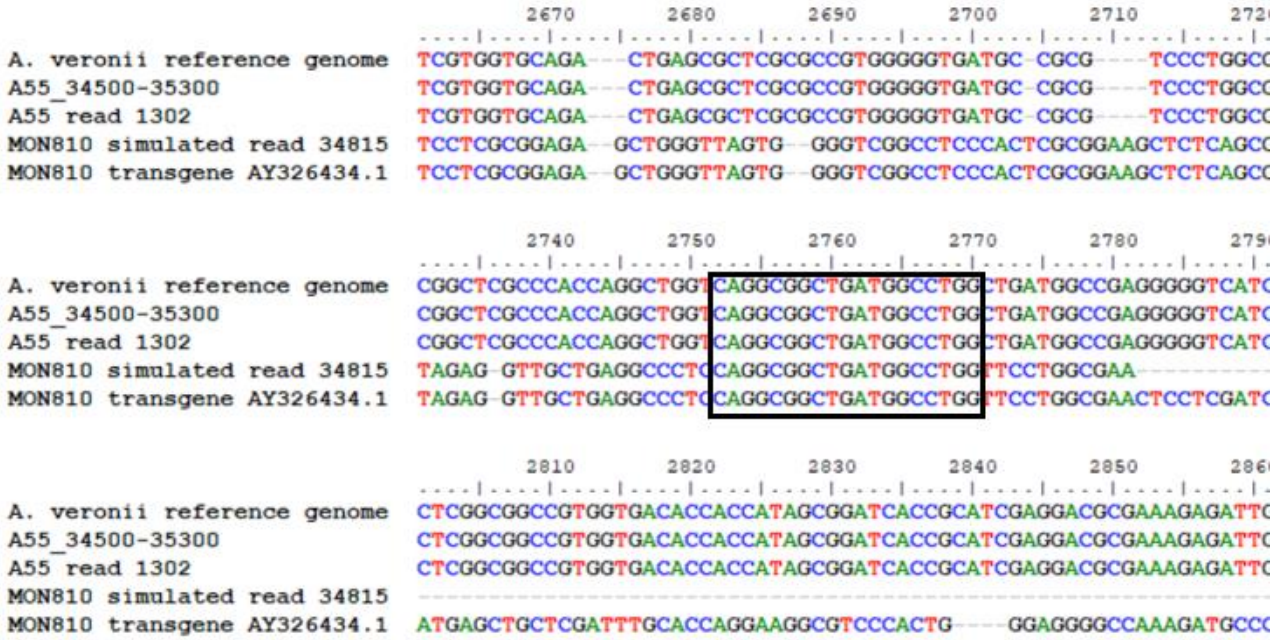


Figure 5.4: An alignment of the reference genome (top line), a region of the A55 assembly (contig 10, region 34500-35300, second line), a MON810 simulated read (third line), and the MON810 transgene sequence (bottom line). The border surrounds a 19bp region which is found in all 4 sequences.

In the Figure above, isolate A55 serves as an example of the results of the BWA MEM alignments. A number of MON810 transgene 'reads' were reported as matches with the A55 assembly, with all of these matches concentrated on contig 44, between position 34 500 and 35 300. When A55 reads were aligned with the MON810 transgene, several reads were reported as matches with the transgene. These reads form part of the A55 assembly in the same region that the MON810 sequences mapped to. However, this region in the A55 assembly also matches more closely to *Aeromonas veronii* reference genomes, as shown in Figure 5.4. A 19 bp region which is common to the A55 assembly, an *A. veronii* reference genome and the MON810 transgene sequence has

been highlighted in the black block in Figure 5.4. Although these regions fulfil the minimum seed length of the BWA algorithm, the relaxed gap penalties were probably responsible for these regions attaining alignment scores high enough to be reported as hits. The region of the assembly in question was identified by RAST as belonging to the Spermidine Putrescine ABC transporter permease component PotB (TC 3.A.1.11.1) gene, a membrane transporter gene found in numerous bacteria.

5.4 Discussion

Previous studies which have searched for historical HGT in environmental isolates before have used methods based on composition patterns and phylogeny, as discussed in the introduction. Many cases were found of genes whose phylogeny or composition did not match that of the rest of the genome, indicative of a different lineage (Ravenhall *et al.*, 2015). Transfer events which have occurred in laboratory settings have been detected using whole genome sequencing to study recombination resulting from transformation, by searching for single nucleotide variants associated with the donor strains in the genomes of transformants (Fani *et al.*, 2014; Maddamsetti & Lenski, 2018; Power *et al.*, 2012).

The purpose of this study was to search for specific, recent HGT events of known sequences, instead of studying the evolutionary history of genomes for all the HGT events in their pasts. To do this, mapping-based approaches which utilised WGS data were used. The first approach used Daisy (Trappe *et al.*, 2016). The detection of the transgene inserts in the HGT-simulation genome indicate that the pipeline and its mapping-based approach can be used to detect specific HGT events. However, no transgene sequences could be detected in the isolates genomes using Daisy. In some cases, Daisy identified candidate regions where split-read SVs were detected in the genomes, though since none of these were associated with transgene sequences and did not pass the remaining steps and quality control measures. These more likely reflect genomic rearrangements and differences between the reference genome and the isolates' genomes. To the best of my knowledge, this is the first time that such a mapping-based HGT detection method has been used on whole genomes of environmental isolates.

A challenge experienced in implementing Daisy was in finding suitable reference genomes for the environmental isolates in this study. For isolate A52 no close reference genome could be found. While reference genomes could be found for isolates B19, C15, and E2, closer matches could be desired. Though the number of draft assemblies in databases such as NCBI and EMBL continue to increase at an impressive rate, close matches for many microbial isolates of environmental origin are not always available. This may have made SVs, particularly split reads, more difficult to

detect, since it is likely that there were regions of the genome where none of the isolate reads mapped. In the case of the present study, the second method confirms the absence of the transgene fragments, regardless of the suitability of the reference genomes.

The BWA MEM approach was similar to the initial steps followed by Croucher *et al.* (2012); Maddamsetti and Lenski (2018); Power *et al.* (2012) among others, whereby isolates underwent whole genome sequencing and the resulting reads were aligned with a reference in order to detect variations associated with recombination. These authors proceeded to use variant analyses to detect mostly nucleotide polymorphisms and mutations, whereas the aim of this study was to determine the insertion of specific fragments. BWA MEM was used to map reads to regions of similar sequence in the assembly or reference genome. However, this approach was also unsuccessful in detecting transgene sequences in the isolates' genome assemblies. Mapping the isolate reads to the MON810 sequence failed too: none of the isolate reads bore enough sequence similarity to the transgenic sequence (beyond the extremely short regions shown in Figure 5.2) to map to it, even though the stringency of the gap and mismatch penalties was lowered. The 19 bp regions of similarity may be sufficient to qualify as regions of microhomology which could facilitate HGT (Harms *et al.*, 2016; Kohli *et al.*, 1999), though they have not done so here and would be considered a rare occurrence.

These results do not agree with those of the previous chapter. In examining why that is, the detection methods used, and the quality of the whole genome sequencing and resulting assemblies were considered. The success of the detection methods in detecting the transgene sequences inserted into the HGT simulation genome indicates that, had transgene sequences been present in the isolate assemblies, they would likely have been detected. The BWA MEM mapping method was able to detect regions of sequence similarity as short as 19 bp, indicating that even short insertions of 50 – 100 bp, which are on the outside of Daisy's capabilities, would have been mapped and detected using the BWA MEM approach if they had been present. We conclude that the transgene sequences are not present in the assemblies of the selected isolates.

That brings us to the assemblies themselves, and the quality thereof – in other words, what is the likelihood that transgene sequences *were* present in the organisms, but did not make it into the assemblies? The assessments with Quast, RAST and BUSCO suggest that the assemblies are fairly complete, and compare well with the completeness of the reference genomes downloaded from the NCBI database (see Figure 5.3). The annotations of the assemblies also show that, in each case, thousands of genes could be identified. There are some genes missing from each genome though, so it is possible that coverage was not complete, and that the regions where transgene sequences were inserted were missed.

The next thing to consider is the results of the previous chapter themselves. What is the likelihood that the transgene sequences were never present in these isolates, and the positive results are due to contamination or PCR artefacts, or that the fragments could have been lost over time? Every care was taken to avoid contamination, and the absence of positive results in the negative controls and hundreds of other isolates is testament to that care. Some of the amplicons from the PCR were also sequenced and aligned with the MON810 transgenes, suggesting that PCR artefacts were not the cause of the positive results. However, there was a large time gap between the PCR tests and the whole genome sequencing, which involved several additional passaging steps as well as a fresh DNA isolation before WGS was done. Loss of recently transferred DNA, particularly if it is not useful to the cell at the time, has been known to occur (Hao & Golding, 2006; Kuo & Ochman, 2009) and may explain the differences in the results.

Most other studies looking for recent transformation events have generated them themselves in the lab, and not in environmental isolates. This approach has the benefits of having the parent/reference strains on hand. Other studies which have studied such recent transfer events have differed from this one in several ways: they have worked with laboratory strains of bacteria (with the parent genome available), and made use of screening procedures such as restoration of antibiotic resistance to detect HGT.

CHAPTER 6

CONCLUSION AND RECOMMENDATIONS

6.1 Conclusion and recommendations

The aim of this project was to determine whether transgene fragments associated with genetically modified maize can be detected in DNA isolated from aquatic organisms, and to annotate sites of insertion if detected, in order to address knowledge gaps related to the interaction between Bt crops and aquatic organisms. Though the latter part of this goal was not realised, the former was. This study shows how transgene sequences can be used as part of monitoring the interaction of Bt crops with their environment, and how DNA-based methods can be used in aquatic ecosystem monitoring more generally, to supplement taxonomy-based monitoring methods. Methods of using whole genome sequencing as a tool to investigate recent horizontal gene transfer in environmental isolates have been explored. This is the first study, to my knowledge, which has used whole genome sequencing and whole genome analysis to attempt detection of HGT of transgenic DNA fragments to environmental isolates. Though transgene fragments were not detected in the isolates' genomes, a set of methods has been outlined which could be utilised by others. High-throughput sequencing will continue to play a vital role in illuminating the intricacies of horizontal gene transfer in different environments and contexts. Knowledge has been collected and collated which fills in part of the gap which persists even after two decades of growing Bt crops. Genetically modified crop development may be progressing beyond single transgene crops, but aquatic ecosystems still represent a blind spot in terms of the ecological context these crops find themselves in. Several key points can be distilled from this thesis:

6.1.1 Aquatic ecosystems are understudied in the context of GM crops

By examining the research done so far, both in South Africa and internationally, it was found that aquatic ecosystems had, to a large degree, been neglected for the first decade of commercial growing of Bt crops (Carstens *et al.*, 2012; Pott *et al.*, 2018; Rosi-Marshall *et al.*, 2007). Although there has been an increase in the amount of studies done since 2007, aquatic ecosystems continue to be left out in many cases. Case in point, a recent review and meta-analysis published in Nature (Pellegrino *et al.*, 2018), which considered the on target and off-target effects of GM maize over 21 years, did not include a single aquatic species in its reckoning. The research which has been done so far has also tended to focus on a limited number of transgenic crops and events, particularly MON810 maize (which is also the transgenic event considered in this study), with numerous others receiving little attention. In chapters 2 and 3, the interaction of Bt crop

material with non-target aquatic organisms was considered. Transgene DNA sequences were detected in several organisms and demonstrated that aquatic invertebrates in the Vaalharts Irrigation Scheme are exposed to Bt plant material. Determining whether there was any detrimental effect as a result of this exposure was outside of the scope of this study, ecotoxicological testing is recommended as a follow-up step. The scale at which Bt crops are currently grown, and at which they are projected to continue to be grown, makes it necessary to consider the implications for the surrounding ecosystems.

6.1.2 Changes in patterns of use of GM crops require changes to risk assessment

The first generation of single transgene crops, are steadily being replaced with stacked events (James, 2016). Projections indicate that stacked events will be the most common form of GM crops through to the 2020s (Parisi *et al.*, 2016). Single transgene events such as MON810 are being replaced with an expanding repertoire of mix-and-match Bt transgenes and herbicide tolerance genes, and an increasing number of other traits such as drought tolerance and modified oil composition are also making their way through the regulatory channels towards commercialization. Single transgene Bt crops were among the pioneers of the GM crop movement, and they have exposed gaps in risk assessment, regulatory systems and environmental monitoring (Carstens *et al.*, 2012), but these crops are being phased out. What has been learned from single transgene crops in terms of monitoring and risk assessment should be applied to the broader context of continued interaction between aquatic ecosystems, agriculture, and the next generations of genetically modified crops. A starting point is the need to investigate combinatorial effects of transgene products, including how they interact with other agrichemicals. Bt toxins are used in combination with each other, as well as herbicide co-technologies and other pesticides such as neonicotinoids (Alford & Krupke, 2018; Huseth *et al.*, 2018). More than that, it is necessary to consider the multiple stressors from high input agricultural systems which are thrown together in the aquatic environment (Alexander *et al.*, 2016; Chapman, 2018; Posthuma *et al.*, 2016; Pott *et al.*, 2018; Schäfer *et al.*, 2016). Examining the potential effects of multiple stressors and inputs is in line with current thinking of aquatic ecosystem monitoring and management, even outside of the context of GM crops. It is important to understand the implications of pest management strategies within their ecological context.

6.1.3 DNA-based methods are useful for monitoring aquatic organisms and ecosystems

In Chapters 3 and 4 of this study, the presence of transgene DNA in the aquatic ecosystem was investigated for two purposes: use of transgene DNA as a surrogate to imply the presence of Bt plant material; and to consider the implications of the release of transgene DNA in terms of HGT,

particularly to aquatic microorganisms. In agreement with studies such as Douville *et al.* (2009), it was found that PCR-based methods were suitable for tracing transgene DNA in various compartments. Furthermore, DNA-based monitoring, when linked to barcoding and meta-barcoding, is strongly advocated for as a method of non-invasive monitoring which can augment current strategies for the improvement of aquatic ecosystem monitoring overall, not just for the study of GM crops.

6.1.4 Use of new technology to study HGT in an environmental context

Following the detection of transgene sequences using a PCR-based approach in Chapter 4, a confirmation step – whole genome sequencing – was put forward in Chapter 5. According to the metrics and parameters used, the draft assemblies of these genomes were of good quality and compared favourably with genomes available in databases such as GenBank (NCBI, 2014). Nevertheless, there were some regions missing (compared to the available reference genomes), so improved sequencing techniques, such as those developed by Oxford Nanopore Techonologies which generate very long reads, could be utilised in future studies to improve sequencing across repetitive regions and reduce fragmented assemblies (Schmid *et al.*, 2018). Though the transgene sequences could not be detected in the whole genomes, a workflow has been demonstrated which could be applied to similar studies in future, in particular the use of modified structural variant detection techniques for the study of short fragment uptake.

6.1.5 Understanding HGT as an implication of releasing transgenic DNA

The results from Chapter 4 suggested that short fragments were responsible for the positive results seen. The failure to detect the fragments in the whole genome assemblies – and the subsequent inability to describe the insertion sites – means these results are inconclusive. Outside of this study, evidence has been accumulating which shows the significance of short fragment recombination in bacterial evolution, including in development of antibiotic resistance (Croucher, Harris *et al.*, 2012; Croucher, Mostowy *et al.*, 2016; Domingues *et al.*, 2012b; Fani *et al.*, 2014; Harms *et al.*, 2016; Maddamsetti & Lenski, 2018; Overballe-Petersen *et al.*, 2013). Recent EFSA guidelines (EFSA, 2015; Gennaro *et al.*, 2017) require that bioinformatics sequence similarity searches should be done to detect regions of sequence similarity (> 200 bp of continuous matching sequence) between the transgene and microbial sequences which might facilitate HGT via homologous recombination. This is an indication that the possibility of HGT is being taken seriously, particularly in the case of marker genes which could contribute to the spread of antibiotic resistance. Most studies of short fragment transfer have mostly been done in lab settings, due to the need to control and in some cases optimise, the conditions. However, to

understand how this process operates under real-world circumstances, experiments which more closely resemble environmental conditions should be explored (Overballe-Petersen *et al.*, 2013; Overballe-Petersen & Willerslev, 2014).

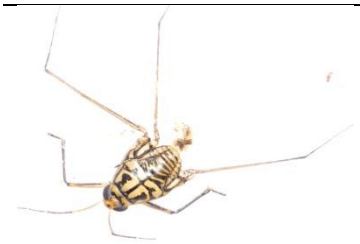
APPENDIX

Appendix I: Table of collected macroinvertebrate specimens

Table A.1 Photographs of collected macroinvertebrate specimens and their corresponding DNA isolation codes (all photos taken by HJ Venter)

Specimen photograph	Classification/Identified as		DNA isolate codes
	Family	**Genus**	
	Aeshnidae	*Anax*	D74-11/04/08mi
	Atyidae	*Caridina*	B84-11/04/08mi B85-11/04/08mi B86-11/04/08mi B87-11/04/08mi C106-11/04/08mi D77-11/04/08mi D78-11/04/08mi D79-11/04/08mi D80-11/04/08mi
	Baetidae		A92-11/04/08mi C103-11/04/08mi F55-11/06/22mi
	Belastomatidae	*Appasus*	A87-11/04/08mi A88-11/04/08mi A89-11/04/08mi A96-11/04/08mi C100-11/04/08mi D65-11/04/08mi D66-11/04/08mi D67-11/04/08mi

			D68-11/04/08mi D69-11/04/08mi D70-11/04/08mi D71-11/04/08mi D72-11/04/08mi D75-11/04/08mi
No photo available	Ceratopogonidae		A108-11/04/08mi
	Chironomidae	*Chironminae*	A101-11/04/08mi
	Coenagrionidae	*Pseudagrion*	C101-11/04/08mi D64-11/04/08mi D73-11/04/08mi F50-11/06/22mi F51-11/06/22mi F54-11/06/22mi F62-11/06/22mi F63-11/06/22mi
No photo available	Corixidae	*Micronecta*	A97-11/04/08mi A98-11/04/08mi F56-11/06/22mi
No photo available	Culcidae		D83-11/04/08mi
	Elmidae		F57-11/06/22mi

	Gerridae	*Eurymetra*	F47-11/06/22mi F49-11/06/22mi F53-11/06/22mi F62-11/06/22mi
	Gomphidae	*Ceratogomphus*	F58-11/06/22mi
	Gyrinidae		C105-11/04/08mi
	Hydrophilidae		D76-11/04/08mi
	Hydropsychidae	*Cheumatopsyche*	F46-11/06/22mi F60-11/06/22mi

	Libellulidae		F64-11/06/22mi
			
	Mesoveliidae		A90-11/04/08mi A91-11/04/08mi C102-11/04/08mi C104-11/04/08mi
			
	Notonectidae	*Anisops*	A93-11/04/08mi A94-11/04/08mi A95-11/04/08mi B81-11/04/08mi B82-11/04/08mi B83-11/04/08mi D62-11/04/08mi
			
	Pleidae		D76-11/04/08mi
			
	Pyralidae		F59-11/06/22mi
			

	Tetragnathidae	A99-11/04/08mi A100-11/04/08mi

REFERENCE LIST

ACB. (African Centre for Biosafety). 2010. ACB Briefing Paper no. 13. The monitoring of enviromental impacts of GMOs in South Africa: a status quo report. Johannesburg.

Acts ***see*** South Africa.

Aimanova, K.G., Zhuang, M. & Gill, S.S. 2006. Expression of Cry1Ac cadherin receptors in insect midgut and cell lines. *Journal of Invertebrate Pathology*, 92(3):178-187.

Al-Gubory, K.H. 2014. Environmental pollutants and lifestyle factors induce oxidative stress and poor prenatal development. *Reproductive BioMedicine Online*, 29(1):17-31.

Alexander, A.C., Culp, J.M., Baird, D.J. & Cessna, A.J. 2016. Nutrient–insecticide interactions decouple density-dependent predation pressure in aquatic insects. *Freshwater Biology*, 61(12):2090-2101.

Alford, A. & Krupke, C. 2018. A Meta-analysis and economic evaluation of neonicotinoid seed treatments and other prophylactic insecticides in Indiana maize From 2000–2015 with IPM recommendations. *Journal of Economic Entomology*, 111(2):689-699.

Alkan, C., Coe, B.P. & Eichler, E.E. 2011. Genome structural variation discovery and genotyping. *Nature Reviews Genetics*, 12(5):363-376.

Alvarez-Alfageme, F., Ortego, F. & Castanera, P. 2009. Bt maize fed-prey mediated effect on fitness and digestive physiology of the ground predator *Poecilus cupreus* L. (Coleoptera: Carabidae). *Journal of Insect Physiology*, 55(2):143-149.

Amann, R.I., Ludwig, W. & Schleifer, K.-H. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological Reviews*, 59(1):143-169.

Ambur, O.H., Frye, S.A. & Tonjum, T. 2007. New functional identity for the DNA uptake sequence in transformation and its presence in transcriptional terminators. *Journal of Bacteriology*, 189(5):2077-2085.

Assaad, F.F. & Signer, E.R. 1990. Cauliflower mosaic virus P35S promoter activity in *Escherichia coli*. *Molecular and General Genetics MGG*, 223(3):517-520.

Avisar, D., Eilenberg, H., Keller, M., Reznik, N., Segal, M., Sneh, B. & Zilberstein, A. 2009. The *Bacillus thuringiensis* delta-endotoxin Cry1C as a potential bioinsecticide in plants. *Plant Science*, 176(3):315-324.

Axelsson, E.P., Hjältén, J., LeRoy, C.J., Whitham, T.G., Julkunen-Tiitto, R. & Wennström, A. 2011. Leaf litter from insect-resistant transgenic trees causes changes in aquatic insect community composition. *Journal of Applied Ecology*, 48(6):1472-1479.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T. & Edwards, R.A. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*, 9(1):75.

Baird, D.J. & Hajibabaei, M. 2012. Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Molecular Ecology*, 21(8):2039-2044.

Baldy, V., Gessner, M.O. & Chauvet, E. 1995. Bacteria, fungi and the breakdown of leaf litter in a large river. *Oikos*, 74(1):93-102.

Baltrus, D.A. 2013. Exploring the costs of horizontal gene transfer. *Trends in Ecology & Evolution*, 28(8):489-495.

Bansal, M.S., Banay, G., Harlow, T.J., Gogarten, J.P. & Shamir, R. 2013. Systematic inference of highways of horizontal gene transfer in prokaryotes. *Bioinformatics*, 29(5):571-579.

Barnes, M.A., Turner, C.R., Jerde, C.L., Renshaw, M.A., Chadderton, W.L. & Lodge, D.M. 2014. Environmental conditions influence eDNA persistence in aquatic systems. *Environmental Science & Technology*, 48(3):1819-1827.

Barrows, G., Sexton, S. & Zilberman, D. 2014. Agricultural biotechnology: the promise and prospects of genetically modified crops. *The Journal of Economic Perspectives*, 28(1):99-119.

Baumgarte, S. & Tebbe, C.C. 2005. Field studies on the environmental fate of the Cry1Ab Bt-toxin produced by transgenic maize (MON810) and its effect on bacterial communities in the maize rhizosphere. *Molecular Ecology*, 14(8):2539-2551.

Beachy, R.N., Fedoroff, N.V., Goldberg, R.B., McHughen, A. & Parrott, W. 2008. The burden of proof: a response to Rosi-Marshall *et al*. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 105(7):2255-2256.

Beketov, M.A., Kefford, B.J., Schäfer, R.B. & Liess, M. 2013. Pesticides reduce regional biodiversity of stream invertebrates. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 110(27):11039-11043.

Benbrook, C.M. 2016. Trends in glyphosate herbicide use in the United States and globally. *Environmental Sciences Europe*, 28(1):1-15.

Bertolla, F. & Simonet, P. 1999. Horizontal gene transfers in the environment: natural transformation as a putative process for gene transfers between transgenic plants and microorganisms. *Research in Microbiology*, 150(6):375-384.

Birkhead, T. 2014. Stormy outlook for long-term ecology studies. *Nature*, 514(7523):405-405.

Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., Douglas, W.Y. & de Bruyn, M. 2014. Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution*, 29(6):358-367.

Bøhn, T., Cuhra, M., Traavik, T., Sanden, M., Fagan, J. & Primicerio, R. 2014. Compositional differences in soybeans on the market: Glyphosate accumulates in Roundup Ready GM soybeans. *Food Chemistry*, 153:207-215.

Bøhn, T., Primicerio, R., Hessen, D.O. & Traavik, T. 2008. Reduced fitness of *Daphnia magna* fed a Bt-transgenic maize variety. *Archives of Environmental Contamination and Toxicology*, 55(4):584-592.

Bøhn, T., Rover, C.M. & Semenchuk, P.R. 2016. *Daphnia magna* negatively affected by chronic exposure to purified Cry-toxins. *Food and Chemical Toxicology*, 91:130-140.

Bøhn, T., Traavik, T. & Primicerio, R. 2010. Demographic responses of *Daphnia magna* fed transgenic Bt-maize. *Ecotoxicology*, 19(2):419-430.

BOLD Systems (Barcode of Life Data Systems). 2014. BOLD Systems Databases. http://www.boldsystems.org/index.php/databases. Date of access: 20 June 2015.

Bolger, A.M., Lohse, M. & Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15):2114-2120.

Böll, S., Schmidt, B.R., Veith, M., Wagner, N., Rödder, D., Weimann, C., Kirschey, T. & Lötters, S. 2013. Anuran amphibians as indicators for changes in aquatic and terrestrial ecosystems following GM crop cultivation: a monitoring guideline. *BioRisk*, 8:39-51.

Bothma, G., Mashaba, C., Mkhonza, N., Chakauya, E. & Chikwamba, R. 2010. GMOs in Africa: opportunities and challenges in South Africa. *GM Crops*, 1(4):175-180.

Bothma, G.C. 2006. Field Experience and Methodologies for Monitoring Environmental Effects of GM Crops in South Africa. *(In* Ghosh, K. & Jepson, P.C., *eds*. Genetically modified organisms in crop production and their effects on the environment: methodologies for monitoring and the way ahead. Rome: FAO (Food and agriculture organization of the United Nations). p. 95-103).

Böttger, R., Schaller, J., Lintow, S. & Dudel, E.G. 2015. Aquatic degradation of Cry1Ab protein and decomposition dynamics of transgenic corn leaves under controlled conditions. *Ecotoxicology and Environmental Safety*, 113:454-459.

Bravo, A., Gill, S.S. & Soberón, M. 2007. Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. *Toxicon*, 49(4):423-435.

Bryan, A. & Swanson, M.S. 2011. Oligonucleotides stimulate genomic alterations of *Legionella pneumophila*. *Molecular Microbiology*, 80(1):231-247.

Bundock, P., den Dulk-Ras, A., Beijersbergen, A. & Hooykaas, P.J. 1995. Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *The EMBO Journal*, 14(13):3206-3214.

Cairns, J., Ruokolainen, L., Hultman, J., Tamminen, M., Virta, M. & Hiltunen, T. 2018. Ecology determines how low antibiotic concentration impacts community composition and horizontal transfer of resistance genes. *Communications Biology*, 1(1):35.

Caquet, T., Roucaute, M., Le Goff, P. & Lagadic, L. 2011. Effects of repeated field applications of two formulations of *Bacillus thuringiensis* var. *israelensis* on non-target saltmarsh invertebrates in Atlantic coastal wetlands. *Ecotoxicology and Environmental Safety*, 74(5):1122-1130.

Carstens, K., Anderson, J., Bachman, P., De Schrijver, A., Dively, G., Federici, B., Hamer, M., Gielkens, M., Jensen, P. & Lamp, W. 2012. Genetically modified crops and aquatic ecosystems: considerations for environmental risk assessment and non-target organism testing. *Transgenic Research*, 21(4):813-842.

Chambers, C.P., Whiles, M.R., Rosi-Marshall, E.J., Tank, J.L., Royer, T.V., Griffiths, N.A., Evans-White, M.A. & Stojak, A.R. 2010. Responses of stream macroinvertebrates to Bt maize leaf detritus. *Ecological Applications*, 20(7):1949-1960.

Chan, C.X., Beiko, R.G., Darling, A.E. & Ragan, M.A. 2009. Lateral transfer of genes and gene fragments in prokaryotes. *Genome Biology and Evolution*, 1:429-438.

Chapman, P.M. 2018. Negatives and Positives: Contaminants and Other Stressors in Aquatic Ecosystems. *Bulletin of Environmental Contamination and Toxicology*, 100(1):3-7.

Chen, I., Christie, P.J. & Dubnau, D. 2005. The ins and outs of DNA transfer in bacteria. *Science*, 310(5753):1456-1460.

Chen, I. & Dubnau, D. 2004. DNA uptake during bacterial transformation. *Nature Reviews Microbiology*, 2(3):241-249.

Chewapreecha, C., Marttinen, P., Croucher, N.J., Salter, S.J., Harris, S.R., Mather, A.E., Hanage, W.P., Goldblatt, D., Nosten, F.H. & Turner, C. 2014. Comprehensive identification of single nucleotide polymorphisms associated with beta-lactam resistance within pneumococcal mosaic genes. *PLoS Genetics*, 10(8):e1004547.

Choi, I.G. & Kim, S.H. 2007. Global extent of horizontal gene transfer. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 104(11):4489-4494.

Claverys, J.-P., Martin, B. & Polard, P. 2009. The genetic transformation machinery: composition, localization, and mechanism. *FEMS Microbiology Reviews*, 33(3):643-656.

Cortez, D., Delaye, L., Lazcano, A. & Becerra, A. 2009. Composition-based methods to identify horizontal gene transfer. *(In* Gogarten, M.B., Gogarten, J.P. & Olendzenski, L.C., *eds*. Horizontal Gene Transfer. Methods in Molecular Biology. Humana Press. p. 215-225).

Coupat, B., Chaumeille-Dole, F., Fall, S., Prior, P., Simonet, P., Nesme, X. & Bertolla, F. 2008. Natural transformation in the *Ralstonia solanacearum* species complex: number and size of DNA that can be transferred. *FEMS Microbiology Ecology*, 66(1):14-24.

Crickmore, N., Baum, J., Bravo, A., Lereclus, D., Narva, K., Sampson, K., Schnepf, E., Sun, M. , Zeigler, D.R 2016. *Bacillus thuringiensis* toxin nomenclature. http://www.btnomenclature.info/ Date of access: 22 March 2016 2016.

Crickmore, N., Zeigler, D., Feitelson, J., Schnepf, E., Van Rie, J., Lereclus, D., Baum, J. & Dean, D. 1998. Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal proteins. *Microbiology and Molecular Biology Reviews*, 62(3):807-813.

Cristescu, M.E. 2014. From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity. *Trends in Ecology & Evolution*, 29(10):566-571.

Croucher, N.J., Harris, S.R., Barquist, L., Parkhill, J. & Bentley, S.D. 2012. A high-resolution view of genome-wide pneumococcal transformation. *PLoS Pathogens*, 8(6):e1002745.

Croucher, N.J., Mostowy, R., Wymant, C., Turner, P., Bentley, S.D. & Fraser, C. 2016. Horizontal DNA transfer mechanisms of bacteria as weapons of intragenomic conflict. *PLoS Biology*, 14(3):e1002394.

Cuhra, M., Traavik, T. & Bøhn, T. 2013. Clone-and age-dependent toxicity of a glyphosate commercial formulation and its active ingredient in *Daphnia magna*. *Ecotoxicology*, 22(2):251-262.

Dallas, H.F. 2007. River health programme: South African Scoring System (SASS) data interpretation guidelines. Department of Water Affairs and Forestry (Resource Quality Services) and the Institute of Natural Resources. http://www.dwa.gov.za/iwqs/rhp/methods/SASS%20 Interpretation%20Guidelines.pdf. Date of access: 11 November 2018.

Dallas, H.F. & Rivers-Moore, N. 2014. Ecological consequences of global climate change for freshwater ecosystems in South Africa: review article. *South African Journal of Science*, 110(5 & 6):1-11.

Darling, A.E., Mau, B., Blattner, F.R. & Perna, N.T. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Research*, 14(7):1394-1403.

Darling, A.E., Mau, B. & Perna, N.T. 2010. Progressive Mauve: multiple genome alignment with gene gain, loss and rearrangement. *PloS One*, 5(6):e11147.

Darmon, E. & Leach, D.R. 2014. Bacterial genome instability. *Microbiology and Molecular Biology Reviews*, 78(1):1-39.

Datta, A. 2013. Genetic engineering for improving quality and productivity of crops. *Agriculture & Food Security*, 2(1):15.

Davis, R.S. & Peterson, R.K. 2008. Effects of single and multiple applications of mosquito insecticides on nontarget arthropods. *Journal of the American Mosquito Control Association*, 24(2):270-280.

Davison, J. 1999. Genetic exchange between bacteria in the environment. *Plasmid*, 42(2):73-91.

de Aldecoa, A.L.I., Zafra, O. & González-Pastor, J.E. 2017. Mechanisms and regulation of extracellular DNA release and its biological roles in microbial communities. *Frontiers in Microbiology*, 8:1390.

De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E. & Taberlet, P. 2014. DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet. *Molecular Ecology Resources*, 14(2):306-323.

De Groot, M.J.A., Bundock, P., Hooykaas, P.J.J. & Beijersbergen, A.G. 1998. *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. *Nature biotechnology*, 16:839–842.

de Vries, J., Heine, M., Harms, K. & Wackernagel, W. 2003. Spread of recombinant DNA by roots and pollen of transgenic potato plants, identified by highly specific biomonitoring using natural transformation of an *Acinetobacter* sp. *Applied and Environmental Microbiology*, 69(8):4455-4462.

de Vries, J. & Wackernagel, W. 2002. Integration of foreign DNA during natural transformation of *Acinetobacter* sp. by homology-facilitated illegitimate recombination. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 99(4):2094-2099.

de Vries, J. & Wackernagel, W. 2005. Microbial horizontal gene transfer and the DNA release from transgenic crop plants. *Plant and Soil*, 266(1-2):91-104.

Deagle, B.E., Eveson, J.P. & Jarman, S.N. 2006. Quantification of damage in DNA recovered from highly degraded samples–a case study on DNA in faeces. *Frontiers in Zoology*, 3(1):11.

DEAT. Department of Environmental Affairs and Tourism. 2009. Environmental risk assessment framework for genetically modified organisms: a guidance document. Pretoria. http://biosafety.org.za/cms/modules/media/scripts/documents/document.handler.php?media_files_id=868. 11 November 2018.

Deiner, K. & Altermatt, F. 2014. Transport distance of invertebrate environmental DNA in a natural river. *PloS one*, 9(2):e88786.

Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P. & Miaud, C. 2011. Persistence of environmental DNA in freshwater ecosystems. *PloS one*, 6(8):e23398.

Dell'Anno, A. & Danovaro, R. 2005. Extracellular DNA plays a key role in deep-sea ecosystem functioning. *Science*, 309(5744):2179-2179.

Dickens, C.W. & Graham, P.M. 2002. The South African Scoring System (SASS) version 5 rapid bioassessment method for rivers. *African Journal of Aquatic Science*, 27(1):1-10.

Dickman, M. 2000. Impacts of a mosquito selective pesticide, Bti, on the macroinvertebrates of a subtropical stream in Hong Kong. *Chemosphere*, 41(1–2):209-217.

Domingues, S., Harms, K., Fricke, W.F., Johnsen, P.J., da Silva, G.J. & Nielsen, K.M. 2012a. Natural transformation facilitates transfer of transposons, integrons and gene cassettes between bacterial species. *PLOS Pathogens*, 8(8):e1002837.

Domingues, S., Nielsen, K.M. & da Silva, G.J. 2012b. Various pathways leading to the acquisition of antibiotic resistance by natural transformation. *Mobile Genetic Elements*, 2(6):257-260.

Douville, M., Gagne, F., Andre, C. & Blaise, C. 2009. Occurrence of the transgenic corn *cry1Ab* gene in freshwater mussels (*Elliptio complanata*) near corn fields: evidence of exposure by bacterial ingestion. *Ecotoxicology and Environmental Safety*, 72(1):17-25.

Douville, M., Gagne, F., Blaise, C. & Andre, C. 2007. Occurrence and persistence of *Bacillus thuringiensis* (Bt) and transgenic Bt corn *cry1Ab* gene from an aquatic environment. *Ecotoxicology and Snvironmental Safety*, 66(2):195-203.

Douville, M., Gagné, F., Masson, L., McKay, J. & Blaise, C. 2005. Tracking the source of *Bacillus thuringiensis* Cry1Ab endotoxin in the environment. *Biochemical Systematics and Ecology*, 33(3):219-232.

DowAgroSciences. 2011. Petition for Determination of Nonregulated Status for Herbicide Tolerant DAS-444Ø6-6 Soybean. (p228). Indianapolis.

Duan, J.J., Marvier, M., Huesing, J., Dively, G. & Huang, Z.Y. 2008. A meta-analysis of effects of Bt crops on honey bees (Hymenoptera: Apidae). *PLoS One*, 3(1):e1415.

Duke, S.O. 2005. Taking stock of herbicide-resistant crops ten years after introduction. *Pest Management Science*, 61(3):211-218.

Duke, S.O., Rimando, A.M., Pace, P.F., Reddy, K.N. & Smeda, R.J. 2003. Isoflavone, glyphosate, and aminomethylphosphonic acid levels in seeds of glyphosate-treated, glyphosate-resistant soybean. *Journal of Agricultural and Food Chemistry*, 51(1):340-344.

Dutra, B.E., Sutera, V.A. & Lovett, S.T. 2007. RecA-independent recombination is efficient but limited by exonucleases. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 104(1):216-221.

DWAF. 2008. Department of Water Affairs and Forestry. National Aquatic Ecosystem Health Monitoring Programme (NAEHMP): River Health Programme (RHP) Implementation Manual. Version 2. Pretoria. http://www.dwaf.gov.za/iwqs/rhp/projectdocuments/river_health_programme_rhp_implementation_manual.pdf. Date of access 11 November 2018.

EFSA. 2004. Opinion of the Scientific Panel on Geneticall Modified Organisms on the use of antibiotic resistance genes as marker genes in genetically modified plants. *The EFSA Journal*, 48:1-18.

EFSA. 2008. Safety and nutritional assessment of GM plants and derived food and feed: the role of animal feeding trials. *Food and Chemical Toxicology*, 46(Supplement 1):pp. S2-S70.

EFSA. 2015. Explanatory note on DNA sequence similarity searches in the context of the assessment of horizontal gene transfer from plants to microorganisms. *EFSA Supporting Publications*, 12(12):916E.

Ekrem, T., Stur, E. & Hebert, P.D.N. 2010. Females do count: Documenting Chironomidae (Diptera) species diversity using DNA barcoding. *Organisms Diversity & Evolution*, 10(5):397-408.

Eliot, T.S. 1915. The Love Song of J. Alfred Prufrock. *Poetry: A Magazine of Verse.* June 1915.

Elkins, C., Thomas, C., Seifert, H. & Sparling, P. 1991. Species-specific uptake of DNA by gonococci is mediated by a 10-base-pair sequence. *Journal of Bacteriology*, 173(12):3911-3913.

Ellington, R.G. 2003. Quantification of the impact of irrigation on the aquifer underlying the Vaalharts Irrigation Scheme. Water Research Commission, Pretoria. (WRC report 1322/2/04).

Epp, L.S., Boessenkool, S., Bellemain, E.P., Haile, J., Esposito, A., Riaz, T., Erseus, C., Gusarov, V.I., Edwards, M.E. & Johnsen, A. 2012. New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. *Molecular Ecology*, 21(8):1821-1833.

Erasmus, A. & Van Den Berg, J. 2014. Effect of Bt-maize expressing Cry1Ab toxin on non-target Coleoptera and Lepidoptera pests of maize in South Africa. *African Entomology*, 22(1):167-179.

Erasmus, A., Van Rensburg, J. & Van Den Berg, J. 2010. Effects of Bt maize on *Agrotis segetum* (Lepidoptera: Noctuidae): a pest of maize seedlings. *Environmental Entomology*, 39(2):702-706.

Fani, F., Leprohon, P., Zhanel, G.G., Bergeron, M.G. & Ouellette, M. 2014. Genomic analyses of DNA transformation and penicillin resistance in *Streptococcus pneumoniae* clinical isolates. *Antimicrobial Agents and Chemotherapy*, 58(3):1397-1403.

Farias, J.R., Andow, D.A., Horikoshi, R.J., Sorgatto, R.J., Fresia, P., dos Santos, A.C. & Omoto, C. 2014. Field-evolved resistance to Cry1F maize by *Spodoptera frugiperda* (Lepidoptera: Noctuidae) in Brazil. *Crop Protection*, 64:150-158.

Fayolle, S., Bertrand, C., Logez, M. & Franquet, É. 2015. Does mosquito control by Bti spraying affect the phytoplankton community? A 5-year study in Camargue temporary wetlands (France). *Annales de Limnologie-International Journal of Limnology*, 51(3):189-198.

Fernandez-Cornejo, J., Hallahan, C., Nehring, R., Wechsler, S. & Grube, A. 2013. Conservation tillage, herbicide use, and genetically engineered crops in the United States: the case of soybeans. *AgBioForum*, 15(3):231-241.

Ferré, J., Van Rie, J. & MacIntosh, S.C. 2008. *(In* Romeis, J., Shelton, A.M., Kennedy, G.G., *ed.* Integration of insect-resistant genetically modified crops within IPM Programs. Progress in Biological Control, vol 5. Dortrecht: Springer. p. 41-85).

Fitzpatrick, D.A. 2012. Horizontal gene transfer in fungi. *FEMS Microbiology Letters*, 329(1):1-8.

Fleeger, J.W., Carman, K.R. & Nisbet, R.M. 2003. Indirect effects of contaminants in aquatic ecosystems. *Science of The Total Environment*, 317(1–3):207-233.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase Subunit 1 from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3(5):294-299.

Fourie, A.J. 2014. Aspects of the biological integrity of the Mutale, Mutshindudi and Tshinane rivers, Limpopo Province. Johannesburg: University of Johannesburg. (Master's dissertation).

Friberg, N. 2010. Pressure-response relationships in stream ecology: introduction and synthesis. *Freshwater Biology*, 55(7):1367-1381.

García-Aljaro, C., Ballesté, E. & Muniesa, M. 2017. Beyond the canonical strategies of horizontal gene transfer in prokaryotes. *Current Opinion in Microbiology*, 38:95-105.

Gebhard, F. & Smalla, K. 1998. Transformation of *Acinetobacter* sp. strain BD413 by transgenic sugar beet DNA. *Applied and Environmental Microbiology*, 64(4):1550-1554.

Gennaro, A., Gomes, A., Herman, L., Nogue, F., Papadopoulou, N. & Tebbe, C. 2017. Explanatory note on DNA sequence similarity searches in the context of the assessment of

horizontal gene transfer from plants to microorganisms. *EFSA Supporting Publications*, 14(7):EN-1273.

Gómez, I., Sánchez, J., Muñoz-Garay, C., Matus, V., Gill, S.S., Soberón, M. & Bravo, A. 2014. *Bacillus thuringiensis* Cry1A toxins are versatile proteins with multiple modes of action: two distinct pre-pores are involved in toxicity. *Biochemical Journal*, 459(2):383-396.

Griffiths, N.A., Tank, J.L., Royer, T.V., Rosi-Marshall, E.J., Whiles, M.R., Chambers, C.P., Frauendorf, T.C. & Evans-White, M.A. 2009. Rapid decomposition of maize detritus in agricultural headwater streams. *Ecological Applications*, 19(1):133-142.

Grohmann, E., Muth, G. & Espinosa, M. 2003. Conjugative plasmid transfer in gram-positive bacteria. *Microbiology and Molecular Biology Reviews: MMBR*, 67(2):277-301, Table of contents.

Gu, J., Bakke, A.M., Valen, E.C., Lein, I. & Krogdahl, Å. 2014. Bt-maize (MON810) and Non-GM soybean meal in diets for Atlantic salmon (*Salmo salar* L.) juveniles–impact on survival, growth performance, development, digestive function, and transcriptional expression of intestinal immune and stress responses. *PloS one*, 9(6):e99932.

Guan, P. & Sung, W.-K. 2016. Structural variation detection using next-generation sequencing data: a comparative technical review. *Methods*, 102:36-49.

Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, 29(8):1072-1075.

Hajibabaei, M., Smith, M., Janzen, D.H., Rodriguez, J.J., Whitfield, J.B. & Hebert, P.D.N. 2006. A minimalist barcode can identify a specimen whose DNA is degraded. *Molecular Ecology Notes*, 6(4):959-964.

Hall, T. 2011. BioEdit: an important software for molecular biology. *GERF Bulletin of Biosciences*, 2(1):60-61.

Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41(41):95-98.

Hanage, W.P. 2016. Not so simple after all: bacteria, their population genetics, and recombination. *Cold Spring Harbor Perspectives in Biology*, 8(7):a018069.

Hao, W. & Golding, G.B. 2006. The fate of laterally transferred genes: life in the fast lane to adaptation or death. *Genome Research*, 16(5):636-643.

Harms, K., Lunnan, A., Hülter, N., Mourier, T., Vinner, L., Andam, C.P., Marttinen, P., Fridholm, H., Hansen, A.J. & Hanage, W.P. 2016. Substitutions of short heterologous DNA segments of intragenomic or extragenomic origins produce clustered genomic polymorphisms. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 113(52):15066-15077.

Hart, R.C. 2001. Rapid estimation of in situ growth rates of *Caridina nilotica* (Crustacea: Decapoda) in Lake Victoria: description and pilot application of a simple, field-compatible technique. *Limnology and Oceanography*, 46(3):692-698.

Hart, R.C., Stewart, B. & Bickerton, I.B. 2001. Decapoda. *(In* Day, J., Stewart, B., De Moor, I. & Louw, A., *eds*. Guides to the freshwater invertebrates of Southern Africa–Vol. 4: Crustacea III Bathynellacea, Amphipoda, Isopoda, Spelaeogriphacea, Tanaidacea and Decapoda. Water Research Commission Report TT141/01. Water Research Commission: Pretoria
p. 87-123).

Harwood, J.D., Wallin, W.G. & Obrycki, J.J. 2005. Uptake of Bt endotoxins by nontarget herbivores and higher order arthropod predators: molecular evidence from a transgenic corn agroecosystem. *Molecular Ecology*, 14(9):2815-2823.

Hebert, P.D.N., Cywinska, A. & Ball, S.L. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270:313-321.

Heinemann, J.A. & Sprague, G.F., Jr. 1989. Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. *Nature*, 340(6230):205-209.

Heinemann, J.A. & Traavik, T. 2004. Problems in monitoring horizontal gene transfer in field trials of transgenic plants. *Nature Biotechnology*, 22(9):1105-1109.

Helassa, N., Quiquampoix, H., Noinville, S., Szponarski, W. & Staunton, S. 2009. Adsorption and desorption of monomeric Bt (*Bacillus thuringiensis*) Cry1Aa toxin on montmorillonite and kaolinite. *Soil Biology and Biochemistry*, 41(3):498-504.

Hernández, M., Pla, M., Esteve, T., Prat, S., Puigdomenech, P. & Ferrando, A. 2003. A specific real-time quantitative PCR detection system for event MON810 in maize YieldGard® based on the 3′-transgene integration sequence. *Transgenic Research*, 12(2):179-189.

Herold, C., Bailey, A.K. & Stewart, S. 1996. Long term salt balance of the Vaalharts Irrigation Scheme. Water Research Commission, Pretoria. (WRC Report 420/1/96.).

Hieber, M. & Gessner, M.O. 2002. Contribution of stream detrivores, fungi, and bacteria to leaf breakdown based on biomass estimates. *Ecology*, 83(4):1026-1038.

Hilbeck, A., Binimelis, R., Defarge, N., Steinbrecher, R., Székács, A., Wickson, F., Antoniou, M., Bereano, P.L., Clark, E.A. & Hansen, M. 2015. No scientific consensus on GMO safety. *Environmental Sciences Europe*, 27(1):1.

Hjältén, J., Axelsson, E.P., Julkunen-Tiitto, R., Wennström, A. & Pilate, G. 2013. Innate and introduced resistance traits in genetically modified aspen trees and their effect on leaf beetle feeding. *PLoS ONE*, 8(9):e73819.

Holderbaum, D.F., Cuhra, M., Wickson, F., Orth, A.I., Nodari, R.O. & Bøhn, T. 2015. Chronic responses of *Daphnia magna* under dietary exposure to leaves of a transgenic (event MON810) Bt–maize hybrid and its conventional near-isoline. *Journal of Toxicology and Environmental Health, Part A*, 78(15):993-1007.

Holland, J. 2004. The environmental consequences of adopting conservation tillage in Europe: reviewing the evidence. *Agriculture, Ecosystems & Environment*, 103(1):1-25.

Holst-Jensen, A., Bertheau, Y., de Loose, M., Grohmann, L., Hamels, S., Hougs, L., Morisset, D., Pecoraro, S., Pla, M., Van den Bulcke, M. & Wulff, D. 2012. Detecting un-authorized genetically modified organisms (GMOs) and derived materials. *Biotechnology Advances*, 30(6):1318-1335.

Hug, K. 2007. Genetically modified organisms: do the benefits outweigh the risks? *Medicina*, 44(2):87-99.

Hülter, N. & Wackernagel, W. 2008. Double illegitimate recombination events integrate DNA segments through two different mechanisms during natural transformation of *Acinetobacter baylyi*. *Molecular Microbiology*, 67(5):984-995.

Hultman, J., Tamminen, M., Pärnänen, K., Cairns, J., Karkman, A. & Virta, M. 2018. Host range of antibiotic resistance genes in wastewater treatment plant influent and effluent. *FEMS Microbiology Ecology*, 94(4):fiy038.

Huseth, A.S., Chappell, T.M., Chitturi, A., Jacobson, A.L. & Kennedy, G.G. 2018. Insecticide resistance signals negative consequences of widespread neonicotinoid use on multiple field crops in the US Cotton Belt. *Environmental Science & Technology*, 52(4):2314-2322.

Hynes, H.B.N. 1975. Edgardo Baldi memorial lecture. The stream and its valley. *Verhandlungen der Internationalen Vereinigung fur theoretische und angewandte Limnologie*, 19:1-15.

Ibrahim, M.A., Griko, N., Junker, M. & Bulla, L.A. 2010. *Bacillus thuringiensis*: a genomics and proteomics perspective. *Bioengineered Bugs*, 1(1):31-50.

Iqbal, Z., Caccamo, M., Turner, I., Flicek, P. & McVean, G. 2012. De novo assembly and genotyping of variants using colored de Bruijn graphs. *Nature Genetics*, 44(2):226-232.

ISAAA (International Service for the Acquisition of Agri-Biotech Applications). 2016. GM Approval Database. http://www.isaaa.org/gmapprovaldatabase/. Date of access: 2016-08-10 2016.

Iversen, M., Grønsberg, I.M., van den Berg, J., Fischer, K., Aheto, D.W. & Bøhn, T. 2014. Detection of transgenes in local maize varieties of small-scale farmers in Eastern Cape, South Africa. *PloS One*, 9(12):e116147.

Jakob, C. & Poulin, B. 2016. Indirect effects of mosquito control using Bti on dragonflies and damselflies (Odonata) in the Camargue. *Insect Conservation and Diversity*, 9(2):161-169.

James, C. 2014. Global Status of Commercialized Biotech/GM Crops: 2014. International Service for the Acquisition of Agri-Biotech Applications, ISAAA Brief 49, Ithaca, NY.

James, C. 2016. Global Status of Commercialized Biotech/GM Crops: 2016. International Service for the Acquisition of Agri-Biotech Applications, ISAAA Brief No. 52, Ithaca, NY.

Janion, C., Bedos, A., Bengtsson, J., Deharveng, L., Jansen van Vuuren, B., Leinaas, H.P., Liu, A., Malmstrom, A., Porco, D. & Chown, S.L. 2011. Springtail diversity in South Africa: research letter. *South African Journal of Science*, 107(11 & 12):1-7.

Jaynes, D.B., Isenhart, T.M. & Parkin, T.B. 2014. Reconnecting riparian buffers with tile drainage (2). Leopold Centre for Sustainable Agriculture. (Completed grant report 463). https://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=1475&context=leopold_grantreports. Date of Access 11 November 2018.

Jensen, P.D., Dively, G.P., Swan, C.M. & Lamp, W.O. 2010. Exposure and nontarget effects of transgenic Bt corn debris in streams. *Environmental Entomology*, 39(2):707-714.

Ji, Y., Ashton, L., Pedley, S.M., Edwards, D.P., Tang, Y., Nakamura, A., Kitching, R., Dolman, P.M., Woodcock, P. & Edwards, F.A. 2013. Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, 16(10):1245-1257.

Jones, M., Ghoorah, A. & Blaxter, M. 2011. jMOTU and Taxonerator: turning DNA Barcode sequences into annotated operational taxonomic units. *PloS One*, 6(4):e19259.

Jørgensen, S. 2012. Evaluating primary (cry1Ab) and secondary effects (deoxynivalenol) of GM maize when fed to zebrafish (*Danio rerio*): investigating growth, intestinal mRNA and white blood cell differentiation. Bergen: The University of Bergen. (Master's thesis).

Jurat-Fuentes, J.L. & Adang, M.J. 2006. Cry toxin mode of action in susceptible and resistant *Heliothis virescens* larvae. *Journal of Invertebrate Pathology*, 92(3):166-171.

Karakoc, E., Alkan, C., O'Roak, B.J., Dennis, M.Y., Vives, L., Mark, K., Rieder, M.J., Nickerson, D.A. & Eichler, E.E. 2012. Detection of structural variants and indels within exome data. *Nature Methods*, 9(2):176-178.

Kay, E., Vogel, T.M., Bertolla, F., Nalin, R. & Simonet, P. 2002. In situ transfer of antibiotic resistance genes from transgenic (transplastomic) tobacco plants to bacteria. *Applied and Environmental Microbiology*, 68(7):3345-3351.

Keeling, P.J. & Palmer, J.D. 2008. Horizontal gene transfer in eukaryotic evolution. *Nature Reviews Genetics*, 9(8):605.

Keese, P. 2008. Risks from GMOs due to horizontal gene transfer. *Environmental Biosafety Research*, 7(03):123-149.

Kohli, A., Griffiths, S., Palacios, N., Twyman, R.M., Vain, P., Laurie, D.A. & Christou, P. 1999. Molecular characterization of transforming plasmid rearrangements in transgenic rice reveals a recombination hotspot in the CaMV 35S promoter and confirms the predominance of microhomology mediated recombination. *The Plant Journal*, 17(6):591-601.

Komoßa, D., Gennity, I. & Sandermann, H. 1992. Plant metabolism of herbicides with C-P bonds: Glyphosate. *Pesticide Biochemistry and Physiology*, 43(2):85-94.

Koonin, E.V. 2009. Darwinian evolution in the light of genomics. *Nucleic Acids Research*, 37(4):1011-1034.

Koonin, E.V. & Wolf, Y.I. 2008. Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world. *Nucleic Acids Research*, 36(21):6688-6719.

Kowalczykowski, S.C. 2000. Initiation of genetic recombination and recombination-dependent replication. *Trends in Biochemical Sciences*, 25(4):156-165.

Kratz, W., Mante, C., Hofmann, F., Shlechtriemen, U., Kuhn, U., Ober, S. & Vögeld, R. 2010. *(In* Breckling, B. & Verhoeven, R., *eds*. Implications of GM-Crop Cultivation at Large Spatial Scales. Frankfurt, Germany: Peter Lang, 2010. p. 21-23).

Kremer, R., Means, N. & Kim, S. 2005. Glyphosate affects soybean root exudation and rhizosphere micro-organisms. *International Journal of Environmental Analytical Chemistry*, 85(15):1165-1174.

Kruger, M., Van Rensburg, J.B.J. & Van den Berg, J. 2009. Perspective on the development of stem borer resistance to Bt maize and refuge compliance at the Vaalharts Irrigation Scheme in South Africa. *Crop Protection*, 28(8):684-689.

Kung, S.H., Retchless, A.C., Kwan, J.Y. & Almeida, R.P. 2013. Effects of DNA size on transformation and recombination efficiencies in *Xylella fastidiosa*. *Applied and Environmental Microbiology*, 79(5):1712-1717.

Kuo, C.H. & Ochman, H. 2009. The fate of new bacterial genes. *FEMS Microbiology Reviews*, 33(1):38-43.

Kuribara, H., Shindo, Y., Matsuoka, T., Takubo, K., Futo, S., Aoki, N., Hirao, T., Akiyama, H., Goda, Y. & Toyada, M. 2002. Novel reference molecules for quantitation of genetically modified maize and soybean. *Journal of AOAC International*, 85(5):1077-1089.

Kurtzman, C.P. & Robnett, C.J. 1998. Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. *Antonie van Leeuwenhoek*, 73(4):331-371.

Lagadic, L., Roucaute, M. & Caquet, T. 2014. Bti sprays do not adversely affect non-target aquatic invertebrates in French Atlantic coastal wetlands. *Journal of Applied Ecology*, 51(1):102-113.

Lagadic, L., Schäfer, R.B., Roucaute, M., Szöcs, E., Chouin, S., de Maupeou, J., Duchet, C., Franquet, E., Le Hunsec, B., Bertrand, C., Fayolle, S., Francés, B., Rozier, Y., Foussadier, R., Santoni, J.-B. & Lagneau, C. 2016. No association between the use of Bti for mosquito control and the dynamics of non-target aquatic invertebrates in French coastal and continental wetlands. *Science of The Total Environment*, 553:486-494.

Lallias, D., Hiddink, J.G., Fonseca, V.G., Gaspar, J.M., Sung, W., Neill, S.P., Barnes, N., Ferrero, T., Hall, N. & Lambshead, P.J.D. 2015. Environmental metabarcoding reveals heterogeneous drivers of microbial eukaryote diversity in contrasting estuarine ecosystems. *The ISME Journal*, 9(5):1208-1221.

Lane, D.J. 1991. 16S/23S rRNA sequencing. (*In*: Strackebrand, E. and Goodfellow, M. Eds. Nucleic Acid Techniques in Bacterial Systematics, Jphn Wiley and Sons, New York, 115-175.)

Langille, M.G. & Brinkman, F.S. 2009. Bioinformatic detection of horizontally transferred DNA in bacterial genomes. *F1000 Biology Reports*, 1:25.

Laramie, M.B., Pilliod, D.S. & Goldberg, C.S. 2015. Characterizing the distribution of an endangered salmonid using environmental DNA analysis. *Biological Conservation*, 183:29-37.

Leray, M. & Knowlton, N. 2015. DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 112(7):2076-2081.

Li, G., Wang, Y., Liu, B. & Zhang, G. 2014. Transgenic *Bacillus thuringiensis* (Bt) rice is safer to aquatic ecosystems than its non-transgenic counterpart. *PLoS One*, 9(8):e104270.

Li, H. 2011. wgsim-Read simulator for next generation sequencing. Github Repository, https://github.com/lh3/wgsim.

Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv Preprint arXiv:1303.3997*. *[q-bio.GN]*

Li, H., Buschman, L.L., Huang, F., Zhu, K.Y., Bonning, B. & Oppert, B. 2007a. DiPel-selected *Ostrinia nubilali*s larvae are not resistant to transgenic corn expressing *Bacillus thuringiensis* CrylAb. *Journal of Economic Entomology*, 100(6):1862-1870.

Li, H. & Durbin, R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25(14):1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. & Durbin, R. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16):2078-2079.

Li, Y.-L., Du, J., Fang, Z.-X. & You, J. 2013. Dissipation of insecticidal Cry1Ac protein and its toxicity to nontarget aquatic organisms. *Journal of Agricultural and Food Chemistry*, 61(46):10864-10871.

Li, Y., Wu, K., Zhang, Y. & Yuan, G. 2007b. Degradation of Cry1Ac protein within transgenic *Bacillus thuringiensis* rice tissues under field and laboratory conditions. *Environmental Entomology*, 36(5):1275-1282.

Li, Y., Zhang, X., Chen, X., Romeis, J., Yin, X. & Peng, Y. 2015. Consumption of Bt rice pollen containing Cry1C or Cry2A does not pose a risk to *Propylea japonica* (Thunberg)(Coleoptera: Coccinellidae). *Scientific Reports*, 5:7679.

Linn, M.D. & Moore, P.A. 2014. The Effects of Bt Corn on Rusty Crayfish (*Orconectes Rusticus*) Growth and Survival. *Archives of Environmental Contamination and Toxicology*, 67(3):436-443.

Liu, H., Fu, Y., Jiang, D., Li, G., Xie, J., Cheng, J., Peng, Y., Ghabrial, S.A. & Yi, X. 2010. Widespread horizontal gene transfer from double-stranded RNA viruses to eukaryotic nuclear genomes. *Journal of Virology*, 84(22):11876-11887.

Liu, W., Lu, H.H., Wu, W., Wei, Q.K., Chen, Y.X. & Thies, J.E. 2008. Transgenic Bt rice does not affect enzyme activities and microbial composition in the rhizosphere during crop development. *Soil Biology and Biochemistry*, 40(2):475-486.

Lorenz, M.G. & Wackernagel, W. 1994. Bacterial gene transfer by natural genetic transformation in the environment. *Microbiological Reviews*, 58(3):563-602.

Loreto, E.L., Carareto, C.M. & Capy, P. 2008. Revisiting horizontal transfer of transposable elements in *Drosophila*. *Heredity*, 100(6):545-554.

Lövei, G.L. & Arpaia, S. 2005. The impact of transgenic plants on natural enemies: a critical review of laboratory studies. *Entomologia Experimentalis et Applicata*, 114(1):1-14.

Mächler, E., Deiner, K., Steinmann, P. & Altermatt, F. 2014. Utility of environmental DNA for monitoring rare and indicator macroinvertebrate species. *Freshwater Science*, 33(4):1174-1183.

Maddamsetti, R. & Lenski, R.E. 2018. Analysis of bacterial genomes from an evolution experiment with horizontal gene transfer shows that recombination can sometimes overwhelm selection. *PLoS Genetics*, 14(1):e1007199.

Mantel, S.K., Hughes, D.A. & Muller, N.W. 2010. Ecological impacts of small dams on South African rivers Part 1: drivers of change-water quantity and quality. *Water SA*, 36(3):351-360.

Marcet-Houben, M. & Gabaldon, T. 2010. Acquisition of prokaryotic genes by fungal genomes. *Trends in Genetics*, 26(1):5-8.

Marvier, M., McCreedy, C., Regetz, J. & Kareiva, P. 2007. A meta-analysis of effects of Bt cotton and maize on nontarget invertebrates. *Science*, 316(5830):1475-1477.

Matsuoka, T., Kawashima, Y., Akiyama, H., Miura, H., Goda, Y., Kusakabe, Y., Isshiki, K. & Toyoda, M. 2000. A method of detecting recombinant DNAs from four lines of genetically modified maize. *Journal of the Food Hygienic Society of Japan*, 41(2):137-143.

McGeoch, M. & Rhodes, J.I. 2006. Ecological risk assessment of genetically modified organisms in South Africa: An assesment of the current policy framework. CIB Occasional Paper. Centre for Invasion Biology, Stellenbosch. (Report no. 2).

Mell, J.C. & Redfield, R.J. 2014. Natural competence and the evolution of DNA uptake specificity. *Journal of Bacteriology*, 196(8):1471-1483.

Mell, J.C., Shumilina, S., Hall, I.M. & Redfield, R.J. 2011. Transformation of natural genetic variation into *Haemophilus influenzae* genomes. *PLoS Pathogens*, 7(7):e1002151.

Mende, D.R., Sunagawa, S., Zeller, G. & Bork, P. 2013. Accurate and universal delineation of prokaryotic species. *Nature Methods*, 10(9):881.

Mendelsohn, M., Kough, J., Vaituzis, Z. & Matthews, K. 2003. Are Bt crops safe? *Nature Biotechnology*, 21(9):1003-1009.

Mensah, P., Muller, W. & Palmer, C. 2012a. Acetylcholinesterase activity in the freshwater shrimp *Caridina nilotica* as a biomarker of Roundup® herbicide pollution of freshwater systems in South Africa. *Water Science and Technology*, 66(2):402-408.

Mensah, P., Palmer, C. & Muller, W. 2012b. Lipid peroxidation in the freshwater shrimp *Caridina nilotica* as a biomarker of Roundup® herbicide pollution of freshwater systems in South Africa. *Water Science and Technology*, 65(9):1660-1666.

Mensah, P.K., Muller, W.J. & Palmer, C.G. 2011. Acute toxicity of Roundup® herbicide to three life stages of the freshwater shrimp *Caridina nilotica* (Decapoda: Atyidae). *Physics and Chemistry of the Earth, Parts A/B/C*, 36(14):905-909.

Mensah, P.K., Muller, W.J. & Palmer, C.G. 2012c. Using growth measures in the freshwater shrimp *Caridina nilotica* as biomarkers of Roundup® pollution of South African freshwater systems. *Physics and Chemistry of the Earth, Parts A/B/C*, 50:262-268.

Mensah, P.K., Palmer, C.G. & Muller, W.J. 2013. Derivation of South African water quality guidelines for Roundup® using species sensitivity distribution. *Ecotoxicology and Environmental Safety*, 96:24-31.

Mesnage, R., Defarge, N., Spiroux de Vendomois, J. & Seralini, G.E. 2014. Major pesticides are more toxic to human cells than their declared active principles. *BioMed Research International*, 2014:179691.

Mohr, K.I. & Tebbe, C.C. 2007. Field study results on the probability and risk of a horizontal gene transfer from transgenic herbicide-resistant oilseed rape pollen to gut bacteria of bees. *Applied Microbiology and Biotechnology*, 75(3):573-582.

Monier, J.M., Bernillon, D., Kay, E., Faugier, A., Rybalka, O., Dessaux, Y., Simonet, P. & Vogel, T.M. 2007. Detection of potential transgenic plant DNA recipients among soil bacteria. *Environmental Biosafety Research*, 6(1-2):71-83.

Moore, A.A. & Palmer, M.A. 2005. Invertebrate biodiversity in agricultural and urban headwater streams: implications for conservation and management. *Ecological Applications*, 15(4):1169-1177.

Nadel, R.L., Slippers, B., Scholes, M.C., Lawson, S.A., Noack, A.E., Wilcken, C.F., Bouvet, J.P. & Wingfield, M.J. 2010. DNA bar-coding reveals source and patterns of Thaumastocoris peregrinus invasions in South Africa and South America. *Biological Invasions*, 12(5):1067-1077.

Naranjo, S.E. 2009. Impacts of Bt crops on non-target invertebrates and insecticide use patterns. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, 4(011):1-11.

NCBI (National Center for Biotechnology Information). 2014. Basic Local Alignment Search Tool (BLAST). http://blast.ncbi.nlm.nih.gov/Blast.cgi Date of access: 22 May 2017.

Nielsen, K.M., Bohn, T. & Townsend, J.P. 2014. Detecting rare gene transfer events in bacterial populations. *Frontiers in Microbiology*, 4:415.

Nielsen, K.M., Bøhn, T. & Townsend, J.P. 2013. Detecting rare gene transfer events in bacterial populations. *Frontiers in Microbiology*, 4:415.

Nielsen, K.M., Calamai, L. & Pietramellara, G. 2006. Stabilization of extracellular DNA and proteins by transient binding to various soil components. *(In* Nannipieri, P. & Smalla, K., *eds*. Nucleic acids and proteins in soil. Berlin, Heidelberg: Springer. Soil Biology, vol 8. p. 141-157).

Nielsen, K.M., Johnsen, P.J., Bensasson, D. & Daffonchio, D. 2007. Release and persistence of extracellular DNA in the environment. *Environmental Biosafety Research*, 6(1-2):37-53.

Nielsen, K.M. & Townsend, J.P. 2004. Monitoring and modeling horizontal gene transfer. *Nature Biotechnology*, 22(9):1110-1114.

Nijkamp, J.F., van den Broek, M.A., Geertman, J.-M.A., Reinders, M.J.T., Daran, J.-M.G. & de Ridder, D. 2012. De novo detection of copy number variation by co-assembly. *Bioinformatics*, 28(24):3195-3202.

Nikoh, N., Tanaka, K., Shibata, F., Kondo, N., Hizume, M., Shimada, M. & Fukatsu, T. 2008. *Wolbachia* genome integrated in an insect chromosome: evolution and fate of laterally transferred endosymbiont genes. *Genome Research*, 18(2):272-280.

Nikolaidis, N., Doran, N. & Cosgrove, D.J. 2013. Plant expansins in bacteria and fungi: evolution by horizontal gene transfer and independent domain fusion. *Molecular Biology and Evolution*, 31(2):376-386.

Niu, Y., Yang, F., Dangal, V. & Huang, F. 2014. Larval survival and plant injury of Cry1F-susceptible,-resistant, and-heterozygous fall armyworm (Lepidoptera: Noctuidae) on non-Bt and Bt corn containing single or pyramided genes. *Crop Protection*, 59:22-28.

OECD (Organisation for Economic Co-operation and Development). 2010. OECD Guidelines for the Testing of Chemicals Proposal for a new Guideline: Sediment-Water Chironomid Life-Cycle Toxicity Test Using Spiked Water or Spiked Sediment. The Organisation for Economic Co-operation and Development: www.oecd.org/. (Test 233).

Oliveira, A.P., Pampulha, M.E. & Bennett, J.P. 2008. A two-year field study with transgenic *Bacillus thuringiensis* maize: effects on soil microorganisms. *Science of the Total Environment*, 405(1):351-357.

Otieno, F. & Adeyemo, J. 2011. Multi-objective cropping pattern in the Vaalharts irrigation scheme. *African Journal of Agricultural Research*, 6(6):1286-1294.

Otomo, P.V., Jansen van Vuuren, B. & Reinecke, S.A. 2009. Usefulness of DNA barcoding in ecotoxicological investigations: resolving taxonomic uncertainties using *Eisenia Malm* 1877 as an example. *Bulletin of Environmental Contamination and Toxicology*, 82(3):261-264.

Overballe-Petersen, S., Harms, K., Orlando, L.A., Mayar, J.V., Rasmussen, S., Dahl, T.W., Rosing, M.T., Poole, A.M., Sicheritz-Ponten, T., Brunak, S., Inselmann, S., de Vries, J., Wackernagel, W., Pybus, O.G., Nielsen, R., Johnsen, P.J., Nielsen, K.M. & Willerslev, E. 2013. Bacterial natural transformation by highly fragmented and damaged DNA. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 110(49):19860-19865.

Overballe-Petersen, S. & Willerslev, E. 2014. Horizontal transfer of short and degraded DNA has evolutionary implications for microbes and eukaryotic sexual reproduction. *Bioessays*, 36(10):1005-1010.

Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Parrello, B. & Shukla, M. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Rresearch*, 42(D1):D206-D214.

Palma, L., Muñoz, D., Berry, C., Murillo, J. & Caballero, P. 2014. *Bacillus thuringiensis* toxins: an overview of their biocidal activity. *Toxins*, 6(12):3296-3325.

Parisi, C., Tillie, P. & Rodríguez-Cerezo, E. 2016. The global pipeline of GM crops out to 2020. *Nature Biotechnology*, 34(1):31.

Parrott, W. 2008. Study of Bt impact on caddisflies overstates its conclusions: response to Rosi-Marshall *et al*. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 105(7):E10; author reply E11.

Paul, J.H. 1999. Microbial gene transfer: an ecological perspective. *Journal of Molecular Microbiology and Biotechnology*, 1(1):45-50.

Pellegrino, E., Bedini, S., Nuti, M. & 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):3113.

Pereira, V., Fernandes, D., Carvalho, G., Benoliel, M., San Romão, M. & Crespo, M.B. 2010. Assessment of the presence and dynamics of fungi in drinking water sources using cultural and molecular methods. *Water Research*, 44(17):4850-4859.

Perry, E.D., Ciliberto, F., Hennessy, D.A. & Moschini, G. 2016. Genetically engineered crops and pesticide use in US maize and soybeans. *Science Advances*, 2(8):e1600850.

Pilacinski, W., Crawford, A., Downey, R., Harvey, B., Huber, S., Hunst, P., Lahman, L., MacIntosh, S., Pohl, M. & Rickard, C. 2011. Plants with genetically modified events combined by conventional breeding: An assessment of the need for additional regulatory data. *Food and Chemical Toxicology*, 49(1):1-7.

Pineyro-Nelson, A., Van Heerwaarden, J., Perales, H.R., Serratos-Hernandez, J.A., Rangel, A., Hufford, M.B., Gepts, P., Garay-Arroyo, A., Rivera-Bustamante, R. & Alvarez-Buylla, E.R. 2009. Transgenes in Mexican maize: molecular evidence and methodological considerations for GMO detection in landrace populations. *Molecular Ecology*, 18(4):750-761.

Pobjecky, N., Rosenberg, G.H., Dinter-Gottlieb, G. & Käufer, N.F. 1990. Expression of the β-glucuronidase gene under the control of the CaMV 35S promoter in *Schizosaccharomyces pombe*. *Molecular and General Genetics MGG*, 220(2):314-316.

Posthuma, L., Dyer, S.D., de Zwart, D., Kapo, K., Holmes, C.M. & Burton Jr, G.A. 2016. Eco-epidemiology of aquatic ecosystems: Separating chemicals from multiple stressors. *Science of the Total Environment*, 573:1303-1319.

Pott, A., Otto, M. & Schulz, R. 2018. Impact of genetically modified organisms on aquatic environments: Review of available data for the risk assessment. *Science of The Total Environment*, 635:687-698.

Poulin, B., Lefebvre, G. & Paz, L. 2010. Red flag for green spray: adverse trophic effects of Bti on breeding birds. *Journal of Applied Ecology*, 47(4):884-889.

Power, P.M., Bentley, S.D., Parkhill, J., Moxon, E.R. & Hood, D.W. 2012. Investigations into genome diversity of *Haemophilus influenzae* using whole genome sequencing of clinical isolates and laboratory transformants. *BMC Microbiology*, 12(1):273.

Prihoda, K.R. & Coats, J.R. 2008. Aquatic fate and effects of *Bacillus thuringiensis* Cry3Bb1 protein: toward risk assessment. *Environmental Toxicology and Chemistry*, 27(4):793-798.

Qing-ling, Z., Yun-he, L.I., Hong-xia, H.U.A., Chang-ju, Y., Hong-jin, W.U. & Yu-fa, P. 2013. Exposure degree of important non-target arthropods to Cry2Aa in Bt rice fields. *Yingyong Shengtai Xuebao - The Journal of Applied Ecology*, 24(6):1647-1651.

Quist, D. & Chapela, I.H. 2001. Transgenic DNA introgressed into traditional maize landraces in Oaxaca, Mexico. *Nature*, 414(6863):541-543.

Randhawa, G.J., Singh, M., Chhabra, R. & Sharma, R. 2010. Qualitative and quantitative molecular testing methodologies and traceability systems for commercialised Bt cotton events and other Bt crops under field trials in India. *Food Analytical Methods*, 3(4):295-303.

Ravenhall, M., Skunca, N., Lassalle, F. & Dessimoz, C. 2015b. Inferring horizontal gene transfer. *PLoS Computational Biology*, 11(5):e1004095.

Raybould, A., Graser, G., Hill, K. & Ward, K. 2012. Ecological risk assessments for transgenic crops with combined insect-resistance traits: the example of Bt11× MIR604 maize. *Journal of Applied Entomology*, 136(1-2):27-37.

Raybould, A. & Vlachos, D. 2011. Non-target organism effects tests on Vip3A and their application to the ecological risk assessment for cultivation of MIR162 maize. *Transgenic Research*, 20(3):599-611.

Reddy, K.N., Rimando, A.M. & Duke, S.O. 2004. Aminomethylphosphonic acid, a metabolite of glyphosate, causes injury in glyphosate-treated, glyphosate-resistant soybean. *Journal of Agricultural and Food Chemistry*, 52(16):5139-5143.

Reddy, K.N. & Zablotowicz, R.M. 2003. Glyphosate-resistant soybean response to various salts of glyphosate and glyphosate accumulation in soybean nodules. *Weed Science*, 51(4):496-502.

Relyea, R. & Hoverman, J. 2006. Assessing the ecology in ecotoxicology: a review and synthesis in freshwater systems. *Ecology Letters*, 9(10):1157-1171.

Relyea, R.A. 2005a. The impact of insecticides and herbicides on the biodiversity and productivity of aquatic communities. *Ecological Applications*, 15(2):618-627.

Relyea, R.A. 2005b. The lethal impact of Roundup on aquatic and terrestrial amphibians. *Ecological Applications*, 15(4):1118-1124.

Richards, T.A., Leonard, G., Soanes, D.M. & Talbot, N.J. 2011. Gene transfer into the fungi. *Fungal Biology Reviews*, 25(2):98-110.

Richards, T.A., Soanes, D.M., Foster, P.G., Leonard, G., Thornton, C.R. & Talbot, N.J. 2009. Phylogenomic analysis demonstrates a pattern of rare and ancient horizontal gene transfer between plants and fungi. *The Plant Cell*, 21(7):1897-1911.

Rivera, A.L., Magaña-Ortíz, D., Gómez-Lim, M., Fernández, F. & Loske, A.M. 2014. Physical methods for genetic transformation of fungi and yeast. *Physics of Life Reviews*, 11(2):184-203.

Rizzo, L., Manaia, C., Merlin, C., Schwartz, T., Dagot, C., Ploy, M.C., Michael, I. & Fatta-Kassinos, D. 2013. Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: a review. *Science of the Total Environment*, 447:345-360.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G. & Mesirov, J.P. 2011. Integrative genomics viewer. *Nature biotechnology*, 29(1):24-26.

Rocha, E.P.C., Cornet, E. & Michel, B. 2005. Comparative and evolutionary analysis of the bacterial homologous recombination systems. *PLoS Genetics*, 1(2):e15.

Romeis, J., Raybould, A., Bigler, F., Candolfi, M.P., Hellmich, R.L., Huesing, J.E. & Shelton, A.M. 2013. Deriving criteria to select arthropod species for laboratory tests to assess the ecological risks from cultivating arthropod-resistant genetically engineered crops. *Chemosphere*, 90(3):901-909.

Rosati, A., Bogani, P., Santarlasci, A. & Buiatti, M. 2008. Characterisation of 3′ transgene insertion site and derived mRNAs in MON810 YieldGard® maize. *Plant Molecular Biology*, 67(3):271-281.

Rosi-Marshall, E.J. & Royer, T.V. 2012. Pharmaceutical compounds and ecosystem function: an emerging research challenge for aquatic ecologists. *Ecosystems*, 15(6):867-880.

Rosi-Marshall, E.J., Tank, J.L., Royer, T.V., Whiles, M.R., Evans-White, M., Chambers, C., Griffiths, N.A., Pokelsek, J. & Stephen, M.L. 2007. Toxins in transgenic crop byproducts may affect headwater stream ecosystems. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 104(41):16204-16208.

Sagstad, A., Sanden, M., Haugland, Ø., Hansen, A.C., Olsvik, P.A. & Hemre, G.I. 2007. Evaluation of stress-and immune-response biomarkers in Atlantic salmon, *Salmo salar* L., fed different levels of genetically modified maize (Bt maize), compared with its near-isogenic parental line and a commercial suprex maize. *Journal of Fish Diseases*, 30(4):201-212.

SANBI (South African National Biodiversity Institute). 2011. Monitoring the environmental impacts of GM maize in South Africa: the outcome of the South Africa-Norway Biosafety Cooperation Project (2008-2010). Department of Environmental Affairs and Tourism, Pretoria. https://www.sanbi.org/documents/monitoring-the-environmental-impacts-of-gm-maize-in-south-africa/. Date of access: 11 November 2018.

Saxena, D., Flores, S. & Stotzky, G. 2002. Vertical movement in soil of insecticidal Cry1Ab protein from *Bacillus thuringiensis*. *Soil Biology and Biochemistry*, 34(1):111-120.

Saxena, D. & Stotzky, G. 2000. Insecticidal toxin from *Bacillus thuringiensis* is released from roots of transgenic Bt corn in vitro and in situ. *FEMS Microbiology Ecology*, 33(1):35-39.

Saxena, D. & Stotzky, G. 2001. Bt corn has a higher lignin content than non-Bt corn. *American Journal of Botany*, 88(9):1704-1706.

Schäfer, R.B., Kühn, B., Malaj, E., König, A. & Gergs, R. 2016. Contribution of organic toxicants to multiple stress in river ecosystems. *Freshwater Biology*, 61(12):2116-2128.

Schmid, M., Frei, D., Patrignani, A., Schlapbach, R., Frey, J.E., Remus-Emsermann, M.N. & Ahrens, C.H. 2018. Pushing the limits of de novo genome assembly for complex prokaryotic genomes harboring very long, near identical repeats. *BioRxiv*:300186.

Schnell, J., Steele, M., Bean, J., Neuspiel, M., Girard, C., Dormann, N., Pearson, C., Savoie, A., Bourbonnière, L. & Macdonald, P. 2015. A comparative analysis of insertional effects in genetically engineered plants: considerations for pre-market assessments. *Transgenic Research*, 24(1):1-17.

Schnepf, E., Crickmore, N.v., Van Rie, J., Lereclus, D., Baum, J., Feitelson, J., Zeigler, D. & Dean, D. 1998. *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiology and Molecular Biology Reviews*, 62(3):775-806.

Schouten, H.J., Krens, F.A. & Jacobsen, E. 2006. Cisgenic plants are similar to traditionally bred plants. *EMBO Reports*, 7(8):750-753.

Schweitzer, J.A., Bailey, J.K., Rehill, B.J., Martinsen, G.D., Hart, S.C., Lindroth, R.L., Keim, P. & Whitham, T.G. 2004. Genetically based trait in a dominant tree affects ecosystem processes. *Ecology Letters*, 7(2):127-134.

Sigsgaard, E.E., Carl, H., Møller, P.R. & Thomsen, P.F. 2014. Monitoring the near-extinct European weather loach in Denmark based on environmental DNA from water samples. *Biological Conservation*, 183:46-52.

Sih, A., Bell, A.M. & Kerby, J.L. 2004. Two stressors are far deadlier than one. *Trends in Ecology & Evolution*, 19(6):274-276.

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V. & Zdobnov, E.M. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31(19):3210-3212.

Sissener, N.H., Johannessen, L.E., Hevrøy, E.M., Wiik-Nielsen, C.R., Berdal, K.G., Nordgreen, A. & Hemre, G.-I. 2010. Zebrafish (*Danio rerio*) as a model for investigating the safety of GM feed ingredients (soya and maize); performance, stress response and uptake of dietary DNA sequences. *British Journal of Nutrition*, 103(01):3-15.

Smets, G., Alcalde, E., Andres, D., Carron, D., Delzenne, P., Heise, A., Legris, G., Parrilla, M.M., Verhaert, J. & Wandelt, C. 2014. The use of existing environmental networks for the post-market monitoring of GM crop cultivation in the EU. *Environmental Science: Processes & Impacts*, 16(7):1754-1763.

Smit, E., Leeflang, P., Glandorf, B., van Elsas, J.D. & Wernars, K. 1999. Analysis of fungal diversity in the wheat rhizosphere by sequencing of cloned PCR-amplified genes encoding 18S rRNA and temperature gradient gel electrophoresis. *Applied and Environmental Microbiology*, 65(6):2614-2621.

Song, W., Steensen, K. & Thomas, T. 2017. HgtSIM: a simulator for horizontal gene transfer (HGT) in microbial communities. *PeerJ*, 5:e4015.

South Africa. 1997. Numbered Acts. 1997. Genetically Modified Organisms Act [No. 15 of 1997]. Government Gazette, 383 (18029): 23 May.

South Africa. 2004. Numbered Acts. 2004. National Environmental Management: Biodiversity Act [Act 10 of 2004]. Government Gazette, 467 (26436): 7 June.

South Africa. 2006. Numbered Acts. 2006. Genetically Modified Organisms Amendment Act [No. 23 of 2006]. Government Gazette, 501 (29803): 17 April.

Spencer, S.J., Tamminen, M.V., Preheim, S.P., Guo, M.T., Briggs, A.W., Brito, I.L., Weitz, D.A., Pitkänen, L.K., Vigneault, F. & Virta, M.P. 2016. Massively parallel sequencing of single cells by epicPCR links functional genes with phylogenetic markers. *The ISME Journal*, 10(2):427.

Stehle, S. & Schulz, R. 2015. Agricultural insecticides threaten surface waters at the global scale. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 112(18):5750-5755.

Stendera, S., Adrian, R., Bonada, N., Cañedo-Argüelles, M., Hugueny, B., Januschke, K., Pletterbauer, F. & Hering, D. 2012. Drivers and stressors of freshwater biodiversity patterns across different ecosystems and scales: a review. *Hydrobiologia*, 696(1):1-28.

Stillwell, L. & Silvanovich, A. 2007. Assessment of Cry1A. 105, Cry2Ab2, Cry3Bb1 and CP4 EPSPS protein levels in the combined corn product MON 89034 x TC1507 x MON 88017 x DAS-59122-7 produced in US field trials during 2006. Monsanto Company, St Louis, Missouri. (Monsanto Study Report MSL0021070). https://www.testbiotech.org/sites/default/files/Stillwell_Silvanowich_2008.pdf. Date of access: 11 November 2018.

Stotzky, G. 2005. Persistence and biological activity in soil of the insecticidal proteins from *Bacillus thuringiensis*, especially from transgenic plants. *Plant and Soil*, 266(1-2):77-89.

Strain, K.E. & Lydy, M.J. 2015. The fate and transport of the Cry1Ab protein in an agricultural field and laboratory aquatic microcosms. *Chemosphere*, 132:94-100.

Strain, K.E., Whiting, S.A. & Lydy, M.J. 2014. Laboratory and field validation of a Cry1Ab protein quantitation method for water. *Talanta*, 128:109-116.

Strickler, K.M., Fremier, A.K. & Goldberg, C.S. 2015. Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. *Biological Conservation*, 183:85-92.

Sun, L., Cai, H., Xu, W., Hu, Y. & Lin, Z. 2002. CaMV 35S promoter directs β-glucuronidase expression in *Ganoderma lucidum* and *Pleurotus citrinopileatus*. *Molecular Biotechnology*, 20(3):239.

Surgan, M., Condon, M. & Cox, C. 2010. Pesticide risk indicators: unidentified inert ingredients compromise their integrity and utility. *Environmental Management*, 45(4):834-841.

Swan, C.M., Jensen, P.D., Dively, G.P. & Lamp, W.O. 2009. Processing of transgenic crop residues in stream ecosystems. *Journal of Applied Ecology*, 46(6):1304-1313.

Swingle, B., Markel, E., Costantino, N., Bubunenko, M.G., Cartinhour, S. & Court, D.L. 2010. Oligonucleotide recombination in Gram-negative bacteria. *Molecular Microbiology*, 75(1):138-148.

Swingle, B.M. 2013. Oligonucleotide Recombination Enabled Site-Specific Mutagenesis in Bacteria. (*In*: Samuelson, J.C., ed. Enzyme Engineering: Methods and Protocols. **Series** Methods in Molecular Biology, vol 978, Totowa, NJ: Humana Press. p. 127-132).

Syvanen, M. 2012. Evolutionary implications of horizontal gene transfer. *Annual Review of Genetics*, 46:341-358.

Tabashnik, B.E. 2015. ABCs of insect resistance to Bt. *PLoS Genetics*, 11(11):e1005646.

Tabashnik, B.E., Van Rensburg, J.B.J. & Carrière, Y. 2009. Field-evolved insect resistance to Bt crops: definition, theory, and data. *Journal of Economic Entomology*, 102(6):2011-2025.

Tabashnik, B.E., Zhang, M., Fabrick, J.A., Wu, Y., Gao, M., Huang, F., Wei, J., Zhang, J., Yelich, A. & Unnithan, G.C. 2015. Dual mode of action of Bt proteins: protoxin efficacy against resistant insects. *Scientific Reports*, 5:15107, doi: 10.1038/srep15107..

Tank, J.L., Rosi-Marshall, E.J., Royer, T.V., Whiles, M.R., Griffiths, N.A., Frauendorf, T.C. & Treering, D.J. 2010. Occurrence of maize detritus and a transgenic insecticidal protein (Cry1Ab) within the stream network of an agricultural landscape. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 107(41):17645-17650.

Tarr, S. 2004. Yeasts from Lesotho: their classification and possible applications. Bloemfontein: University of the Free State. (Doctoral thesis).

Tattini, L., D'Aurizio, R. & Magi, A. 2015. Detection of genomic structural variants from next-generation sequencing data. *Frontiers in Bioengineering and Biotechnology*, 3:92.

Taverniers, I., Papazova, N., Bertheau, Y., De Loose, M. & Holst-Jensen, A. 2008. Gene stacking in transgenic plants: towards compliance between definitions, terminology, and detection within the EU regulatory framework. *Environmental Biosafety Research*, 7(4):197-218.

Thagwana, M.S. 2010. Trends in women's participation in agriculture at Tshiombo Irrigation Scheme, Limpopo province. Cape Town: University of the Western Cape. (Master's dissertation).

Then, C. 2010. Risk assessment of toxins derived from *Bacillus thuringiensis*—synergism, efficacy, and selectivity. *Environmental Science and Pollution Research*, 17(3):791-797.

Thomas, C.M. & Nielsen, K.M. 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature Reviews Microbiology*, 3(9):711-721.

Thomsen, P.F. & Willerslev, E. 2014. Environmental DNA–An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, 183:4-18.

Thorvaldsdóttir, H., Robinson, J.T. & Mesirov, J.P. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, 14(2):178-192.

Tomer, M., Boomer, K., Porter, S., Gelder, B., James, D. & Mclellan, E. 2015. Agricultural conservation planning framework: 2. Classification of riparian buffer design types with application to assess and map stream corridors. *Journal of Environmental Quality*, 44(3):768-779.

Torres, J.B. & Ruberson, J.R. 2008. Interactions of *Bacillus thuringiensis* Cry1Ac toxin in genetically engineered cotton with predatory heteropterans. *Transgenic Research*, 17(3):345-354.

Traavik, T., Nielsen, K.M. & Quist, D. 2009. Genetically engineered cells and organisms: substantially equivalent or different? TWN Biotechnology & Biosafety Series 9. Penang: Third World Network.

Trappe, K., Emde, A.-K., Ehrlich, H.-C. & Reinert, K. 2014. Gustaf: detecting and correctly classifying SVs in the NGS twilight zone. *Bioinformatics*, 30(24):3484-3490.

Trappe, K., Marschall, T. & Renard, B.Y. 2016. Detecting horizontal gene transfer by mapping sequencing reads across species boundaries. *BioRxiv*:039495.

Turner, C.R., Uy, K.L. & Everhart, R.C. 2014. Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biological Conservation*, 183:93-102.

USDA-ERS. (United States Department of Agriculture Economic Research Service). 2015. Adoption of Genetically Engineered Crops in the U.S. https://www.ers.usda.gov/data-products/adoption-of-genetically-engineered-crops-in-the-us.aspx Date of access: 10 July 2015.

Vachon, V., Laprade, R. & Schwartz, J.-L. 2012. Current models of the mode of action of *Bacillus thuringiensis* insecticidal crystal proteins: a critical review. *Journal of Invertebrate Pathology*, 111(1):1-12.

Van Averbeke, W. 2012. Performance of smallholder irrigation schemes in the Vhembe District of South Africa. *(In* Kumar, M., *ed.* Problems, Perspectives and Challenges of Agricultural Water Management. Rijecka: InTech. p. 413-438).

Van den Berg, J. 2013. Evolution in action: field-evolved resistance of African stem borer to Bt maize. *Outlooks on Pest Management*, 24(5):236-239.

van Frankenhuyzen, K. 2013. Cross-order and cross-phylum activity of Bacillus thuringiensis pesticidal proteins. *Journal of Invertebrate Pathology*, 114(1):76-85.

van Pijkeren, J.-P. & Britton, R.A. 2012. High efficiency recombineering in lactic acid bacteria. *Nucleic Acids Research*, 40(10):e76-e76.

Venter, H.J. & Bezuidenhout, C.C. 2016. DNA-based identification of aquatic invertebrates useful in the South African context? *South African Journal of Science*, 112(5/6):4 Pages.

Vernikos, G.S. & Parkhill, J. 2006. Interpolated variable order motifs for identification of horizontally acquired DNA: revisiting the *Salmonella* pathogenicity islands. *Bioinformatics*, 22(18):2196-2203.

Viljoen, C. 2013. Letter to the Editor. *Food and Chemical Toxicology*, 59:809-810.

Wandeler, H., Bahylova, J. & Nentwig, W. 2002. Consumption of two Bt and six non-Bt corn varieties by the woodlouse Porcellio scaber. *Basic and Applied Ecology*, 3(4):357-365.

Wang, J.-M., Chen, X.-P., Liang, Y.-Y., Zhu, H.-J., Ding, J.-T. & Peng, Y.-F. 2014a. Influence of transgenic rice expressing a fused Cry1Ab/1Ac protein on frogs in paddy fields. *Ecotoxicology*, 23(9):1619-1628.

Wang, J., Chen, X., Li, Y., Su, C., Ding, J. & Peng, Y. 2014b. Green algae (*Chlorella pyrenoidosa*) adsorbs *Bacillus thurigiensis* (Bt) toxin, Cry1Ca insecticidal protein, without an effect on growth. *Ecotoxicology and Environmental safety*, 106:6-10.

Wang, J., Chen, X., Li, Y., Zhu, H., Ding, J. & Peng, Y. 2014c. Effect of straw leachates from Cry1Ca-expressing transgenic rice on the growth of *Chlorella pyrenoidosa*. *Environmental Toxicology and Chemistry*, 33(5):1156-1162.

Wang, Y., Hu, H., Huang, J., Li, J., Liu, B. & Zhang, G. 2013a. Determination of the movement and persistence of Cry1Ab/1Ac protein released from Bt transgenic rice under field and hydroponic conditions. *Soil Biology and Biochemistry*, 58:107-114.

Wang, Y., Huang, J., Hu, H., Li, J., Liu, B. & Zhang, G. 2013b. Field and laboratory studies on the impact of two Bt rice lines expressing a fusion protein Cry1Ab/1Ac on aquatic organisms. *Ecotoxicology and Environmental safety*, 92:87-93.

Ward, J.V. 1998. Riverine landscapes: biodiversity patterns, disturbance regimes, and aquatic conservation. *Biological Conservation*, 83(3):269-278.

Waterhouse, R.M., Seppey, M., Simão, F.A., Manni, M., Ioannidis, P., Klioutchnikov, G., Kriventseva, E.V. & Zdobnov, E.M. 2018. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Molecular Biology and Evolution*, 35(3):543-548.

Wellington, E.M.H., Boxall, A.B.A., Cross, P., Feil, E.J., Gaze, W.H., Hawkey, P.M., Johnson-Rollings, A.S., Jones, D.L., Lee, N.M., Otten, W., Thomas, C.M. & Williams, A.P. 2013. The role of the natural environment in the emergence of antibiotic resistance in Gram-negative bacteria. *The Lancet Infectious Diseases*, 13(2):155-165.

Whitham, T.G., Bailey, J.K., Schweitzer, J.A., Shuster, S.M., Bangert, R.K., LeRoy, C.J., Lonsdorf, E.V., Allan, G.J., DiFazio, S.P. & Potts, B.M. 2006. A framework for community and ecosystem genetics: from genes to ecosystems. *Nature Reviews Genetics*, 7(7):510-523.

Whiting, S.A., Strain, K.E., Campbell, L.A., Young, B.G. & Lydy, M.J. 2014. A multi-year field study to evaluate the environmental fate and agronomic effects of insecticide mixtures. *Science of the Total Environment*, 497:534-542.

Wolfenbarger, L.L., Naranjo, S.E., Lundgren, J.G., Bitzer, R.J. & Watrud, L.S. 2008. Bt crop effects on functional guilds of non-target arthropods: a meta-analysis. *PLoS One*, 3(5):e2118.

Wolt, J.D. & Peterson, R.K.D. 2010. Prospective formulation of environmental risk assessments: probabilistic screening for Cry1A(b) maize risk to aquatic insects. *Ecotoxicology and Environmental Safety*, 73(6):1182-1188.

Wu, W.-X., Ye, Q.-F. & Min, H. 2004a. Effect of straws from Bt-transgenic rice on selected biological activities in water-flooded soil. *European Journal of Soil Biology*, 40(1):15-22.

Wu, W.-X., Ye, Q.-F., Min, H., Duan, X.-J. & Jin, W.-M. 2004b. Bt-transgenic rice straw affects the culturable microbiota and dehydrogenase and phosphatase activities in a flooded paddy soil. *Soil Biology and Biochemistry*, 36(2):289-295.

Xiao, M., Dong, S., Li, Z., Tang, X., Chen, Y., Yang, S., Wu, C., Ouyang, D., Fang, C. & Song, Z. 2015. Effects of water management practices on residue decomposition and degradation of Cry1Ac protein from crop-wild Bt rice hybrids and parental lines during winter fallow season. *Ecotoxicology and Environmental Safety*, 122:275-289.

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.-H., Whitman, W.B., Euzéby, J., Amann, R. & Rosselló-Móra, R. 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature Reviews Microbiology*, 12(9):635.

Yiallouros, M., Storch, V. & Becker, N. 1999. Impact of *Bacillus thuringiensis* var. *israelensis* on larvae of *Chironomus thummi thummi* and *Psectrocladius psilopterus* (Diptera: Chironomidae). *Journal of Invertebrate Pathology*, 74(1):39-47.

Yu, H., Li, Y., Li, X. & Wu, K. 2014a. Arthropod abundance and diversity in transgenic Bt soybean. *Environmental Entomology*, 43(4):1124-1134.

Yu, H., Romeis, J., Li, Y., Li, X. & Wu, K. 2014b. Acquisition of Cry1Ac protein by non-target arthropods in Bt soybean fields. *PLoS One*, 9(8):e103973.

Zhu, H.-J., Chen, Y., Li, Y.-H., Wang, J.-M., Ding, J.-T., Chen, X.-P. & Peng, Y.-F. 2015. A 90 day safety assessment of genetically modified rice expressing Cry1Ab/1Ac protein using an aquatic animal model. *Journal of Agricultural and Food Chemistry*, 63(14):3627-3633.

Züghart, W., Beismann, H. & Schröder, W. 2013. Tools for a scientifically rigorous and efficient monitoring of genetically modified organisms (GMOs)–VDI Guidelines to ensure high quality of GMO-monitoring data. *BioRisk*, 8:3-13.

Züghart, W., Benzler, A., Berhorn, F., Sukopp, U. & Graef, F. 2007. Determining indicators, methods and sites for monitoring potential adverse effects of genetically modified plants to the environment: the legal and conceptional framework for implementation. *Euphytica*, 164(3):845-852.

Zupan, J.R. & Zambryski, P. 1995. Transfer of T-DNA from *Agrobacterium* to the plant cell. *Plant Physiology*, 107(4):1041-1047.